7506–7521 *Nucleic Acids Research, 2018, Vol. 46, No. 15* *Published online 16 July 2018*
doi: 10.1093/nar/gky631

Incorporating uracil and 5-halouracils into short peptide nucleic acids for enhanced recognition of A–U pairs in dsRNAs

Kiran M. Patil[1,†], Desiree-Faye Kaixin Toh[1,†], Zhen Yuan[1,†], Zhenyu Meng[1,†], Zhiyu Shu[1], Haiping Zhang[2], Alan Ann Lerk Ong[1], Manchugondanahalli S. Krishna[1], Lanyuan Lu[2], Yunpeng Lu[1,*] and Gang Chen[1,*]

[1]Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371 and [2]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551

Received January 05, 2018; Revised June 09, 2018; Editorial Decision June 29, 2018; Accepted July 03, 2018

ABSTRACT

Double-stranded RNA (dsRNA) structures form triplexes and RNA-protein complexes through binding to single-stranded RNA (ssRNA) regions and proteins, respectively, for diverse biological functions. Hence, targeting dsRNAs through major-groove triplex formation is a promising strategy for the development of chemical probes and potential therapeutics. Short (e.g., 6–10 mer) chemically-modified Peptide Nucleic Acids (PNAs) have been developed that bind to dsRNAs sequence specifically at physiological conditions. For example, a PNA incorporating a modified base thio-pseudoisocytosine (L) has an enhanced recognition of a G–C pair in an RNA duplex through major-groove L·G–C base triple formation at physiological pH, with reduced pH dependence as observed for C^+·G–C base triple formation. Currently, an unmodified T base is often incorporated into PNAs to recognize a Watson–Crick A–U pair through major-groove T·A–U base triple formation. A substitution of the 5-methyl group in T by hydrogen and halogen atoms (F, Cl, Br, and I) causes a decrease of the pK_a of N3 nitrogen atom, which may result in improved hydrogen bonding in addition to enhanced base stacking interactions. Here, we synthesized a series of PNAs incorporating uracil and halouracils, followed by binding studies by non-denaturing polyacrylamide gel electrophoresis, circular dichroism, and thermal melting. Our results suggest that replacing T with uracil and halouracils may enhance the recognition of an A–U pair by PNA·RNA_2 triplex formation in a sequence-dependent manner, underscoring the importance of local stacking interactions. Incorporating bromouracils and chlorouracils into a PNA results in a significantly reduced pH dependence of triplex formation even for PNAs containing C bases, likely due to an upshift of the apparent pK_a of N3 atoms of C bases. Thus, halogenation and other chemical modifications may be utilized to enhance hydrogen bonding of the adjacent base triples and thus triplex formation. Furthermore, our experimental and computational modelling data suggest that PNA·RNA_2 triplexes may be stabilized by incorporating a ^{Br}UL step but not an L^{Br}U step, in dsRNA-binding PNAs.

INTRODUCTION

RNAs often form complex secondary and tertiary structures, and in turn, interact with proteins and metabolites for diverse regulatory and catalytic functions (1–3). Double-stranded RNAs (dsRNAs) are essential components of RNA secondary structures (4–7). RNA triplex structures, formed between dsRNA regions and single-stranded RNA (ssRNA) regions, are emerging as an important tertiary structure motif (8–13). It is thus of great potential to develop ligands that can selectively and sequence-specifically recognize viral and cellular dsRNA regions as chemical probes and therapeutics.

Targeting dsRNAs through sequence-specific major-groove triplex formation is a promising strategy and numerous chemical modifications have been developed to enhance

*To whom correspondence should be addressed. Tel: +65 6592 2549; Fax: +65 6791 1961; Email: RNACHEN@ntu.edu.sg
Correspondence may also be addressed to Yunpeng Lu. Email: YPLu@ntu.edu.sg
†The authors wish it to be known that, in their opinion, the first four authors should be regarded as Joint First Authors.

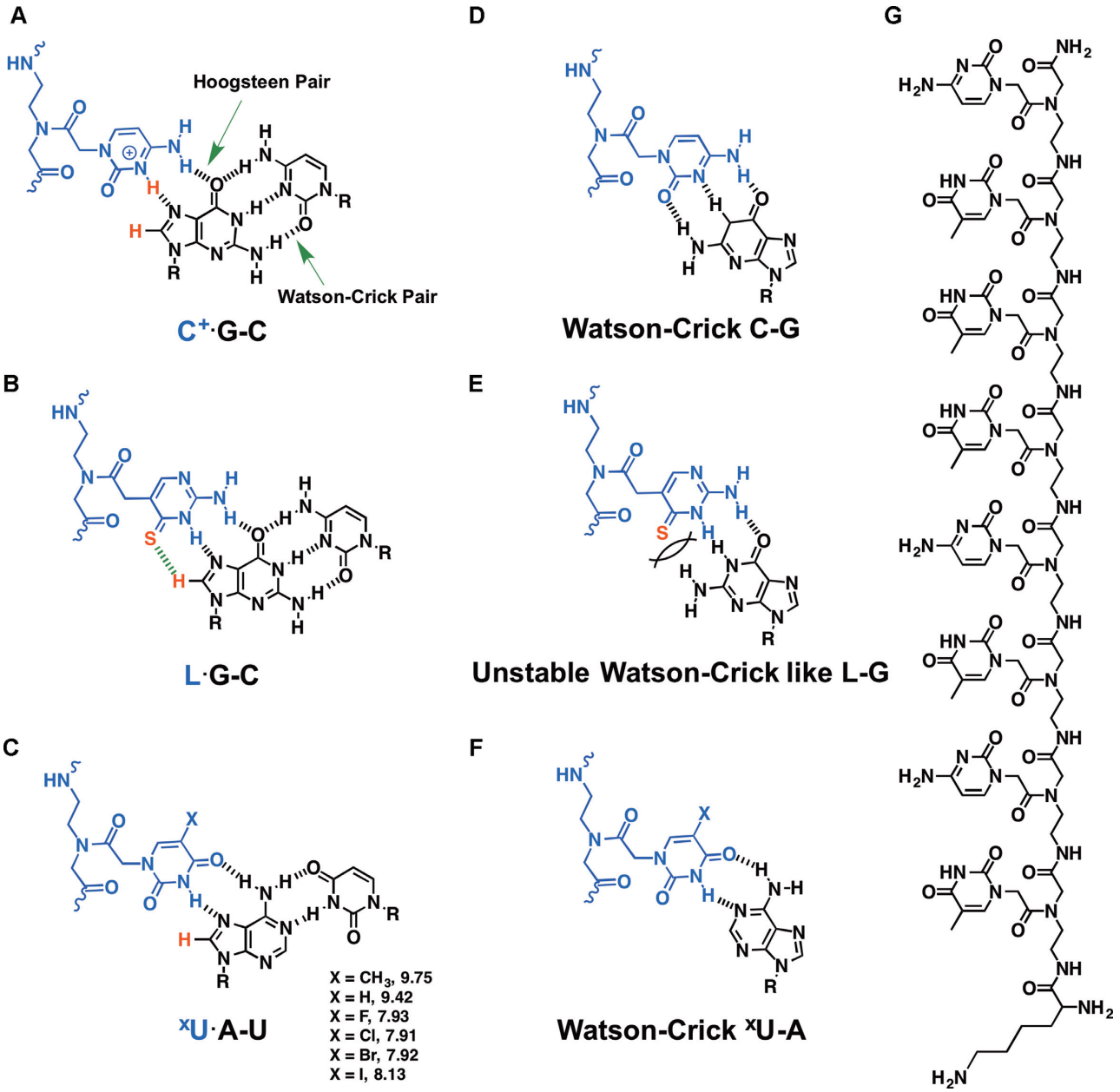


Figure 1. Chemical structures of base triples, base pairs, and PNA P1. The letter R represents the sugar-phosphate backbone of RNA. Hydrogen bonding interactions are indicated by black dashed lines. (**A–C**) Base triples of C^+·G–C, L·G–C, and ^{X}U·A–U. The green dashed line in panel B indicates enhanced van der Waals interaction between the atoms shown in red. The values shown in panel C are the pK_a values of the N3 atoms measured for the free bases (43). (**D–F**) Structures of Watson–Crick C–G, unstable Watson–Crick-like L–G due to a steric clash as indicated, and Watson–Crick ^{X}U–A pairs. (**G**) Chemical structure of PNA P1 (NH_2-Lys-TCTCTTTC-$CONH_2$). The N-terminal lysine residue has an L configuration.

the binding affinity and specificity of Triplex-Forming Oligonucleotides (TFOs) (14–19). Peptide Nucleic Acids (PNAs) (20) typically have a neutral peptide-like backbone (Figure 1) and are advantageous with strong resistance against nucleases and proteases, and enhanced binding to natural nucleic acids. PNAs were originally designed to bind to double-stranded DNAs (dsDNAs) to form major-groove PNA·dsDNA (PNA·DNA–DNA, with the '·' and '–' representing Hoogsteen and Watson–Crick pairs, respectively, see Figure 1A–C for the base pairing interactions) triplexes. However, PNAs have been shown to form Watson–Crick duplexes with complementary DNA, RNA, or PNA in an antiparallel (with the C-terminus of a PNA aligned with 5′ end of DNA/RNA) or parallel (with the N-terminus of a PNA aligned with 5′ end of DNA/RNA) orientation (20–22). In addition, PNAs can be involved in the formation of parallel major-groove triplexes with various compositions including PNA·DNA–PNA, PNA·RNA–PNA, PNA·DNA–DNA, and PNA·RNA–RNA (13,23–36). It is interesting to note that short PNAs show selective binding to dsRNAs over dsDNAs, suggesting PNAs' great potential as dsRNA-specific binders (13,28–37).

To recognize a Watson–Crick G–C base pair in an RNA duplex, one may design a major-groove C^+·G–C triple (Figure 1A). However, to form a Hoogsteen C^+·G pair often requires a relatively low pH (<6) for the protonation of the N3 atom of the C base (28,33,38). Modified PNA bases such as thio-pseudoisocytosine (L, Figure 1B) and 2-aminopyridine (M) have close-to-neutral pK_a values for the corresponding nitrogen atom and thus allow enhanced recognition of a G–C pair in RNA duplexes through L·G–C and M·G–C base triple formation, respectively, at physiological pH (29,33,39). In addition, short modified PNAs incorporating L bases show selective binding toward dsRNAs over ssRNAs (Figure 1B, D, E) (33,36), suggesting that L-modified PNAs have a great potential for probing and targeting dsRNAs.

To recognize an inverted Watson–Crick C–G pair in an RNA duplex, a guanidine-modified PNA monomer (Q monomer) (34) or other monomers may be incorporated into PNAs through PNA·RNA–RNA triplex formation. Similar to L, Q is a cytosine derivative favoring Q·C–G triple formation, while destabilizing Watson–Crick-like Q–G pair formation (34–37). To recognize an inverted Watson–Crick U–A pair in an RNA duplex, a PNA strand incorporating an E (3-oxo-2,3-dihydropyridazine) base has been used through E·U–A base triple formation (39–41).

An unmodified T base has been incorporated into PNAs to recognize a Watson–Crick A–U pair through major-groove T·A–U base triple formation (Figure 1C) (13,20,28,38,39,42). Enhancing the recognition of Watson–Crick A–U base pairs is crucial due to the relatively high abundance of A–U pairs in dsRNAs. A substitution of the 5-methyl group in T by hydrogen or halogen atoms (F, Cl, Br, and I, designated as ^{X}U in Figure 1C) results in a decrease of the pK_a of N3 nitrogen atom (43) and thus may strengthen the hydrogen bond formed involving the N3 atom. Thus, ^{X}U substitution of T in a PNA is expected to enhance Hoogsteen and Watson–Crick hydrogen bonding interactions (Figure 1C, F) (44,45). The relatively large halogen atoms may also contribute to base stacking interactions (44). Here, we report the binding properties of PNAs incorporating uracil and 5-halouracils (5-fluorouracil (^{F}U), 5-chlorouracil (^{Cl}U), 5-bromouracil (^{Br}U) and 5-iodouracil (^{I}U) for the recognition of A–U base pairs in dsRNAs. We also tested if incorporating adjacent L and ^{X}U bases can cooperatively enhance the binding affinities of dsRNA-binding PNAs for targeting microRNA-198 (miRNA-198 or miR-198) hairpin precursor (46,47) and HIV-1 −1 ribosomal frameshift inducing mRNA hairpin (48–50).

MATERIALS AND METHODS

General methods

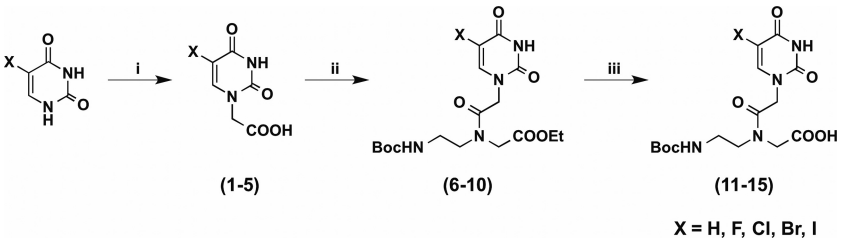
Reagents and solvents used were obtained from commercial sources and used without further purification. All organic reactions were monitored with the use of thin layer chromatography (TLC) using aluminum sheets silica gel 60 F254 (Merck). Compounds were purified by flash column chromatography using silica gel with ethyl acetate/petroleum ether mixture as the eluting solvent. All ^{1}H and ^{13}C NMR spectra were obtained at room temperature on 300, 400 (100 MHz, ^{13}C) or 500 MHz Bruker spectrometer. The chemical shifts (δ) are shown in parts per million (ppm). The residual solvent peaks were used as references for the ^{1}H (chloroform-d: 7.26; dimethyl sulfoxide-d_6: 2.50) and ^{13}C (chloroform-d: 77.0; dimethyl sulfoxide-d_6: 39.5) NMR spectra. The mass spectra of the compounds were obtained via liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI) and high-resolution mass spectrometry (electron ionization) (HRMS-EI). Reverse-phase high performance liquid chromatography (RP-HPLC) purified RNA and DNA oligonucleotides were purchased from Sigma-Aldrich Singapore.

