Incorporating G-C Pair-Recognizing Guanidinium into PNAs for Sequence and Structure Specific Recognition of dsRNAs over dsDNAs and ssRNAs

Manchugondanahalli S. Krishna,[1] Zhenzhang Wang,[2] Liangzhen Zheng,[3] Jogesh Bowry,[1,4] Alan Ann Lerk Ong,[1] Yuguang Mu,[3] Mookkan Prabakaran,[2] Gang Chen,[1,*]

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

[2]Temasek Life Science Laboratory, 1 Research Link, National University of Singapore, 117604, Singapore

[3]School of Biological Sciences, Nanyang Technological University, 637551, Singapore

[4]Department of Chemistry, University of Southampton, Southampton, SO17 1BJ, UK

Correspondence should be addressed to G.C. (RNACHEN@ntu.edu.sg)

Key words: RNA duplex; hydrogen bonding; stacking; RNA targeting; triplex

ABSTRACT

Recognition of RNAs at physiological conditions is important for developing chemical probes and therapeutic ligands. Nucleobase modified dsRNA-binding PNAs (dbPNAs) are promising for the recognition of dsRNAs in a sequence and structure specific manner at near-physiological conditions. Guanidinium is often present in proteins and small molecules for the recognition of G bases in nucleic acids, in cell-penetrating carriers, as well as in bioactive drug molecules, which might be due to the fact that guanidinium is amphiphilic with unique hydrogen bonding and stacking properties. We hypothesized that a simple guanidinium moiety can be directly incorporated into PNAs for facilitating enhanced molecular recognition of G-C pairs in dsRNAs and for improved bioactivity. We grafted guanidinium moiety directly into a PNA monomer (designated as R) using a two-carbon linker as guided by computational modeling studies. The synthetic scheme of PNA R monomer is relatively simple compared to that of previously reported L monomer. We incorporated the R residue into various dbPNAs for binding studies. dbPNAs incorporated with R residues are excellent in sequence specifically recognizing G-C pairs in dsRNAs over dsDNA and ssRNAs. We demonstrated that R residue is compatible with unmodified T and C and previously developed modified L and Q residues in dbPNAs for targeting model dsRNAs, influenza A viral panhandle duplex structure, and HIV-1 frameshift site RNA hairpin. Furthermore, R residues enhance the cellular uptake of PNAs.

INTRODUCTION

RNAs form diverse structures for interacting with metabolites and proteins, for catalysing RNA splicing and protein synthesis, and for regulating gene expression. Thus, developing chemical compounds for targeting desired RNA structures has great potential in biomedical applications.[1-4] The strategy of directly targeting double-stranded RNA (dsRNA) by triplex-formation is relatively underexplored mainly due to the weak binding at physiological conditions for the traditional triplex-forming oligonucleotides (TFOs).[5-8]

Nucleobase-modified dsRNA-binding peptide nucleic acids (dbPNAs) are promising in targeting dsRNAs in a sequence and structure specific manner (**Figures 1 and S1**).[9-18] Substituting C with a modified nucleobase may alleviate the pH dependence of G-C pair recognition through the formation of C^+·G-C triples (**Figures 1b-d and S1a,b**).[19, 20] For example, 2-aminopyridine (M, **Figure S1b**) has a monomer pK_a of 6.7 for the protonation of M to form M^+ and M-modified PNAs show enhanced recognition of G-C pairs in dsRNAs.[15, 21] M^+ residues also enhance cellular uptake of PNAs.[16, 17]

Pseudoisocytosine (J, **Figure S1a**) is a modified neutral nucleobase for the recognition of G-C pairs include.[22] J-modified PNAs have reduced pH dependence but relatively weakened recognition of G-C pair through Hoogsteen pairing.[11, 15] We reported the incorporation of thio-pseudoisocytosine (L, **Figure 1c,d**) into dbPNAs for the recognition of G-C pairs in dsRNAs.[11] The unique and selective base pairing (Hoogsteen over Watson-Crick, **Figures 1c,d and S1c**) and stacking properties of the L base result in L-modified dbPNAs to be selective in binding to dsRNAs

over ssRNAs and dsDNAs at near-physiological conditions. However, the synthesis of PNA L monomer is nontrivial and incorporation of L into PNAs does not improve cell permeability.