Synthesis of PNA monomers

The detailed synthesis procedures for the PNA monomers are shown in Supplementary Data (Supplementary Figures S1–S15). The PNA monomers were synthesized based on the previously reported methods (Scheme 1) (51–53). Commercially available uracil or 5-halouracil was reacted with chloro- or bromo-acetic acid in the presence of potassium hydroxide to yield respective compounds **1**–**5**. Compounds **1–5** were coupled with PNA backbone ethyl *N*-(2-Boc-aminoethyl)glycinate in the presence of Hexafluorophosphate Benzotriazole Tetramethyl Uronium (HBTU) and *N*-methylmorpholine (NMM), in Dimethylformamide (DMF) to give respective compounds **6–10**. Finally, the ethyl ester group was hydrolyzed by using aq. LiOH in water–THF solvent. The formed lithium salt of carboxylic acid was neutralized and the mixture was acidified using diluted HCl for the desired PNA monomers **11–15**.

Synthesis of PNA oligomers

The PNA thymine (T) and cytosine (C) monomers containing the standard aminoethyl glycine backbone were purchased from ASM Research Chemicals. The PNA monomer L was synthesized following our previously reported method (33). PNA oligomers were synthesized manually using Boc chemistry via a Solid-Phase Peptide Synthesis (SPPS) protocol (37). 4-Methylbenzhydrylamine hydrochloride (MBHA·HCl) polystyrene resins were used. The loading value used for the synthesis of the oligomers was 0.3 mmol/g and acetic anhydride was used as the capping reagent. Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and *N,N*-diisopropylethylamine (DIPEA) were used as the coupling reagent. The oligomerization of PNA was monitored by Kaiser test. Cleavage of the PNA oligomers was done using trifluoroacetic acid (TFA) and trifluoromethanesulfonic acid (TFMSA) method, after which the oligomers were precipitated with diethyl ether, dissolved in deionized water and purified by RP-HPLC using water–CH_3CN–0.1% TFA as the mobile phase. Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) analysis was used to characterize the oligomers (Supplementary Table S1 and Figure S16), with the use of α-cyano-4-hydroxycinnamic acid (CHCA) as the



Scheme 1. Reagentsand conditions for the syntheses of uracil and modified uracil PNA monomers: (i) Bromoacetic acid, KOH, water, 40°C, 1 h, 65–70%. (ii) Ethyl *N*-(2-Boc-aminoethyl)glycinate, Hexafluorophosphate Benzotriazole Tetramethyl Uronium (HBTU), *N*-methylmorpholine (NMM), anhydrous dimethylformamide (DMF), room temperature (rt), 3–5 h, 60–70%. (iii) 1 M aq. LiOH, tetrahydrofuran (THF), rt, 1 h, 1 M HCl, 0°C, 78–80%.

sample crystallization matrix. The extinction coefficients of ^{X}U were assumed to be the same as that of T in calculating the extinction coefficients of the PNA oligomers (37).

Non-denaturing polyacrylamide gel electrophoresis

Non-denaturing (12%) polyacrylamide gel electrophoresis (PAGE) experiments (37) were conducted with an incubation buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5, 7.7 or 8.0. The loading volumes for the samples containing RNA and DNA hairpins were 20 or 25 μl and 10 μl, respectively. The samples were prepared by snap cooling of the hairpin, followed by annealing with PNA oligomers by slow cooling from 65°C to room temperature and incubation at 4°C overnight. Prior to loading the samples into the wells, 35% glycerol (20% of the total volume) was added to the sample mixtures. 1× Tris–borate–EDTA (TBE) buffer, pH 8.3 was used as the running buffer for all gel experiments. The gels were run at 4°C at 250 V for 5 h. The gels were then stained with ethidium bromide for 30 min and imaged by the Typhoon Trio Variable Mode Imager.

Circular dichroism

Circular dichroism (CD) spectroscopic characterization was carried out by using a JASCO model J-1500-150 spectropolarimeter and a quartz cell (optical path length: 10 mm). The concentrations of RNA and PNA are 2 and 4 or 20 μM, respectively. RNA was snap cooled from 95°C in an incubation buffer containing 100 mM NaCl, 10 mM sodium phosphate, pH 7.0. The RNA hairpin was annealed with PNA by incubating at 65°C for 10 min followed by slow cooling to room temperature. The scanning rate was 50 nm/min and the data were averaged from six scans. The measurements were done at room temperature.

UV-absorbance-detected thermal melting

UV-absorbance-detected thermal melting experiments were conducted using the Shimadzu UV-2550 UV-Vis spectrophotometer with the use of an 8-microcell cuvette. The absorbance at 260 nm was recorded with the temperature increasing from 15 to 95°C followed by the temperature decreasing from 95 to 15°C. The temperature ramp rate is 0.5°C/min. The optical path length of the 8-microcell cuvette is 1 cm. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM NaH_2PO_4, at pH 7.5. All samples contain 5 μM RNA and 5 μM PNA in 130 μl buffer. The samples containing the ssRNA and PNA were annealed by slow cooling from 95°C to room temperature, followed by incubation at 4°C overnight. Data were normalized at high temperature and the melting temperatures were determined based on the Gaussian fits of the first derivatives of the curves.

Confocal microscopy studies

HeLa cells (1×10^5) were plated in a petri-dish in 500 μl Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS). The cells were grown at 37°C, 5% CO_2 for 24 h. The carboxyfluorescein-labeled PNAs were added to the cell cultures, followed by incubation for 24 h. A final concentration of 0.5 mg/l Hochest 33258 were added and incubated for 1 h. The cells were washed twice with PBS. Fresh OPTIMEM medium was added and the cells were immediately visualized on 63× objective of Zeiss LSM 800 confocal microscope using 405 and 488 nm lasers for Hoechst and carboxyfluorescein, respectively.

Computational modelling of PNA·RNA–RNA triplex structures

The crystal structure (PDB ID: 1PNN) of a PNA·DNA–PNA triplex (54) was used as a starting structure for the modelling of the PNA·RNA–RNA triplex structures. The backbones of the Watson–Crick PNA–DNA duplex were replaced with the RNA backbone using Discovery Studio 2016 (Dassault Systèmes, San Diego). The geometries of the base triples (C·G–C and T·A–T) without backbone attached were also built based on the crystal structure, with the C·G–C and T·A–T triples replaced with L·G–C and T·A–U triples, respectively, using Discovery Studio 2016. The new base triple structures were optimized under B3LYP/6-31G* scheme using Gaussian 09 (Gaussian, Inc., Wallingford, CT, USA). The optimized base triples were then used to replace the base triples in the original PNA·DNA–PNA crystal structure to obtain the starting structure for a PNA·RNA–RNA triplex structure containing the PNA sequence of AcNH-TLTLTTTL-$CONH_2$ with the N-terminal amine capped by an acetyl group. The RNA duplex contains 11 bp (corresponding to base pairs from A4–U29 to U14–A19, see Figure 2A) with the sequences of (5′AGAGAGAAAGU3′ and 5′ACUUUCUCUCU3′). We

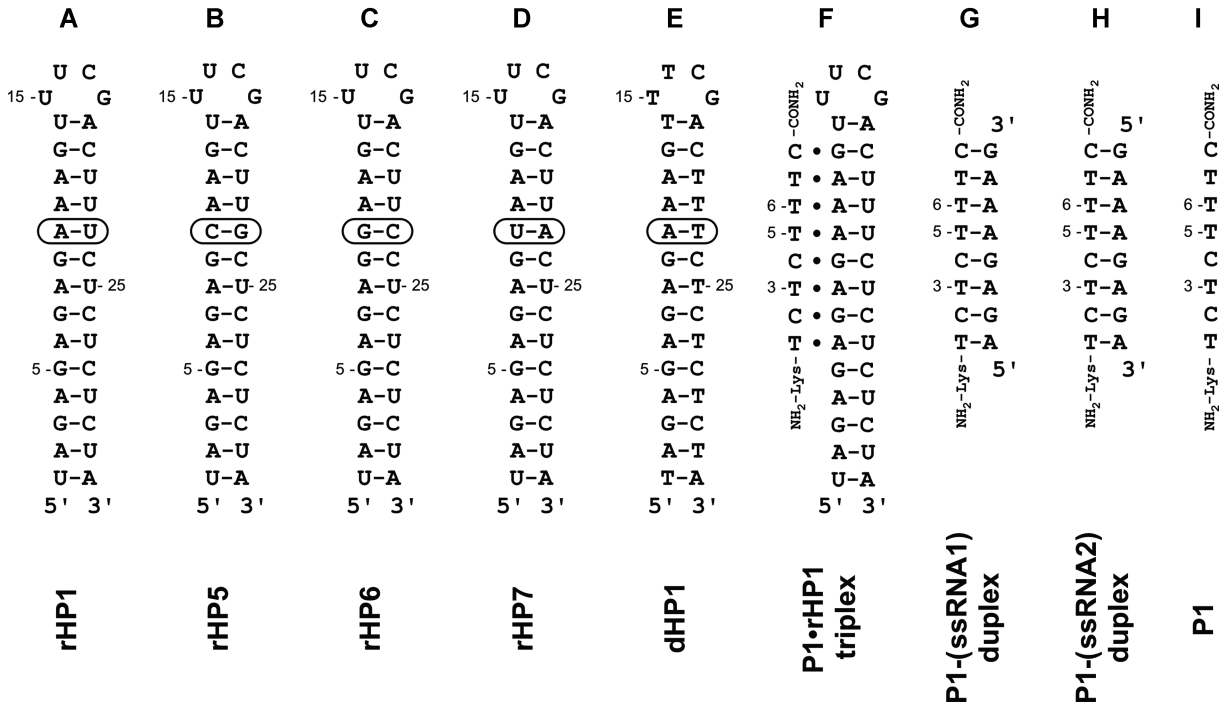


Figure 2. Sequences and structures of RNAs, a DNA and PNAs studied in this paper. (**A–D**) Model RNA hairpins rHP1, rHP5, rHP6 and rHP7. (**E**) Model DNA hairpin dHP1. (**F**) A model PNA·RNA_2 triplex formed between PNA P1 and rHP1. (**G**) Parallel PNA–RNA duplex formed between PNA P1 and ssRNA1. (**H**) Anti-parallel PNA–RNA duplex formed between PNA P1 and ssRNA2. (**I**) PNA P1. The T residues at position 3, 5 and 6 have different neighboring bases and were replaced with U or 5-halouracils in this study.

then applied energy minimization and molecular dynamics (MD) simulations to optimize the triplex structure using GROMACS 5.1 package (55). The force field used for RNA residues was Amber99bsc0 with χOL3 modification (56). The force field parameters for unmodified PNA can be retrieved from REDDB database (57). For the modified PNA residues, the atomic charges can be fit using RED server (58), and the topological parameters were adapted from Amber10 (University of California, San Francisco) Generalized Amber Force Field (GAFF) library (59). Unless otherwise noted, we applied the base–base hydrogen bond restraints throughout the simulations.

The starting PNA·RNA–RNA triplex structure was first solvated in a periodic boundary box, followed by an energy minimization, an NVT (constant temperature and volume) equilibrium, and NPT (constant temperature and pressure) equilibrium step, with the final temperature set to be 298K and pressure set to be 1 atm. To further optimize the backbone conformation, we applied a 40-ns simulation protocol (298 K→398 K→298 K→398 K→298 K), with all the base-base hydrogen bonds restrained. Thereafter, a 100-ns product MD was performed, with the base-base hydrogen bonds of the terminal A14–U19 pair (corresponding to the base pair near the hairpin loop) restrained. The clustered triplex structure was obtained from the last 10 ns of the MD simulation. Discovery Studio 2016 was used to build the triplex structures containing the other PNA sequences by replacing the corresponding PNA bases without further optimization.

RESULTS AND DISCUSSION

Substitution of T with U or halouracils in an unmodified PNA enhances binding to a dsRNA

We carried out non-denaturing PAGE to characterize the triplex formation (Supplementary Figures S17–S28). An unmodified PNA (P1, Figures 1G and 2, Table 1) has been previously shown by non-denaturing PAGE to bind to an RNA hairpin rHP1 (K_d = 5.1 μM at 200 mM NaCl, pH 7.5) (33). In this study, we used a decreased concentration of rHP1 (0.25 μM instead of 1 μM as used previously) and obtained a slightly decreased K_d (2.1 μM, Figure 3A, Table 2). The replacement of the single thymine at position 3 in P1 by uracil (U-3) results in an enhanced binding toward rHP1 (with the K_d value decreased from 2.1 to 1.2 μM at pH 7.5, Table 2, Supplementary Figures S17 and S18). Replacing two T residues at the third and fifth positions in P1 with U bases (U-3,5) further enhances the binding affinity toward rHP1 with a K_d value of 0.6 μM at pH 7.5 (Figure 3B, Table 2). The observed stabilizing effect is consistent with the fact that a hydrogen bond (formed between N3 of T and N7 of A) in the PNA·RNA Hoogsteen T·A can be enhanced by shifting the monomer N3 pK_a value from ∼9.8 for T toward neutrality for U (Figure 1C) (43,44).

It is of note that a destabilizing effect was previously observed in forming a DNA·dsDNA triplex upon DNA T to DNA U substitution in the Hoogsteen strand (44,60). The T-to-U destabilizing effect in a DNA Hoogsteen strand was attributed to the loss of the favorable hydrophobic effect and stacking involving the methyl groups of T residues in the major groove (44,60). It is likely that the favorable

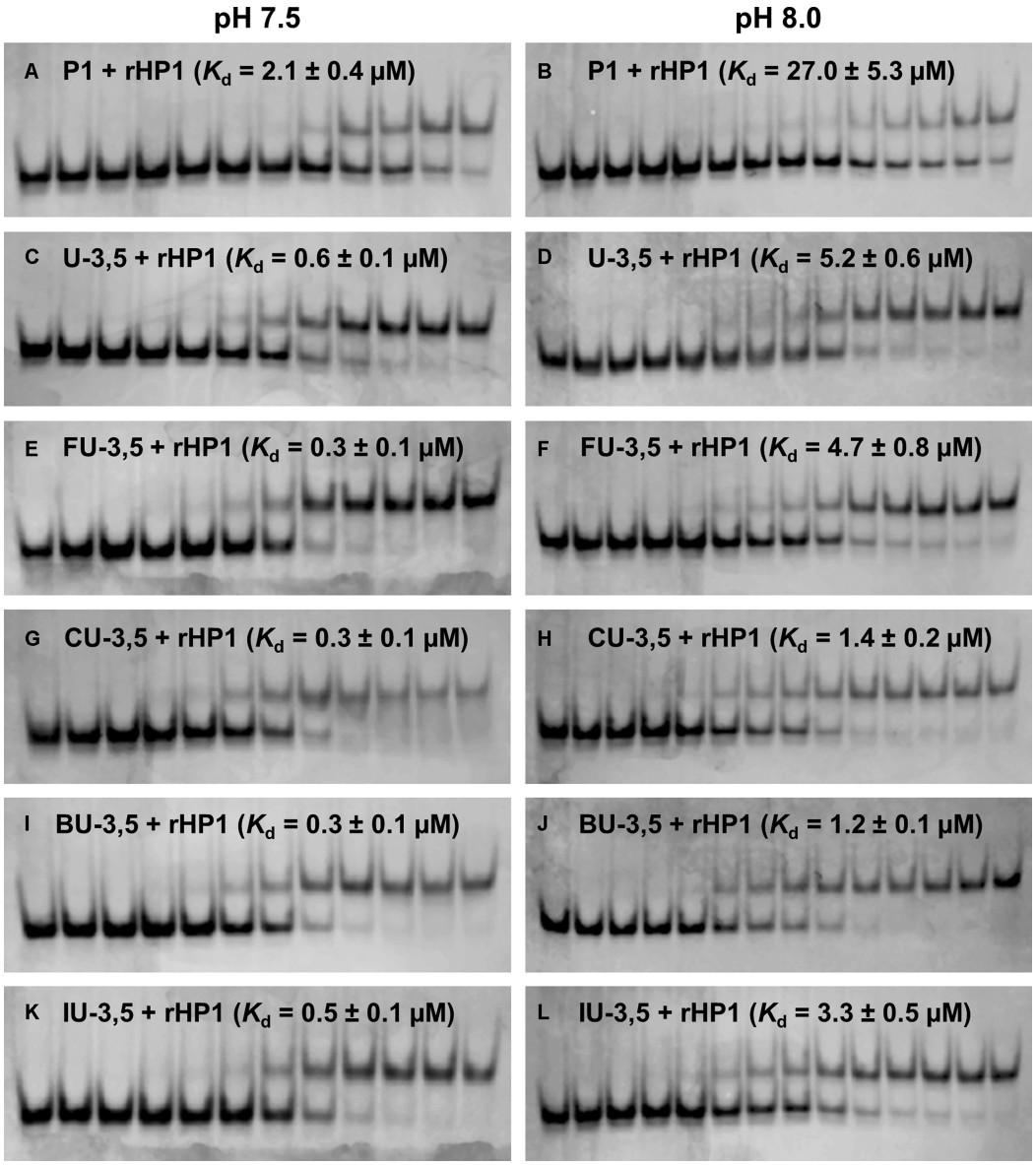


Figure 3. Non-denaturing PAGE (12wt%) study of various PNAs binding to rHP1. The gels contain a running buffer of 1× TBE, pH 8.3 and was run for 5 h at 250 V. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5 (left column) or pH 8.0 (right column). rHP1 loaded is at 0.25 μM in 25 μl. The PNA concentrations in lanes from left to right are (left column, **A**, **C**, **E**, **G**, **I**, **K**) 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 2, 4, 10 and 20 μM, respectively, and (right column, **B**, **D**, **F**, **H**, **J**, **L**) 0, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, 20, 28 and 50 μM, respectively.

stacking and hydrophobic effect of the methyl groups of T residues within a PNA strand are not as significant in PNA·dsRNA triplexes. It is also possible that the stacking interactions in a PNA·RNA_2 triplex may be different from a DNA·DNA_2 triplex. Previous studies suggest that the binding of the third strand to an RNA or DNA duplex to form a major-groove triplex induces a subtle conformational change. The resulting structure of the parental RNA or DNA duplex is different from either the A-form or B-form structure, although the base-base Watson–Crick pairing geometries remain largely the same (8,10,61–64). It is thus conceivable that the stacking interactions in a PNA·RNA_2 triplex may be different from RNA·RNA_2 or DNA·DNA_2 triplex.

The 5-halouracil substitution in PNAs may result in enhanced base stacking, in addition to strengthened hydrogen bonding interaction. We tested the effects of single halouracil substitutions at varied positions (positions 3, 5, and 6, Figure 2F, Table 1) of the unmodified 8-mer PNA P1 on PNA·RNA_2 triplex formation. The T residues of PNA P1 at positions 3, 5 and 6 are flanked by two cytosine residues, cytosine and thymine, and two thymine residues, respectively. Among all the singly-modified PNAs at position 3, BU-3 (with ^{Br}U modification) shows the largest enhancement in binding to rHP1 (K_d = 0.4 μM at pH 7.5, Table 2, Supplementary Figures S17 and S18). It is probable that the bromine atom in a ^{Br}U base may have optimized electronegativity and atom size, for enhanced Hoog-

Table 1. Sequences of the PNAs studied in this paper

PNA	Sequence
P1	NH_2-Lys-TCTCTTTC-$CONH_2$
U-3	NH_2-Lys-TC**U**CTTTC-$CONH_2$
U-3,5	NH_2-Lys-TC**U**C**U**TTC-$CONH_2$
FU-3	NH_2-Lys-TC^{F}UCTTTC-$CONH_2$
FU-5	NH_2-Lys-TCTC^{F}UTTC-$CONH_2$
FU-6	NH_2-Lys-TCTCT^{F}UTC-$CONH_2$
FU-3,5	NH_2-Lys-TC^{F}UC^{F}UTTC-$CONH_2$
CU-3	NH_2-Lys-TC^{Cl}UCTTTC-$CONH_2$
CU-3,5	NH_2-Lys-TC^{Cl}UC^{Cl}UTTC-$CONH_2$
BU-3	NH_2-Lys-TC^{Br}UCTTTC-$CONH_2$
BU-5	NH_2-Lys-TCTC^{Br}UTTC-$CONH_2$
BU-6	NH_2-Lys-TCTCT^{Br}UTC-$CONH_2$
BU-3,5	NH_2-Lys-TC^{Br}UC^{Br}UTTC-$CONH_2$
IU-3	NH_2-Lys-TC^{I}UCTTTC-$CONH_2$
IU-3,5	NH_2-Lys-TC^{I}UC^{I}UTTC-$CONH_2$
miR-C,T	NH_2-Lys-CCCTCT-$CONH_2$
miR-C,BU	NH_2-Lys-CCC^{Br}UC^{Br}U-$CONH_2$
miR-L,T	NH_2-Lys-LLLTLT-$CONH_2$
miR-L,BU	NH_2-Lys-LLL^{Br}UL^{Br}U-$CONH_2$
HIV-T,T	NH_2-Lys-LLTTLL-$CONH_2$
HIV-BU,BU	NH_2-Lys-LL^{Br}U^{Br}ULL-$CONH_2$
HIV-FU,FU	NH_2-Lys-LL^{F}U^{F}ULL-$CONH_2$
HIV-BU,T	NH_2-Lys-LL^{Br}UTLL-$CONH_2$
HIV-T,BU	NH_2-Lys-LLT^{Br}ULL-$CONH_2$
HIV-T,T-cf[a]	CF-Lys-LLTTLL-$CONH_2$
HIV-BU,BU-cf[a]	CF-Lys-LL^{Br}U^{Br}ULL-$CONH_2$

[a]Carboxyfluorescein is attached on the backbone amine group of lysine residue through amide bond coupling.

steen hydrogen bonding (through pK_a reduction) and base stacking interactions, respectively (44). The stabilization effect of a single ^{Br}U base substitution is essentially position independent, as the K_d values for PNAs with singly modified ^{Br}U at positions 3, 5, and 6 are within a narrow range of 0.4-0.5 μM. Thus, ^{Br}U may be incorporated in between T and/or C residues in a PNA for enhanced binding to dsRNAs.

However, the binding affinity toward rHP1 of singly-modified PNAs containing ^{F}U residue was found to be position dependent (Table 2, Supplementary Figures S17 and S18). The PNA containing a ^{F}U modification at position 3 (FU-3, with ^{F}U in between two cytosine, C^{F}UC) has a K_d value of 1.1 μM, whereas, the PNA with the ^{F}U residue at position 5 (FU-5, flanked by cytosine and thymine, C^{F}UT) or 6 (FU-6, flanked by two thymine residues, T^{F}UT) has a tighter binding (K_d = 0.3 μM). It is likely that a fluorine atom has favorable interactions with an adjacent T residue, e.g., a C–H···F–C hydrogen bond or favorable electrostatic interaction (65–67) may form between the F atom and the methyl group of a T residue on the C-terminal side (see the computationally modelled triplex structure below).

It is interesting to note that singly- (at position 3, 5, or 6) and doubly-modified PNAs (at positions 3 and 5) generally show similar K_d values at pH 7.5 (Table 2, Figure 4A, and Supplementary Figures S17 and S18), suggesting that the stabilizing effect of an internal ^{X}U modification may propagate beyond the nearest neighbor and the next nearest neighbor stacking partners in a triplex. Such non-nearest neighbor effect has been previously revealed for the formation of duplexes with other modifications and mutations by bulk thermal melting and single-molecule mechanical unfolding experiments (68–73).

We tested if the modified PNAs form pre-organized helix structure by CD spectroscopy as previously observed for a backbone-modified PNA (74). Our CD spectroscopic data suggest the formation of PNA·dsRNA triplexes (Supplementary Figure S29A), which is in agreement with previously reported results (23,28,38). However, a ^{Br}U-modified PNA alone does not show helicity, indicating that the PNAs are not preorganized as a helical structure for triplex formation. It is possible that a ^{X}U modification reduces the flexibility of PNAs resulting in a reduced entropic penalty for binding to an RNA duplex. Taken together, our results indicate that a single ^{Br}U or multiple ^{Br}U (separated by more than two residues) residues may be incorporated into a PNA for enhanced binding to dsRNAs.

Substitution of T with U or halouracils in an unmodified PNA reduces pH dependence of PNA·dsRNA triplex formation

Due to a relatively low pK_a of 4.5 for N3 atom of a monomer cytosine, for a PNA containing one or more C bases, the formation of protonated C$^+$·G–C base triple (Figure 1A) and PNA·dsRNA triplex is expected to be pH dependent. It is known that the apparent pK_a of internal C residues in a TFO within an RNA·dsRNA triplex or a DNA·dsDNA triplex is shifted up near or above neutrality primarily driven by the triplex formation (8,75). Since substituting a RNA TFO with a PNA significantly enhances triplex formation (33), the apparent pK_a of internal C residues in a PNA·dsRNA triplex may also be up shifted toward or above neutrality.

Our PAGE studies show that the binding affinity of PNA P1 to rHP1 decreases significantly upon increasing pH with the K_d values of 2.1, 15.6, and 27.0 μM, respectively, at pH 7.5, 7.7, and 8.0 (Figure 4B, Table 2, Supplementary Figures S17–S22), which is consistent with the fact that PNA P1 has two internal and one terminal C residues. It is of note that our PAGE data reveal relatively less significant pH dependent binding for the doubly-modified PNAs (Table 2, Figures 3 and 4B). For example, with the pH varied from 7.5 to 8.0, PNA BU-3,5 has K_d values ranging from 0.3 to 1.2 μM.