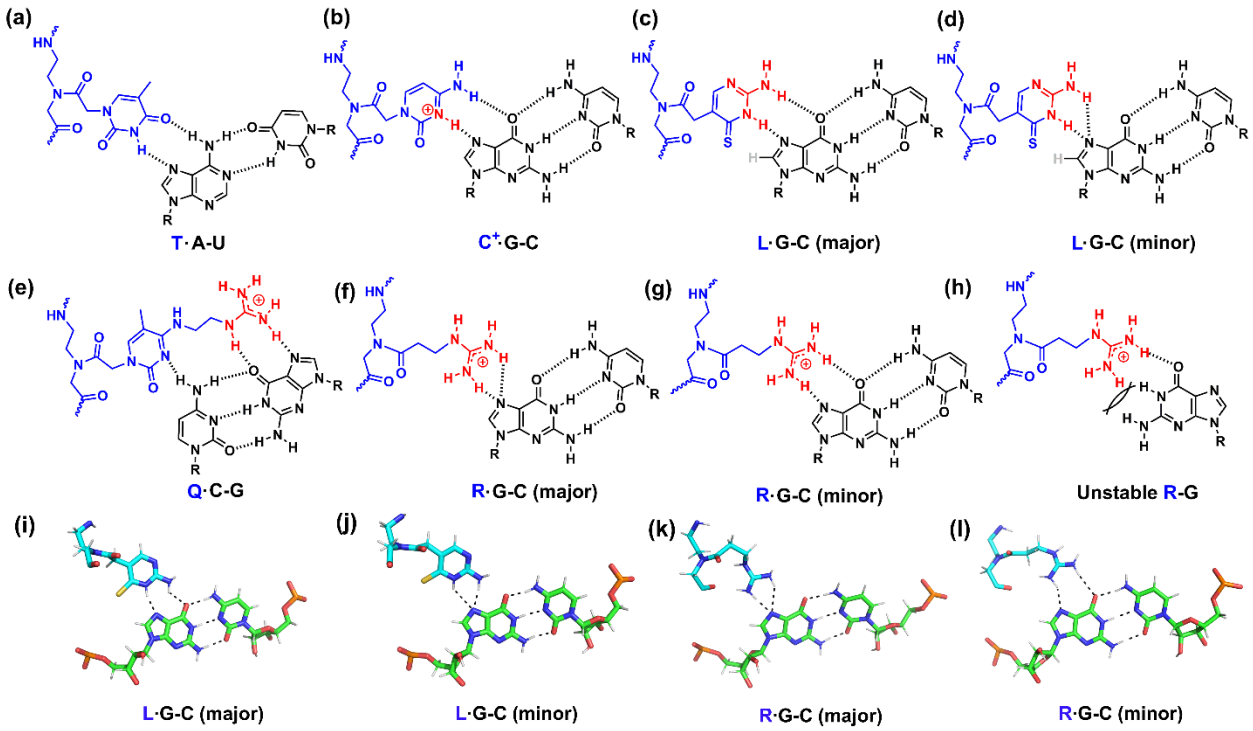


Figure 1. Base triples and a potential Watson–Crick-like base pair. The guanidinium moiety serving as hydrogen bond donors is highlighted in red. The hydrogen atoms shown in gray may be involved in van der Waals interactions in enhancing base triple formation (see panels **c** and **d**). (**a-g**) Hydrogen bonding patterns for base triples formed between PNA bases (T, C, L, Q, and R) and RNA Watson–Crick base pairs. (**h**) Unstable Watson–Crick-like R-G pair. (**i-l**) Modeled base triple structures of L·G-C and R·G-C. Two populations of structures with the formation of at least one hydrogen bond are shown.

Guanidinium (protonated form of guanidine) has unique hydrogen bonding and interesting stacking properties.[23, 24] Guanidinium (arginine residue) is present in proteins often involved in hydrogen bonding and stacking interactions with nucleic acids.[25-28] Various compounds incorporating guanidinium moiety show important biological activities.[29-37] Incorporation of the

guanidinium moiety facilitates enhanced cellular uptake of various cargos.[13, 23, 38-49] Free guanidinium has recently been found to be tightly bound in riboswitches.[50, 51]

Guanidine modified aromatic moieties have been incorporated into artificial nucleic acids for the recognition of single-stranded G residues as G-clamps[52, 53] or C-G pairs through Q·C-G triple (**Figure 1e**) or other triple formation.[13, 14, 54-56] Our modeling studies (**Figure 1f,g,k,l**) suggest that guanidinium may be directly incorporated into PNAs (without an extra aromatic group) for enhanced recognition of G-C pairs in dsRNAs. We grafted guanidinium moiety into a PNA monomer (designated as R) through a simple two-carbon linker (**Figure 1f,g**) followed by incorporating the R residue into various dbPNAs for binding studies. We further tested the compatibility of R in combination with C, T, L, and Q residues for targeting model dsRNAs and dsRNA structures found in influenza and HIV-1 viral RNAs.