We have previously discovered (33) that a PNA incorporating a modified base thio-pseudoisocytosine (L) has an enhanced recognition of a G–C pair in an RNA duplex through L·G–C base triple formation at physiological pH, and selective binding toward dsRNAs over ssRNAs (Figure 1B, E). The pK_a of a monomer L is around 9 (33,76,77). It is remarkable to note that, a PNA containing three L and no C residues (NH_2-Lys-TLTLTTTL-$CONH_2$) has a binding affinity (0.2 μM at pH 7.5) (34), similar to that of PNA BU-3,5 (NH_2-Lys-TC^{Br}UC^{Br}UTTC-$CONH_2$) containing two ^{Br}U and three C residues (0.3 μM at pH 7.5, Figure 3I, Table 2). The results indicate that upon substitution of T with ^{Br}U, a PNA·dsRNA triplex containing C$^+$·G–C base triples may become more resistant against high-pH destabilization. It is likely that enhancing hydrogen bonding and stacking involving ^{X}U·A–U base triples (compared to PNAs without ^{X}U substitution) may in turn favor C$^+$·G–C base triple formation and reduce the pH dependence by

Table 2. K_d values (in μM) for binding of PNAs to model RNA and DNA hairpins obtained by non-denaturing PAGE[a]

PNA[b]	rHP1			rHP5	rHP6	rHP7	dHP1
	pH 7.5	pH 7.7	pH 8.0			pH 7.7	
P1	2.1 ± 0.4	15.6 ± 3.6	27.0 ± 5.3	NB	NB	NB	NB
U-3	1.2 ± 0.3	3.1 ± 0.5	-	NB	NB	NB	NB
U-3,5	0.6 ± 0.1	1.6 ± 0.3	5.2 ± 0.6	NB	NB	NB	NB
FU-3	1.1 ± 0.2	2.5 ± 0.5	-	NB	NB	NB	NB
FU-5	0.3 ± 0.1	-	-	-	-	-	-
FU-6	0.3 ± 0.1	-	-	-	-	-	-
FU-3,5	0.3 ± 0.1	0.9 ± 0.2	4.7 ± 0.8	NB	NB	NB	NB
CU-3	0.6 ± 0.2	0.9 ± 0.2	-	NB	NB	NB	NB
CU-3,5	0.3 ± 0.1	0.5 ± 0.1	1.4 ± 0.2	NB	NB	NB	NB
BU-3	0.4 ± 0.1	1.0 ± 0.2	-	NB	NB	NB	NB
BU-5	0.4 ± 0.1	-	-	-	-	-	-
BU-6	0.5 ± 0.1	-	-	-	-	-	-
BU-3,5	0.3 ± 0.1	0.4 ± 0.1	1.2 ± 0.1	NB	NB	NB	NB
IU-3	0.8 ± 0.2	1.1 ± 0.2	-	NB	NB	NB	NB
IU-3,5	0.5 ± 0.1	1.1 ± 0.2	3.3 ± 0.5	NB	NB	NB	NB

[a]Incubation buffers at three different pH's were used: 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5, 7.7 or 8.0. '-' indicates that the data were not measured in this study. 'NB' indicates that no binding was observed.
[b]The letters and numbers indicate the type and position(s) of modifications. For instance, BU-3,5 indicates that T residues at positions 3 and 5 (from N-terminus to C-terminus) in P1 (NH_2-Lys-TCTCTTTC-$CONH_2$) have been replaced by 5-bromouracil.

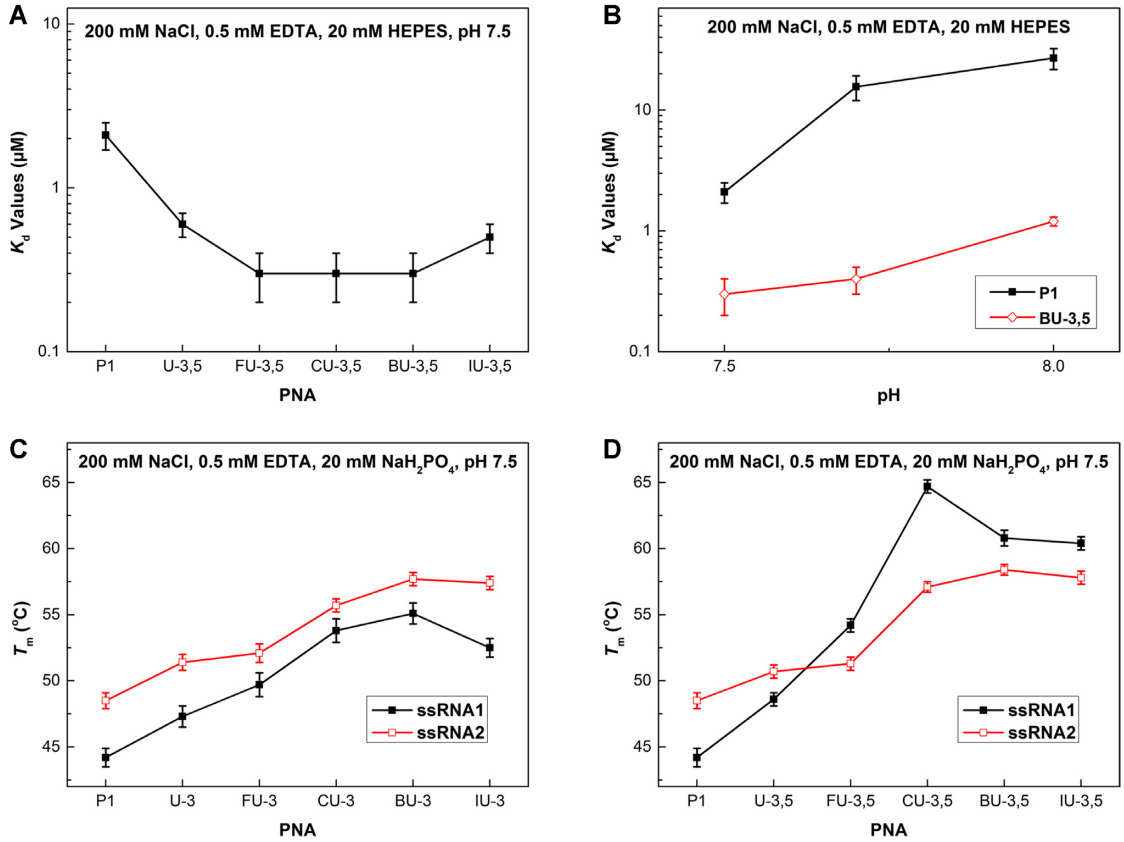


Figure 4. Summary of binding properties. (**A**) Comparison of K_d values of rHP1 binding to PNA P1 and doubly-modified PNAs obtained by PAGE. (**B**) K_d value versus pH (7.5, 7.7, and 8.0) for rHP1 binding to PNA P1 and BU-3,5 obtained by PAGE. Compared to unmodified PNA P1, PNA BU-3,5 shows a reduced pH dependence in binding to rHP1 to form a PNA·RNA_2 triplex. (**C**) Comparison of thermal melting temperature (T_m) values (heating) of PNA P1 and singly-modified PNAs binding to ssRNA1 and ssRNA2. (**D**) Comparison of T_m values (heating) of PNA P1 and doubly-modified PNAs binding to ssRNA1 and ssRNA2 (see Figure 2G, H).

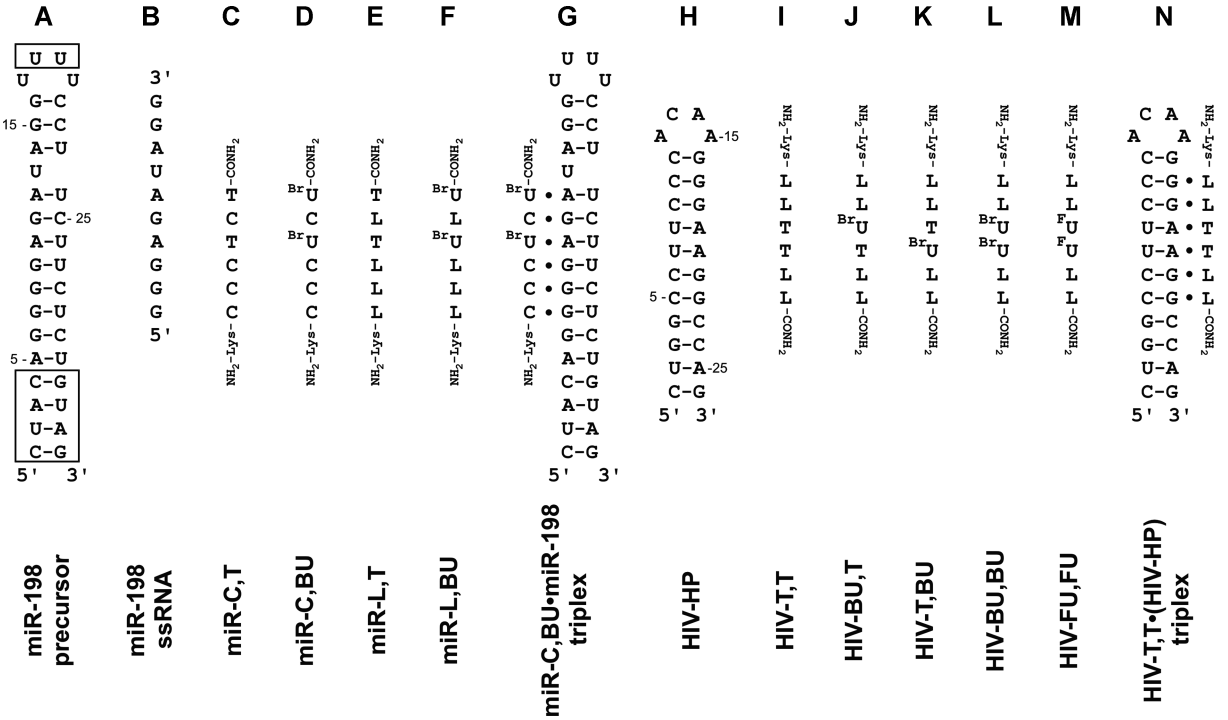


Figure 5. Sequences and structures of miR-198 hairpin precursor construct and HIV-1 ribosomal frameshift-inducing hairpin and PNAs for targeting the two RNAs. (**A**) miR-198 hairpin precursor construct. The residues shown in the boxes are added for the binding study. (**B**) A single-stranded fragment of miR-198 hairpin precursor. (**C–F**) PNAs targeting miR-198 hairpin precursor. (**G**) A PNA·RNA_2 triplex formed between PNA miR-C,BU and miR-198 hairpin precursor. (**H**) HIV-1 frameshift-inducing RNA hairpin. (**I–M**) PNAs targeting HIV-1 frameshift-inducing hairpin. (**N**) A PNA·RNA_2 triplex formed between PNA HIV-T,T and HIV-1 frameshift-inducing hairpin.

further shifting up the pK_a of C residues in a PNA·dsRNA triplex through the coupling between stacking and hydrogen bonding interactions (78). However, the K_d value of PNA FU-3,5 binding to rHP1 is still relatively pH dependent 0.3 and 4.7 μM at pH 7.5 and 8.0, respectively), probably because ^{F}U stabilizes triplex formation mainly through enhancing base triple formation but not by shifting the pK_a of adjacent C residues. Taken together, our results suggest that one may simply incorporate halouracils (e.g. ^{Br}U and ^{Cl}U) next to the C bases in PNAs to enhance PNA·dsRNA triplex formation and reduces pH dependence.

PNAs incorporating modified T residues retain sequence specificity in binding to dsRNA and show no appreciable binding to dsDNA

We next tested the sequence specificity of the ^{X}U modified PNAs. No triplex formation was observed at pH 7.7 between modified PNAs and mutated hairpins with an A–U pair in rHP1 (opposite to PNA position 5) substituted with a C–G pair (rHP5), a G–C pair (rHP6), or a U–A pair (rHP7) (Figure 2, Supplementary Figures S23 and S24). Our CD data confirm that PNA BU-3,5 can sequence specifically bind to rHP1 with no binding to rHP2 (Supplementary Figure S29A). rHP2 has the G9–C24 pair in rHP1 replaced with a C–G pair. The results suggest that the ^{X}U modified PNAs have a high sequence specificity through ^{X}U·A–U and C^{+}·G–C triple formation. Furthermore, P1 and the ^{X}U modified PNAs show no triplex formation at pH 7.7 with a DNA hairpin (dHP1, Figure 2E, Supplementary Figure S25), which is consistent with previously reported results (28,33,34). The preferential binding of PNAs to an RNA duplex over a DNA duplex may suggest that the relatively deep and narrow RNA duplex major groove is more compatible for accommodating a PNA. Taken together, it is promising to incorporate ^{X}U residues into PNAs for the enhanced and sequence-specific recognition of A–U pairs in RNA duplexes.

Substitution of T with U or halouracils in an unmodified PNA enhances binding to ssRNAs

It is expected that, similar to stabilizing a triplex by incorporating Hoogsteen ^{X}U·A pairs, incorporating Watson–Crick ^{X}U–A pairs would also stabilize Watson–Crick PNA–RNA duplex (45). We tested by thermal melting the binding of the ^{X}U-modified PNAs to ssRNAs through Watson–Crick duplex formation (Supplementary Figures S30 and S31). Compared to P1, the singly- and doubly-modified PNA sequences have higher melting temperatures (based on the heating curves), with ^{Cl}U and ^{Br}U modifications being the most stabilizing (Figure 4C, D).

Compared to ^{F}U base, ^{Cl}U and ^{Br}U may be more ideal for optimizing the coupled stacking and hydrogen bonding interactions (78,79) for the sequences tested here, with the PNAs containing pyrimidine residues only. The stabilization effect of ^{X}U substitution on PNA–RNA duplex was also observed previously for a different PNA sequence (with the ^{X}U residue flanked by two G residues), although relatively less variation among ^{F}U, ^{Br}U, and ^{Cl}U was observed

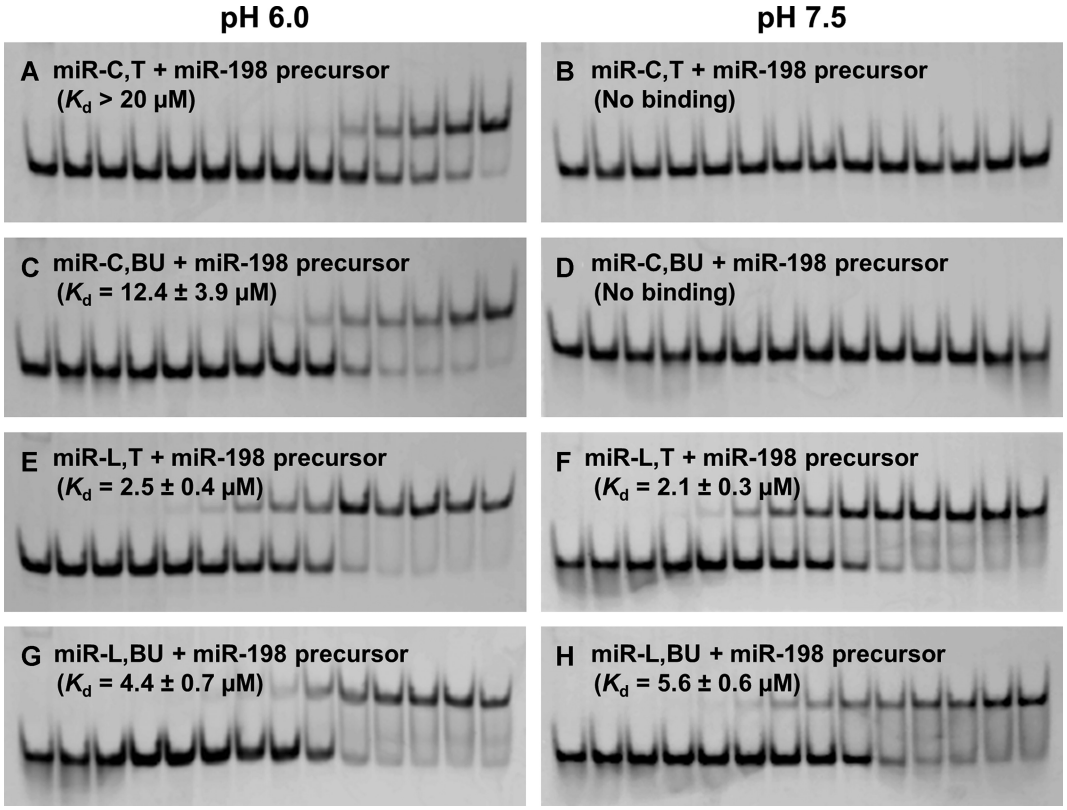


Figure 6. Non-denaturing PAGE (12 wt%) study of PNAs binding to miR-198 hairpin precursor. The gels were run with a running buffer of 1× TBE, pH 8.3 for 5 h at 250 V. The incubation buffer contains 200 mM NaCl, 0.5 mM EDTA, 20 mM MES, pH 6.0 (left column, **A**, **C**, **E**, **G**), or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5 (right column, **B**, **F**, **F**, **H**). The loaded miR-198 precursor RNA hairpin is at 1 μM in 20 μl. The PNA concentration in lanes from left to right are 0, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, 20, 28 and 50 μM, respectively. (**A–D**) miR-C,T and miR-C,BU bind to miR-198 hairpin precursor at pH 6.0 with K_d values of (>20) and (12.4 ± 3.9) μM, respectively. (**E-H**) miR-L,T and miR-L,BU bind to miR-198 hairpin precursor at both pH 6.0 (**E**, **G**) and pH 7.5 (**F**, **H**). An L^{Br}U step may be unfavorable compared to a C^{Br}U step for PNA·RNA$_2$ triplex formation, probably due to unfavorable stacking between N1 nitrogen of L and bromine atom of ^{Br}U (see Figures 8 and 9D).

(45). It is interesting to note that previous NMR studies suggest that a Watson–Crick ^{X}U–A pair is locally more dynamic with enhanced base pair opening kinetics and proton transfer within a base pair (80–83). Thus, enhanced local dynamics (as revealed by NMR) may contribute favorable entropy to the global stabilizing effect (observed in the UV-absorbance detected thermal melting experiment). We observed a relatively large hysteresis between the heating and cooling curves in the thermal melting experiment (Supplementary Figures S30 and S31). A comparison of the T_m values of the cooling curves suggests a relatively less significant stabilizing effect of ^{X}U substitution. The results may indicate that ^{X}U substitutions in a PNA strand stabilize a Watson–Crick PNA–RNA duplex mainly through enhancing molecular interactions in the duplex and thus decreasing unfolding kinetics.

Enhancing PNA–RNA Watson–Crick duplex formation may result in the invasion of a RNA–RNA duplex. Our PAGE data suggest that lowering NaCl concentration from 200 to 10 mM results in enhanced PNA·RNA–RNA triplex formation, but no strand invasion complex formation (Supplementary Figure S28A, B). The PAGE data suggest that addition of 2 mM $MgCl_2$ results in the destabilization of PNA·dsRNA triplexes (Supplementary Figure S28C, D). However, PNA BU-3,5 still shows a relatively tight binding even in the presence of 2 mM $MgCl_2$ (K_d values of 0.9 μM with 2 mM $MgCl_2$ versus 0.3 μM without 2 mM $MgCl_2$ (see Figure 3I and Supplementary Figure S28D)). DNA hairpin dHP1 shows no binding to the PNAs at 200 mM NaCl, but the formation of strand invasion complexes at 10 mM NaCl (Supplementary Figure S28E). Thus, substitution of T with ^{X}U is a promising strategy for enhancing the recognition of dsRNAs over dsDNAs at physiological conditions.

Targeting miR-198 hairpin precursor structure

We next incorporated ^{Br}U into PNAs to target miR-198 hairpin precursor structure (46,47). We designed a series of 6-mer PNAs for the recognition of a duplex region of miR-198 hairpin precursor (Figure 5A–G). Our PAGE results suggest that substitution of the two T residues in miR-C,T with ^{Br}U (miR-C,BU) results in an enhanced binding at pH 6.0 (from >20 μM to 12.4 μM), although no binding was observed at pH 7.5 (Figure 6A–D, Supplementary Figure S26). It is probable that simultaneous protonation of three consecutive C residues may be unfavorable with relatively low apparent pK_a values of the three C residues.

Substituting all four C residues in miR-C,T with L (miR-L,T, see Figure 1B, E for the structure of L) results in an enhanced binding at both pH 6.0 and 7.5 (Figure 6A, B, E,

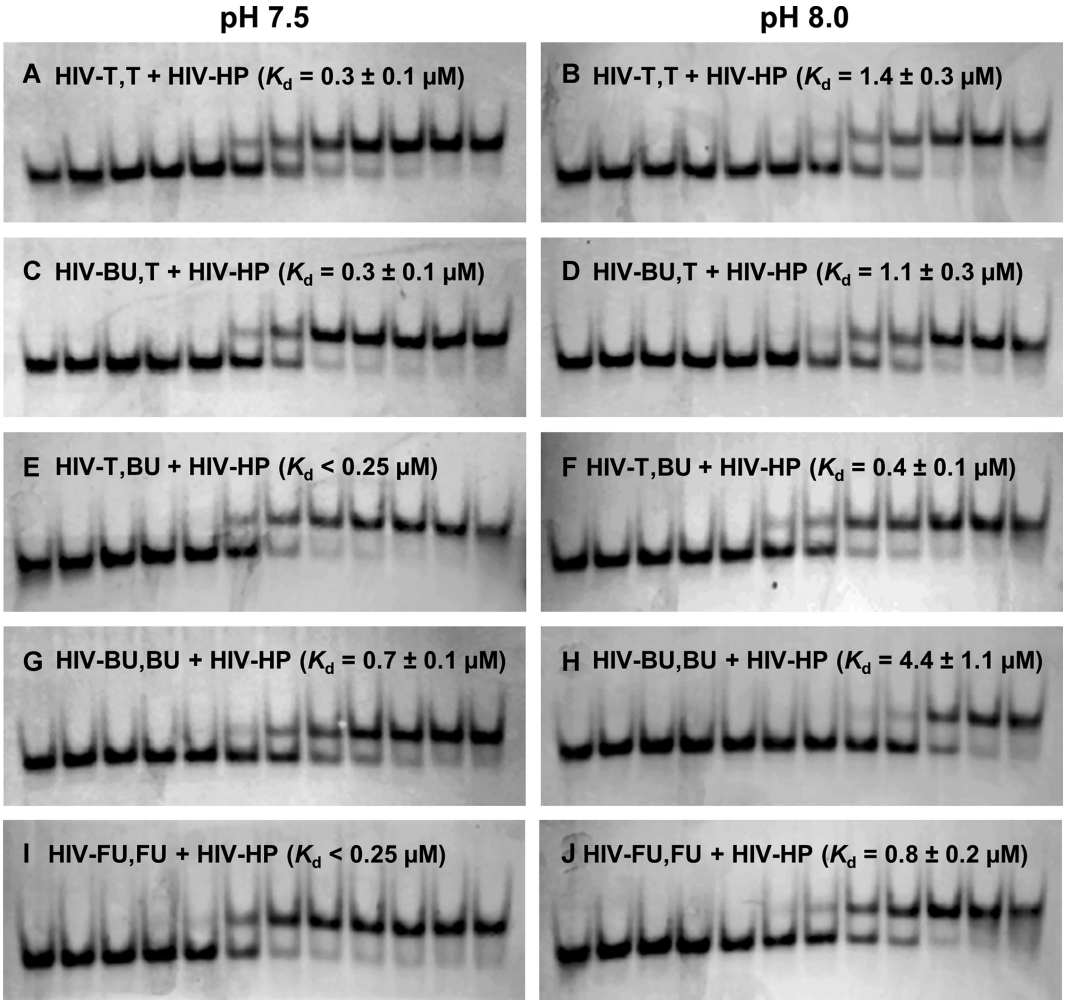


Figure 7. Non-denaturing PAGE (12 wt%) study of various PNAs binding to HIV-1 frameshift-inducing hairpin at pH 7.5 and 8.0. The gels were run with a running buffer of 1× TBE, pH 8.3 for 5 h at 250 V. The incubation buffer contains 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5 (left column, **A**, **C**, **E**, **G**, **I**) or pH 8.0 (right column, **B**, **D**, **F**, **H**, **J**). The loaded HIV-HP is at 0.25 μM in 25 μl. The PNA concentration in lanes from left to right are 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 2, 4, 10 and 20 μM, respectively.

F, Supplementary Figure S26). Consistent with our previous results (33,34), minimal pH dependence was observed for the L-modified miR-L,T. Thus, a neutral L residue is advantageous in substituting consecutive C residues to enhance the recognition of consecutive G–C pairs. However, further replacing the T residues in miR-L,T with ^{Br}U (miR-L,BU) destabilizes the PNA·RNA–RNA triplex (from 2.1 μM to 5.6 μM at pH 7.5, Figure 6E–H). The results indicate that adjacent L and ^{Br}U residues may not be cooperative in stabilizing a PNA·RNA–RNA triplex, possibly due to unfavorable stacking interactions in an L^{Br}U step (see below the computationally modeled triplex structures and stacking patterns).

Our CD spectroscopic data confirm the formation of PNA·RNA–RNA triplex formation involving the relatively short 6-mer PNAs (Supplementary Figure S29B). Interestingly, both our UV-absorbance-detected thermal melting and CD data suggest that the short PNAs show no strong binding to miR-198 ssRNA (5′-GGGAGAUAGG-3′), a 5′ fragment of miR-198 hairpin precursor (Figure 5B, Supplementary Figures S29C and S32). Thus, short PNAs are promising for selectively targeting dsRNAs over ssRNAs.