MATERIALS AND METHODS

Computational modeling

The PNA·dsRNA triplex containing the PNA sequence H_2N-Lys-TLTRTTTL-$CONH_2$ was generated based on the model for a PNA·dsRNA triplex containing the PNA sequence H_2N-Lys-TLTLTTTL-$CONH_2$.[57, 58] We also replaced the two-carbon base-backbone linker in R with a three-carbon linker (R′) to obtain a PNA·dsRNA triplex containing the PNA sequence H_2N-Lys-TLTR′TTTL-$CONH_2$. We modelled the RNA molecules with Amber force field χOL3.[59, 60] and the PNA residues with Amber GAFF force field[61] and AM1-BCC atomic charges.[62] TIP3P explicit solvent[63] and counter ions (Na^+) were added to the simulation boxes. Next, we performed a 1000-

step energy minimization using steepest descent algorithm.[64] For each model, three independent 20-ns NPT simulations were carried out with random initial velocities. The cutoff for van der Waals and electrostatics was 1.2 nm. All the simulations were performed with OpenMM simulation package.[65] The three trajectories for each system were merged by removing the first 10% of the trajectories, and then the hydrogen bonds were calculated with Mdtraj,[66] considering the distance between hydrogen and hydrogen bond acceptor within 0.25 nm and angle between hydrogen bond acceptor, hydrogen and hydrogen bond donor larger than 120 °.

General experimental methods

The required reagents, starting materials, and all the solvents were obtained from the available commercial vendors and used without further purification. All the reactions were performed using the high performance liquid chromatography (HPLC) grade solvents at room temperature unless otherwise mentioned. The reaction progress was monitored by thin layer chromatography (TLC) on aluminum sheets silica gel 60 F254 (Merck) and liquid chromatography-mass spectrometry with electrospray ionization source (LCMS-ESI). Column chromatographic purification was done for Fmoc-, t-butyl ester compound **5**, using ethyl acetate/hexane solvent systems on flash silica gel 230–400 mesh size. The intermediate compounds and the final monomer were characterized by high-resolution mass spectrometry (electron ionization) (HRMS-EI), and ^{1}H and ^{13}C NMR spectra on a 400 MHz (100 MHz, ^{13}C) or 500 MHz Bruker spectrometer. The NMR chemical shifts (δ) are expressed in parts per million (ppm). The residual solvent peaks were used as references for the ^{1}H (chloroform-d: 7.26; dimethyl sulfoxide-d6: 2.50) and ^{13}C (chloroform-d: 77.0; dimethyl sulfoxide-d6: 39.5) NMR spectra.

Synthesis of PNA monomer R

The synthetic protocol for the PNA R monomer is outlined in **Scheme 1** with the details shown in **Schemes S1-S3** and **Figures S2-S6**. The starting materials **1** and **2** were synthesized using the previously reported procedures.[67, 68] Compound **4** was obtained by a 4-hour reaction of the di-Cbz protected pyrazole guanidine **1** with crushed 3-aminopropionic acid **3** using DIPEA in DCM at room temperature. The obtained white solid **4** was reacted with **2** in the presence of HBTU and NMM in anhydrous DMF. The reaction was carried out under nitrogen atmosphere at room temperature followed by a column purification using EtOAc/Hexane (v/v; 3/10) solvent system to give precursor t-butyl ester compound **5**. Compound **5** was treated with 50% TFA-DCM for 1 hour, followed by the removal of the solvent with a subsequent precipitation using cold di-ethyl ether. The product was filtered, washed, and dried in vacuum to yield the required PNA R monomer **6**.