Targeting HIV-1 ribosomal frameshift inducing mRNA hairpin

We incorporated ^{Br}U into PNAs to target HIV-1 viral mRNA hairpin structure (Figure 5H) critical for stimulating −1 ribosomal frameshift (48–50). Our previous PAGE data suggest that a short 6-mer PNA (NH_2-Lys-LLTTLL-$CONH_2$, HIV-T,T, Figure 5I) can bind to the HIV-1 frameshift inducing hairpin (K_d = 1.1 μM at 200 mM NaCl, pH 7.5) (33). Here, we used a decreased RNA hairpin concentration of 0.25 μM (instead of 1 μM) and obtained the binding affinities more accurately (K_d = 0.3 μM for HIV-T,T at 200 mM NaCl, pH 7.5, Figure 7A). To compare the binding affinities of the PNAs with various substitutions, we also measured the binding at pH 8.0 (Figure 7B, D, F, H, J). We made two PNAs with single ^{Br}U substitutions (HIV-BU,T (NH_2-Lys-LL^{Br}UTLL-$CONH_2$) and HIV-T,BU (NH_2-Lys-LLT^{Br}ULL-$CONH_2$)) to probe

the sequence dependent stabilization effect of ^{Br}U substitution (Figures 5J, K and 7C–F). PNA HIV-BU,T shows similar binding affinities compared to PNA HIV-T,T (with K_d values of 1.1 versus 1.4 μM at 200 mM NaCl, pH 8.0, Figure 7A–D). The results are consistent with our binding data for PNAs targeting miR-198 hairpin precursor (Figure 6E–H) and the potential unfavorable stacking interactions for an L^{Br}U step present in PNA HIV-BU,T (see below the computationally modeled triplex structures and stacking patterns). Significantly, we observed enhanced PNA·RNA–RNA triplex formation for PNA HIV-T,BU compared to PNA HIV-T,T and HIV-BU,T (with K_d values of 0.4 versus 1.4 and 1.1 μM at 200 mM NaCl, pH 8.0, Figure 7B, D, F). The stabilization effect of PNA HIV-T,BU compared to PNA HIV-BU,T results mainly from the ^{Br}UL step in HIV-T,BU, compared to the L^{Br}U step in HIV-BU,T. The ^{Br}UT and T^{Br}U steps seem to have essentially the same stabilizing effect as observed for the rHP1-targeting PNAs containing no L residues (Table 2). Clearly, the thermodynamic effect of ^{Br}U substitution for T in L-containing PNAs is sequence dependent. One may stabilize a PNA·RNA–RNA triplex by incorporating a ^{Br}UL step instead of an L^{Br}U step.

Interestingly, PNA HIV-BU,BU (NH_2-Lys-LL^{Br}U^{Br}ULL-$CONH_2$) has a weakened binding compared to HIV-T,T (with K_d values of 4.4 versus 1.4 μM at pH 8.0, Figure 7B, H). Clearly, the thermodynamic contributions of consecutive ^{Br}U residues in a dsRNA-binding PNA are not additive, likely due to a local distortion of the helical structure with altered stacking geometry (84), although Br atom has a radius similar to that of a methyl group in T (85). In addition, stacking of two consecutive ^{Br}U residues in a dsRNA-binding PNA might result in the keto-enol tautomeric shift of the ^{Br}U base (43,86,87), and thus weaken PNA·RNA–RNA triplex formation. It is interesting that incorporating two consecutive ^{F}U residues in PNA HIV-FU,FU (NH_2-Lys-LL^{F}U^{F}ULL-$CONH_2$) shows a stabilizing effect in triplex formation as compared to PNA HIV-T,T (with K_d values of 0.8 versus 1.4 μM at pH 8.0, see Figure 7B, J). It is probable that the stabilization effect of ^{F}U substitution is mainly through enhancing base triple formation without significantly affecting stacking interactions with adjacent triples.

Extensive studies have been done for facilitating the cellular uptake bioactive PNAs and other peptides and nucleic acids (32,34,88–100). We focused on testing the effect of bromination of PNA on cellular uptake as previous studies suggest that fluorination of PNA T base does not significantly enhance cellular uptake of PNAs (101,102). Our confocal microscopic imaging studies show that the carboxyfluorescein-labelled PNAs HIV-T,T-cf and HIV-BU,BU-cf are both essentially cell impermeable (Supplementary Figure S33).

Computational modeling of PNA·dsRNA triplex structures

There is no high-resolution three-dimensional structure available for a PNA·dsRNA triplex. To gain further insights into the potential interactions responsible for the triplex formation stability, we computationally modeled the PNA·dsRNA triplexes using the existing crystal structure of PNA·DNA–PNA triplex as a starting structure (Figures 8 and 9A). We used the PNA sequence of AcNH-TLTLTTTL-$CONH_2$ to model the initial PNA·dsRNA triplex structure, because an L·G–C triple is relatively more stable than a C·G-C triple. We consider the modelled triplex structure to be reasonably accurate for providing the structural insights into the potential molecular interactions responsible for a PNA·dsRNA triplex formation. Our modeled PNA·dsRNA triplex structure suggests that the methyl group and halogen atom in T and ^{X}U bases may be potentially stacked with an adjacent base on the N-terminal side (Figures 8 and 9). Such a stacking interaction may explain the (de)stabilizing effects experimentally observed in our study.

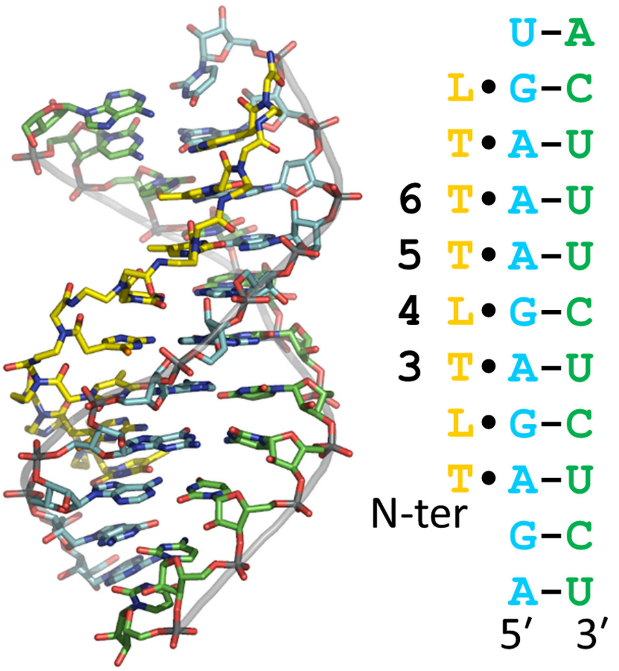


Figure 8. Modelled three-dimensional structure of a PNA·RNA–RNA triplex. We used the PNA sequence of AcNH-TLTLTTTL-$CONH_2$) because an L·G–C triple is relatively more stable than a C·G–C triple. The carbon atoms of the PNA strand are shown in yellow. The carbon atoms of the two RNA strands are shown in cyan and green, respectively. The hydrogen atoms are omitted for clarity. The stacking patterns of the base triples involving PNA residues from positions 3–6 are shown in Figure 9.

For example, a stabilizing effect may be expected if a base stack involves a C^{X}U step (from N-terminus to C-terminus, with X = halogen atoms), because, compared to a CT step, a C^{X}U step has an enhanced stacking in addition to improved Hoogsteen hydrogen bonding (Figure 9B, C). The modelled structures also imply that, for a ^{X}UT step (Figure 9C, D), favorable interactions may form between the T methyl group and ^{X}U halogen atom, e.g. a C–H···F–C hydrogen bond or favorable electrostatic interaction (65–67) as discussed above.

The modelled structures may also explain why an L^{Br}U step is less favorable than a ^{Br}UL step (Figures 5–7). For example, in an L^{Br}U step, charge-charge repulsion may be expected due to the overlap in a stacking geometry between the partially negatively charged N1 nitrogen atom of L and Br atom of ^{Br}U in the major groove (Figure 9D). Taken to-

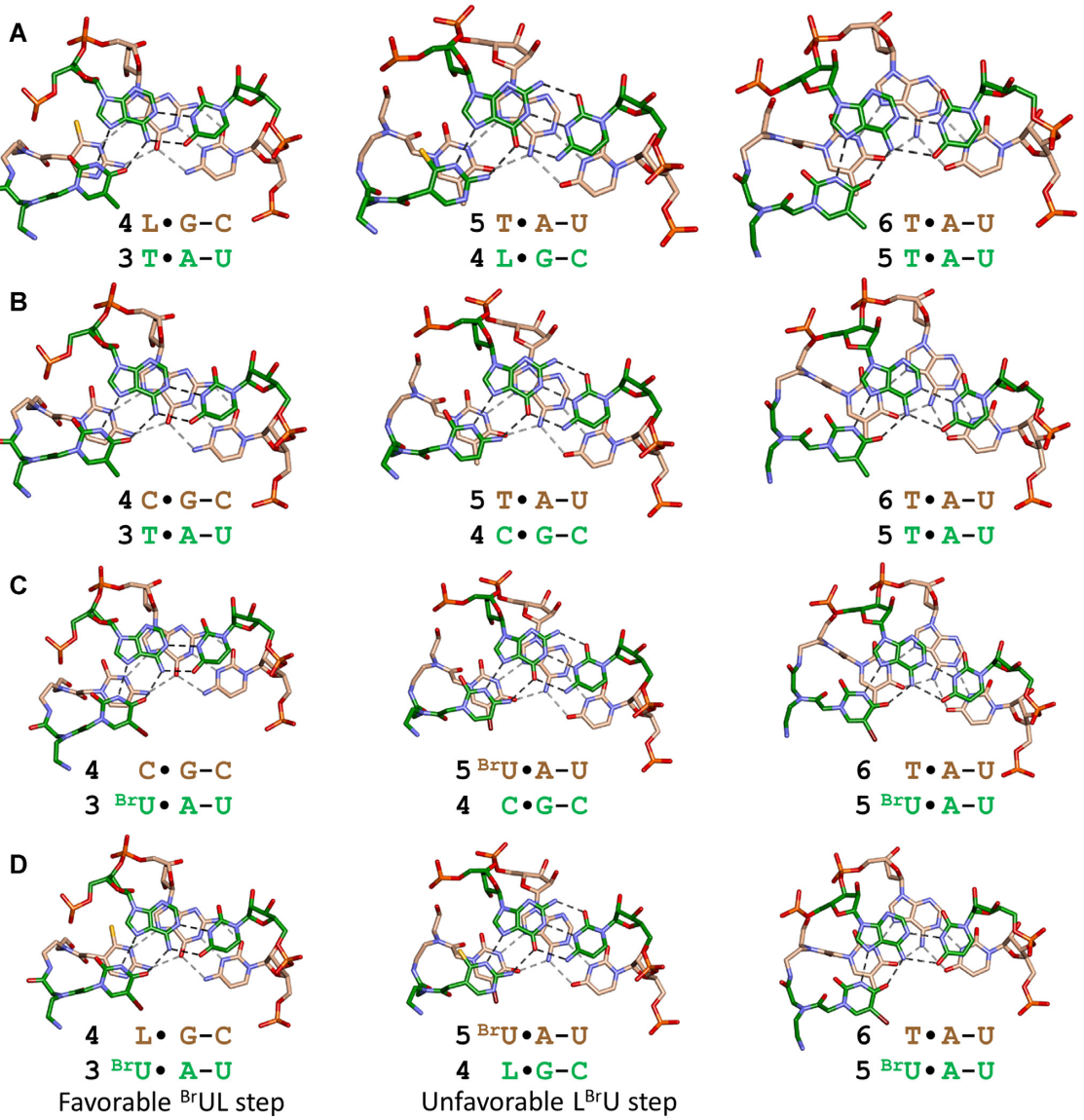


Figure 9. Potential stacking patterns of PNA·RNA–RNA base triples based on the modeled triplex structures (see Figure 8). The PNA strand is shown on the bottom left of each stacking figure. The stacking patterns shown in panel A are based on the modeled structure (with the PNA sequence of AcNH-TLTLTTTL-$CONH_2$) as shown in Figure 8. The triplex structures shown in panels B–D have the PNA sequences of AcNH-TCTCTTTC-$CONH_2$. (**B**) AcNH-TC^{Br}UC^{Br}UTTC-$CONH_2$ (**C**), and AcNH-TL^{Br}UL^{Br}UTTL-$CONH_2$ (**D**), respectively, and were built by replacing the corresponding PNA bases based on the structure shown in Figure 8 without further optimization. Only the base triples involving PNA residues 3–6 are shown. The modelled structure implies that a ^{F}UT step may be stabilized by a potential C–H···F–C hydrogen bond or favorable electrostatic interaction (65–67). The modelled stacking geometries shown in panel D suggest that a ^{Br}UL step is more favorable than an L^{Br}U step, likely because in an L^{Br}U step, there is an overlap between a nitrogen atom in L and bromine atom in ^{Br}U, which are both partially negatively charged.

gether, our experimental and computational modelling results suggest that stacking interactions are critical in stabilizing PNA·dsRNA triplex structures. One may incorporate ^{Br}U upstream of an L residue, but avoid having a ^{Br}U residue downstream of an L residue in designing dsRNA-binding PNAs.

CONCLUSIONS

In summary, substitution of T with U in PNAs can enhance the binding of short PNAs to dsRNAs through major-groove triplex formation, which is opposite to the destabilizing effect observed for DNA·DNA–DNA triplex formation (44,60). The 5-position halogenation of U can further enhance the triplex formation (with ^{Br}U substitution being the most stabilizing), presumably due to the optimized effects of improved hydrogen bonding and stacking interactions. Our results also suggest that one may incorporate a single ^{Br}U or multiple nonconsecutive ^{Br}U residues into C base-containing PNAs to enhance PNA·dsRNA triplex formation and reduce pH dependence. The reduced pH dependence may be due to an upshift of the apparent pK_a of C bases through favorable stacking interactions (78,79) with halouracil bases. It is significant that incorporating a ^{Br}UL step in a PNA (from N-terminus to C-terminus) can coop-

eratively enhance PNA·dsRNA triplex formation. Incorporating an L^{Br}U step in a PNA, however, may not stabilize PNA·dsRNA triplex formation (compared to an LT step). It is probably due to unfavorable stacking of an L^{Br}U step with the overlap of partially negatively charged N1 nitrogen in L and Br atom in ^{Br}U, as implicated in our modelling studies (Figures 8 and 9). Incorporating ^{X}U residues into PNAs also results in an enhanced binding to ssRNAs but no binding to dsDNAs at near-physiological buffer conditions.

Our work will facilitate rational design of various chemically-modified bases into PNAs for the enhanced and selective recognition of RNA duplexes containing varied base pairs. For example, it is promising that one may design dsRNA-binding PNAs by incorporating ^{Br}U residues not immediately downstream of L residues, which are advantageous in selectively binding to dsRNAs over ssRNAs (see Figure 1B, E) (13,33,35,36). Our work provides the foundation for developing modified PNAs as chemical probes and therapeutic ligands for targeting functionally important RNA duplex structures including those found in miRNA hairpin precursors and viral RNAs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Singapore Ministry of Education (MOE) Tier 1 [RGT3/13, RG42/15, RG152/17 to G.C.]; MOE Tier 2 [MOE2013-T2-2-024 and MOE2015-T2-1-028 to G.C.]. Funding for open access charge: Singapore Ministry of Education (MOE) Tier 1 [RG152/17 to G.C.].
Conflict of interest statement. None declared.

REFERENCES

1. Cech,T.R. and Steitz,J.A. (2014) The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*, **157**, 77–94.
2. Velagapudi,S.P., Gallo,S.M. and Disney,M.D. (2014) Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat. Chem. Biol.*, **10**, 291–297.
3. Deveson,I.W., Hardwick,S.A., Mercer,T.R. and Mattick,J.S. (2017) The dimensions, dynamics, and relevance of the mammalian noncoding transcriptome. *Trends Genet.*, **33**, 464–478.
4. Mathews,D.H., Moss,W.N. and Turner,D.H. (2010) Folding and finding RNA secondary structure. *Cold Spring Harb. Perspect. Biol.*, **2**, a003665.
5. Leontis,N.B. and Westhof,E. (2003) Analysis of RNA motifs. *Curr. Opin. Struct. Biol.*, **13**, 300–308.
6. Cannone,J.J., Subramanian,S., Schnare,M.N., Collett,J.R., D'Souza,L.M., Du,Y., Feng,B., Lin,N., Madabusi,L.V., Muller,K.M. *et al.* (2002) The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, **3**, 2.
7. Kwok,C.K., Tang,Y., Assmann,S.M. and Bevilacqua,P.C. (2015) The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem. Sci.*, **40**, 221–232.
8. Theimer,C.A., Blois,C.A. and Feigon,J. (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell*, **17**, 671–682.
9. Shefer,K., Brown,Y., Gorkovoy,V., Nussbaum,T., Ulyanov,N.B. and Tzfati,Y. (2007) A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Mol. Cell. Biol.*, **27**, 2130–2143.
10. Mitton-Fry,R.M., DeGregorio,S.J., Wang,J., Steitz,T.A. and Steitz,J.A. (2010) Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science*, **330**, 1244–1247.
11. Abu Almakarem,A.S., Petrov,A.I., Stombaugh,J., Zirbel,C.L. and Leontis,N.B. (2012) Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res.*, **40**, 1407–1423.
12. Bacolla,A., Wang,G. and Vasquez,K.M. (2015) New perspectives on DNA and RNA triplexes as effectors of biological activity. *PLoS Genet.*, **11**, e1005696.
13. Devi,G., Zhou,Y., Zhong,Z., Toh,D.-F.K. and Chen,G. (2015) RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdiscip. Rev. RNA*, **6**, 111–128.
14. Zhou,Y., Kierzek,E., Loo,Z.P., Antonio,M., Yau,Y.H., Chuah,Y.W., Geifman-Shochat,S., Kierzek,R. and Chen,G. (2013) Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. *Nucleic Acids Res.*, **41**, 6664–6673.
15. Semenyuk,A., Darian,E., Liu,J., Majumdar,A., Cuenoud,B., Miller,P.S., Mackerell,A.D. Jr and Seidman,M.M. (2010) Targeting of an interrupted polypurine:polypyrimidine sequence in mammalian cells by a triplex-forming oligonucleotide containing a novel base analogue. *Biochemistry*, **49**, 7867–7878.
16. Ohkubo,A., Yamada,K., Ito,Y., Yoshimura,K., Miyauchi,K., Kanamori,T., Masaki,Y., Seio,K., Yuasa,H. and Sekine,M. (2015) Synthesis and triplex-forming properties of oligonucleotides capable of recognizing corresponding DNA duplexes containing four base pairs. *Nucleic Acids Res.*, **43**, 5675–5686.
17. Fox,K.R. and Brown,T. (2011) Formation of stable DNA triplexes. *Biochem. Soc. Trans.*, **39**, 629–634.
18. Wang,S. and Kool,E.T. (1995) Relative stabilities of triple helices composed of combinations of DNA, RNA and 2′-O-methyl-RNA backbones: chimeric circular oligonucleotides as probes. *Nucleic Acids Res.*, **23**, 1157–1164.
19. Vuyisich,M. and Beal,P.A. (2000) Regulation of the RNA-dependent protein kinase by triple helix formation. *Nucleic Acids Res.*, **28**, 2369–2374.
20. Nielsen,P.E., Egholm,M., Berg,R.H. and Buchardt,O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **254**, 1497–1500.
21. Wittung,P., Nielsen,P.E., Buchardt,O., Egholm,M. and Norden,B. (1994) DNA-like double helix formed by peptide nucleic acid. *Nature*, **368**, 561–563.
22. Egholm,M., Buchardt,O., Christensen,L., Behrens,C., Freier,S.M., Driver,D.A., Berg,R.H., Kim,S.K., Norden,B. and Nielsen,P.E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature*, **365**, 566–568.
23. Wittung,P., Nielsen,P. and Norden,B. (1997) Extended DNA-recognition repertoire of peptide nucleic acid (PNA): PNA-dsDNA triplex formed with cytosine-rich homopyrimidine PNA. *Biochemistry*, **36**, 7973–7979.
24. Demidov,V.V., Protozanova,E., Izvolsky,K.I., Price,C., Nielsen,P.E. and Frank-Kamenetskii,M.D. (2002) Kinetics and mechanism of the DNA double helix invasion by pseudocomplementary peptide nucleic acids. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 5953–5958.
25. Bentin,T., Hansen,G.I. and Nielsen,P.E. (2006) Structural diversity of target-specific homopyrimidine peptide nucleic acid-dsDNA complexes. *Nucleic Acids Res.*, **34**, 5790–5799.
26. Hansen,M.E., Bentin,T. and Nielsen,P.E. (2009) High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers. *Nucleic Acids Res.*, **37**, 4498–4507.
27. Lonkar,P., Kim,K.H., Kuan,J.Y., Chin,J.Y., Rogers,F.A., Knauert,M.P., Kole,R., Nielsen,P.E. and Glazer,P.M. (2009) Targeted correction of a thalassemia-associated beta-globin mutation induced by pseudo-complementary peptide nucleic acids. *Nucleic Acids Res.*, **37**, 3635–3644.
28. Li,M., Zengeya,T. and Rozners,E. (2010) Short peptide nucleic acids bind strongly to homopurine tract of double helical RNA at pH 5.5. *J. Am. Chem. Soc.*, **132**, 8676–8681.
29. Zengeya,T., Gupta,P. and Rozners,E. (2012) Triple-helical recognition of RNA using 2-aminopyridine-modified PNA at physiologically relevant conditions. *Angew. Chem. Int. Ed.*, **51**, 12593–12596.

30. Zengeya,T., Gupta,P. and Rozners,E. (2014) Sequence selective recognition of double-stranded RNA using triple helix-forming peptide nucleic acids. *Methods Mol. Biol.*, **1050**, 83–94.
31. Endoh,T., Hnedzko,D., Rozners,E. and Sugimoto,N. (2016) Nucleobase-modified PNA suppresses translation by forming a triple helix with a hairpin structure in mRNA in vitro and in cells. *Angew. Chem. Int. Ed.*, **55**, 899–903.
32. Hnedzko,D., McGee,D.W., Karamitas,Y.A. and Rozners,E. (2017) Sequence-selective recognition of double-stranded RNA and enhanced cellular uptake of cationic nucleobase and backbone-modified peptide nucleic acids. *RNA*, **23**, 58–69.
33. Devi,G., Yuan,Z., Lu,Y., Zhao,Y. and Chen,G. (2014) Incorporation of thio-pseudoisocytosine into triplex-forming peptide nucleic acids for enhanced recognition of RNA duplexes. *Nucleic Acids Res.*, **42**, 4008–4018.
34. Toh,D.K., Devi,G., Patil,K.M., Qu,Q., Maraswami,M., Xiao,Y., Loh,T.P., Zhao,Y. and Chen,G. (2016) Incorporating a guanidine-modified cytosine base into triplex-forming PNAs for the recognition of a C–G pyrimidine-purine inversion site of an RNA duplex. *Nucleic Acids Res.*, **44**, 9071–9082.
35. Patil,K.M. and Chen,G. (2016) In: Jurga,S, Erdmann,VA and Barciszewski,J (eds). *Modified Nucleic Acids in Biology and Medicine*. Springer International Publishing, Cham, pp. 299–317.
36. Puah,R.Y., Jia,H., Maraswami,M., Kaixin Toh,D.F., Ero,R., Yang,L., Patil,K.M., Lerk Ong,A.A., Krishna,M.S., Sun,R. *et al.* (2018) Selective binding to mRNA duplex regions by chemically modified peptide nucleic acids stimulates ribosomal frameshifting. *Biochemistry*, **57**, 149–159.
37. Toh,D.-F.K., Patil,K.M. and Chen,G. (2017) Sequence-specific and selective recognition of double-stranded RNAs over single-stranded RNAs by chemically modified peptide nucleic acids. *J. Vis. Exp.*, **127**, e56221.
38. Sato,T., Sato,Y. and Nishizawa,S. (2016) Triplex-forming peptide nucleic acid probe having thiazole orange as a base surrogate for fluorescence sensing of double-stranded RNA. *J. Am. Chem. Soc.*, **138**, 9397–9400.
39. Tahtinen,V., Granqvist,L., Murtola,M., Stromberg,R. and Virta,P. (2017) ^{19}F NMR spectroscopic analysis of the binding modes in triple-helical peptide nucleic acid (PNA)/microRNA complexes. *Chem. Eur. J.*, **23**, 7113–7124.
40. Eldrup,A.B., Dahl,O. and Nielsen,P.E. (1997) A novel peptide nucleic acid monomer for recognition of thymine in triple-helix structures. *J. Am. Chem. Soc.*, **119**, 11116–11117.
41. Gupta,P., Zengeya,T. and Rozners,E. (2011) Triple helical recognition of pyrimidine inversions in polypurine tracts of RNA by nucleobase-modified PNA. *Chem. Commun.*, **47**, 11125–11127.
42. Felsenfeld,G., Davies,D.R. and Rich,A. (1957) Formation of a 3-stranded polynucleotide molecule. *J. Am. Chem. Soc.*, **79**, 2023–2024.
43. Privat,E.J. and Sowers,L.C. (1996) A proposed mechanism for the mutagenicity of 5-formyluracil. *Mutat. Res.*, **354**, 151–156.
44. Povsic,T.J. and Dervan,P.B. (1989) Triple helix formation by oligonucleotides on DNA extended to the physiological pH range. *J. Am. Chem. Soc.*, **111**, 3059–3061.
45. Liu,C.-D., Wang,J.-H., Xie,Y. and Chen,H. (2017) Synthesis and DNA/RNA complementation studies of peptide nucleic acids containing 5-halouracils. *Med. Chem. Commun.*, **8**, 385–389.
46. Kozomara,A. and Griffiths-Jones,S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.*, **42**, D68–D73.
47. Sundaram,G.M., Common,J.E., Gopal,F.E., Srikanta,S., Lakshman,K., Lunny,D.P., Lim,T.C., Tanavde,V., Lane,E.B. and Sampath,P. (2013) 'See-saw' expression of microRNA-198 and FSTL1 from a single transcript in wound healing. *Nature*, **495**, 103–106.
48. Brakier-Gingras,L., Charbonneau,J. and Butcher,S.E. (2012) Targeting frameshifting in the human immunodeficiency virus. *Expert Opin. Ther. Targets*, **16**, 249–258.
49. Hilimire,T.A., Bennett,R.P., Stewart,R.A., Garcia-Miranda,P., Blume,A., Becker,J., Sherer,N., Helms,E.D., Butcher,S.E., Smith,H.C. *et al.* (2016) N-Methylation as a strategy for enhancing the affinity and selectivity of RNA-binding peptides: Application to the HIV-1 frameshift-stimulating RNA. *ACS Chem. Biol.*, **11**, 88–94.
50. Brierley,I. and Dos Ramos,F.J. (2006) Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res.*, **119**, 29–42.
51. Lohse,B., Ramanujam,P.S., Hvilsted,S. and Berg,R.H. (2005) Photodimerization in pyrimidine-substituted dipeptides. *J. Pept. Sci.*, **11**, 499–505.
52. Ferrer,E., Shevchenko,A. and Eritja,R. (2000) Synthesis and hybridization properties of DNA–PNA chimeras carrying 5-bromouracil and 5-methylcytosine. *Bioorg. Med. Chem.*, **8**, 291–297.
53. Jacobsen,J.R., Cochran,A.G., Stephans,J.C., King,D.S. and Schultz,P.G. (1995) Mechanistic studies of antibody-catalyzed pyrimidine dimer photocleavage. *J. Am. Chem. Soc.*, **117**, 5453–5461.
54. Betts,L., Josey,J.A., Veal,J.M. and Jordan,S.R. (1995) A nucleic acid triple helix formed by a peptide nucleic acid-DNA complex. *Science*, **270**, 1838–1841.
55. Abraham,M.J., Murtola,T., Schulz,R., Páll,S., Smith,J.C., Hess,B. and Lindahl,E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25.
56. Zgarbova,M., Otyepka,M., Sponer,J., Mladek,A., Banas,P., Cheatham,T.E. 3rd and Jurecka,P. (2011) Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.*, **7**, 2886–2902.
57. Dupradeau,F.Y., Cezard,C., Lelong,R., Stanislawiak,E., Pecher,J., Delepine,J.C. and Cieplak,P. (2008) R.E.DD.B.: a database for RESP and ESP atomic charges, and force field libraries. *Nucleic Acids Res.*, **36**, D360–D367.
58. Vanquelef,E., Simon,S., Marquant,G., Garcia,E., Klimerak,G., Delepine,J.C., Cieplak,P. and Dupradeau,F.Y. (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.*, **39**, W511–W517.
59. Wang,J., Wolf,R.M., Caldwell,J.W., Kollman,P.A. and Case,D.A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.*, **25**, 1157–1174.
60. Soto,A.M., Rentzeperis,D., Shikiya,R., Alonso,M. and Marky,L.A. (2006) DNA intramolecular triplexes containing dT → dU substitutions: unfolding energetics and ligand binding. *Biochemistry*, **45**, 3051–3059.
61. Radhakrishnan,I. and Patel,D.J. (1994) Solution structure of a pyrimidine.purine.pyrimidine DNA triplex containing T·AT, C^+·GC and G·TA triples. *Structure*, **2**, 17–32.
62. Tarkoy,M., Phipps,A.K., Schultze,P. and Feigon,J. (1998) Solution structure of an intramolecular DNA triplex linked by hexakis(ethylene glycol) units: d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-TCTCTCTT). *Biochemistry*, **37**, 5810–5819.
63. Gotfredsen,C.H., Schultze,P. and Feigon,J. (1998) Solution structure of an intramolecular pyrimidine-purine-pyrimidine triplex containing an RNA third strand. *J. Am. Chem. Soc.*, **120**, 4281–4289.
64. Asensio,J.L., Dosanjh,H.S., Jenkins,T.C. and Lane,A.N. (1998) Thermodynamic, kinetic, and conformational properties of a parallel intermolecular DNA triplex containing 5′ and 3′ junctions. *Biochemistry*, **37**, 15188–15198.
65. Guo,F., Li,Q. and Zhou,C. (2017) Synthesis and biological applications of fluoro-modified nucleic acids. *Org. Biomol. Chem.*, **15**, 9552–9565.
66. Dickerhoff,J. and Weisz,K. (2017) Nonconventional C–H···F hydrogen bonds support a tetrad flip in modified G-quadruplexes. *J. Phys. Chem. Lett.*, **8**, 5148–5152.
67. Anzahaee,M.Y., Watts,J.K., Alla,N.R., Nicholson,A.W. and Damha,M.J. (2011) Energetically important C–H···F–C pseudohydrogen bonding in water: evidence and application to rational design of oligonucleotides with high binding affinity. *J. Am. Chem. Soc.*, **133**, 728–731.
68. Barnes,T.W. 3rd and Turner,D.H. (2001) Long-range cooperativity in molecular recognition of RNA by oligodeoxynucleotides with multiple C5-(1-propynyl) pyrimidines. *J. Am. Chem. Soc.*, **123**, 4107–4118.
69. Znosko,B.M., Barnes,T.W. 3rd, Krugh,T.R. and Turner,D.H. (2003) NMR studies of DNA single strands and DNA:RNA hybrids with and without 1-propynylation at C5 of oligopyrimidines. *J. Am. Chem. Soc.*, **125**, 6090–6097.

70. Kierzek,E., Ciesielska,A., Pasternak,K., Mathews,D.H., Turner,D.H. and Kierzek,R. (2005) The influence of locked nucleic acid residues on the thermodynamic properties of 2′-O-methyl RNA/RNA heteroduplexes. *Nucleic Acids Res.*, **33**, 5082–5093.
71. Carlucci,M., Kierzek,E., Olejnik,A., Turner,D.H. and Kierzek,R. (2009) Chemical synthesis of LNA-2-thiouridine and its influence on stability and selectivity of oligonucleotide binding to RNA. *Biochemistry*, **48**, 10882–10893.
72. McCann,M.D., Lim,G.F., Manni,M.L., Estes,J., Klapec,K.A., Frattini,G.D., Knarr,R.J., Gratton,J.L. and Serra,M.J. (2011) Non-nearest-neighbor dependence of the stability for RNA group II single-nucleotide bulge loops. *RNA*, **17**, 108–119.
73. Zhong,Z., Soh,L.H., Lim,M.H. and Chen,G. (2015) A U·U pair-to-U·C pair mutation-induced RNA native structure destabilisation and stretching-force-induced RNA misfolding. *ChemPlusChem*, **80**, 1267–1278.
74. Dragulescu-Andrasi,A., Rapireddy,S., Frezza,B.M., Gayathri,C., Gil,R.R. and Ly,D.H. (2006) A simple gamma-backbone modification preorganizes peptide nucleic acid into a helical structure. *J. Am. Chem. Soc.*, **128**, 10258–10267.
75. Asensio,J.L., Lane,A.N., Dhesi,J., Bergqvist,S. and Brown,T. (1998) The contribution of cytosine protonation to the stability of parallel DNA triple helices. *J. Mol. Biol.*, **275**, 811–822.
76. Kan,L.S., Lin,W.C., Yadav,R.D., Shih,J.H. and Chao,I. (1999) NMR studies of the tautomerism in pseudoisocytidine. *Nucleos. Nucleot.*, **18**, 1091–1093.
77. Ono,A., Tso,P.O.P. and Kan,L.S. (1991) Triplex formation of oligonucleotides containing 2′-O-methylpseudoisocytidine in substitution for 2′-deoxycytidine. *J. Am. Chem. Soc.*, **113**, 4032–4033.
78. Manalo,M.N., Perez,L.M. and LiWang,A. (2007) Hydrogen-bonding and pi-pi base-stacking interactions are coupled in DNA, as suggested by calculated and experimental trans-Hbond deuterium isotope shifts. *J. Am. Chem. Soc.*, **129**, 11298–11299.
79. Frankel,E.A., Strulson,C.A., Keating,C.D. and Bevilacqua,P.C. (2017) Cooperative interactions in the hammerhead ribozyme drive pKa shifting of G12 and its stacked base C17. *Biochemistry*, **56**, 2537–2548.
80. Sowers,L.C., Eritja,R., Kaplan,B.E., Goodman,M.F. and Fazakerley,G.V. (1987) Structural and dynamic properties of a fluorouracil-adenine base pair in DNA studied by proton NMR. *J. Biol. Chem.*, **262**, 15436–15442.
81. Kremer,A.B., Mikita,T. and Beardsley,G.P. (1987) Chemical consequences of incorporation of 5-fluorouracil into DNA as studied by NMR. *Biochemistry*, **26**, 391–397.
82. Fazakerley,G.V., Sowers,L.C., Eritja,R., Kaplan,B.E. and Goodman,M.F. (1987) Structural and dynamic properties of a bromouracil-adenine base pair in DNA studied by proton NMR. *J. Biomol. Struct. Dyn.*, **5**, 639–650.
83. Theruvathu,J.A., Kim,C.H., Rogstad,D.K., Neidigh,J.W. and Sowers,L.C. (2009) Base pairing configuration and stability of an oligonucleotide duplex containing a 5-chlorouracil-adenine base pair. *Biochemistry*, **48**, 7539–7546.
84. Holroyd,L.F. and van Mourik,T. (2015) Stacking of the mutagenic base analogue 5-bromouracil: energy landscapes of pyrimidine dimers in gas phase and water. *Phys. Chem. Chem. Phys.*, **17**, 30364–30370.
85. Hicke,B.J., Willis,M.C., Koch,T.H. and Cech,T.R. (1994) Telomeric protein-DNA point contacts identified by photo-cross-linking using 5-bromodeoxyuridine. *Biochemistry*, **33**, 3364–3373.
86. Kimsey,I.J., Petzold,K., Sathyamoorthy,B., Stein,Z.W. and Al-Hashimi,H.M. (2015) Visualizing transient Watson–Crick-like mispairs in DNA and RNA duplexes. *Nature*, **519**, 315–320.
87. Samijlenko,S.P., Yurenko,Y.P., Stepanyugin,A.V. and Hovorun,D.M. (2010) Tautomeric equilibrium of uracil and thymine in model protein-nucleic acid contacts. Spectroscopic and quantum chemical approach. *J. Phys. Chem. B*, **114**, 1454–1461.
88. Avitabile,C., Accardo,A., Ringhieri,P., Morelli,G., Saviano,M., Montagner,G., Fabbri,E., Gallerani,E., Gambari,R. and Romanelli,A. (2015) Incorporation of naked peptide nucleic acids into liposomes leads to fast and efficient delivery. *Bioconjugate Chem.*, **26**, 1533–1541.
89. Lattig-Tunnemann,G., Prinz,M., Hoffmann,D., Behlke,J., Palm-Apergi,C., Morano,I., Herce,H.D. and Cardoso,M.C. (2011) Backbone rigidity and static presentation of guanidinium groups increases cellular uptake of arginine-rich cell-penetrating peptides. *Nat. Commun.*, **2**, 453.
90. Ohmichi,T., Kuwahara,M., Sasaki,N., Hasegawa,M., Nishikata,T., Sawai,H. and Sugimoto,N. (2005) Nucleic acid with guanidinium modification exhibits efficient cellular uptake. *Angew. Chem. Int. Ed.*, **44**, 6682–6685.
91. Shiraishi,T. and Nielsen,P.E. (2011) Improved cellular uptake of antisense peptide nucleic acids by conjugation to a cell-penetrating peptide and a lipid domain. *Methods Mol. Biol.*, **751**, 209–221.
92. Wexselblatt,E., Esko,J.D. and Tor,Y. (2014) On guanidinium and cellular uptake. *J. Org. Chem.*, **79**, 6766–6774.
93. Wu,C.H., Weng,M.H., Chang,H.C., Li,J.H. and Cheng,R.P. (2014) Effect of each guanidinium group on the RNA recognition and cellular uptake of Tat-derived peptides. *Bioorg. Med. Chem.*, **22**, 3016–3020.
94. Zhou,P., Wang,M., Du,L., Fisher,G.W., Waggoner,A. and Ly,D.H. (2003) Novel binding and efficient cellular uptake of guanidine-based peptide nucleic acids (GPNA). *J. Am. Chem. Soc.*, **125**, 6878–6879.
95. Ma,X., Devi,G., Qu,Q., Toh,D.F., Chen,G. and Zhao,Y. (2014) Intracellular delivery of antisense peptide nucleic acid by fluorescent mesoporous silica nanoparticles. *Bioconjugate Chem.*, **25**, 1412–1420.
96. Das,I., Desire,J., Manvar,D., Baussanne,I., Pandey,V.N. and Decout,J.L. (2012) A peptide nucleic acid-aminosugar conjugate targeting transactivation response element of HIV-1 RNA genome shows a high bioavailability in human cells and strongly inhibits tat-mediated transactivation of HIV-1 transcription. *J. Med. Chem.*, **55**, 6021–6032.
97. Gupta,A., Bahal,R., Gupta,M., Glazer,P.M. and Saltzman,W.M. (2016) Nanotechnology for delivery of peptide nucleic acids (PNAs). *J. Control Release*, **240**, 302–311.
98. Torres,A.G., Fabani,M.M., Vigorito,E., Williams,D., Al-Obaidi,N., Wojciechowski,F., Hudson,R.H., Seitz,O. and Gait,M.J. (2012) Chemical structure requirements and cellular targeting of microRNA-122 by peptide nucleic acids anti-miRs. *Nucleic Acids Res.*, **40**, 2152–2167.
99. Doyle,D.F., Braasch,D.A., Simmons,C.G., Janowski,B.A. and Corey,D.R. (2001) Inhibition of gene expression inside cells by peptide nucleic acids: effect of mRNA target sequence, mismatched bases, and PNA length. *Biochemistry*, **40**, 53–64.
100. Cheng,C.J., Bahal,R., Babar,I.A., Pincus,Z., Barrera,F., Liu,C., Svoronos,A., Braddock,D.T., Glazer,P.M., Engelman,D.M. *et al.* (2015) MicroRNA silencing for cancer therapy targeted to the tumour microenvironment. *Nature*, **518**, 107–110.
101. Ellipilli,S., Palvai,S. and Ganesh,K.N. (2016) Fluorinated peptide nucleic acids with fluoroacetyl side chain bearing 5-(F/CF_3)-uracil: synthesis and cell uptake studies. *J. Org. Chem.*, **81**, 6364–6373.
102. Ellipilli,S. and Ganesh,K.N. (2015) Fluorous peptide nucleic acids: PNA analogues with fluorine in backbone (gamma-CF_2-apg-PNA) enhance cellular uptake. *J. Org. Chem.*, **80**, 9185–9191.