Solid phase synthesis of PNA oligomers

The unmodified N-(2-aminoethyl)glycine PNA (aegPNA) monomers were acquired from ASM Research Chemicals and used as such for the synthesis of PNA oligomers. The L and Q monomers were synthesized as per the previously reported procedures.[11, 13] The synthesis of PNA oligomers was done on Methylbenzhydrylamine hydrochloride (MBHA·HCl) polystyrene-based resin as solid support. The initial loading value of 1.5–1.7 mmol/g was brought down to 0.35 mmol/g using acetic anhydride as capping agent. Except for the R monomer, the oligomer synthesis was performed using Boc-strategy in the presence of (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBop) and N,N-Diisopropylethylamine (DIPEA) as the coupling agents. The Fmoc deprotection was done using 20% piperidine in DMF.

The final cleavage of the oligomer from the resin was done using the high-low trifluoroacetic acid (TFA)-trifluoromethanesulfonic acid (TFMSA) method once the coupling has been done for all the unmodified and modified PNA monomers. The crude PNA oligomers were precipitated with ice-cold diethyl ether and dissolved in de-ionized water. The purification of PNA oligomers was done by reverse-phase HPLC (RP-HPLC) using water-CH_3CN-0.1% TFA as the mobile phase. The oligomers were characterized by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) spectrometry with α-cyano-4-hydroxycinnamic acid (CHCA) as sample crystallization matrix. The extinction co-efficient of PNA T, C, Q, L, and R residues are 8.8, 7.3, 7.3, 7.3, and 0 $mM^{-1}cm^{-1}$, respectively, at 260 nm.[10, 13]

Nondenaturing polyacrylamide gel electrophoresis

The HPLC-purified RNA and DNA oligonucleotides were acquired from Sigma-Aldrich, Singapore. The nondenaturing polyacrylamide gel electrophoresis (PAGE) (12%) experiments were performed using an incubation buffer containing 200 mM NaCl, 0.5 mM EDTA, 20 mM MES, pH 6.0, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5, or 200 mM NaCl, 0.5 mM EDTA, 2 mM $MgCl_2$, 20 mM HEPES, pH 7.5, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 8.0. For the Cy3-labelled rHP1-Cy3 and rHP3-Cy3, the samples were prepared using 50 nM RNA with varied PNA concentrations of 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 µM in 20 µL of buffer. For the unlabeled RNA/DNAs, the concentration was 0.25 µM with varied PNA concentrations of 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 µM in 25 µL of buffer. Snap cooling was done for RNA/DNA hairpins by directly dipping the sample tubes from 95 ^oC into ice bath followed by annealing with PNAs by slow cooling from 65 ^oC to room temperature. The sample tubes were incubated overnight at 4 ^oC before the loading. To all

the samples, 35% glycerol (20% of the total loaded volume) was added followed by vortexing for a few seconds before the loading into sample wells of the gels. Using a running buffer 1× TBE (Tris–Borate–EDTA), pH 8.3, the gels were run at 4 ^{o}C for 6 h with a voltage set at 250 V for all the experiments. The gels were imaged using a Typhoon imager directly for the Cy3-labelled (rHP1-Cy3 and rHP3-Cy3) or after ethidium bromide staining for 30 min for the unlabeled RNA/DNA hairpins.

Fluorescence binding study

The fluorescence emission spectra were recorded using a 1 cm square quartz cuvette on a Varian Cary Eclipse fluorescence spectrophotometer. The required 2-aminopurine-labelled RNA and DNA oligonucleotides purified by RP-HPLC were acquired from Sigma-Aldrich, Singapore. All the fluorescence experiments were carried out using 1 μM RNAs or DNAs and varied concentrations of PNAs at 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 μM in 70 μL buffer having 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. PNA alone samples at 10 μM were used as controls. The samples were prepared by slow cooling of RNAs/DNAs hairpins from 95 ^{o}C to room temperature, followed by the addition of PNA oligomers. The solutions were then vortexed for 10 min and left at room temperature for 1 hour and incubated at 4 ^{o}C overnight. All the emission spectra were recorded at room temperature over a wavelength range of 330–550 nm, at a fixed excitation wavelength of 303 nm with a slit width of 10 nm.

Thermal melting

The UV-absorbance-detected thermal melting experiments were conducted with the use of an 8-microcell cuvette using the Shimadzu UV-2550 UV-Vis spectrophotometer. The absorbance

values at 260 nm with the temperature increasing from 15 to 95 ^{o}C followed by the temperature decreasing from 95 to 15 ^{o}C were recorded. The optical path length of the 8-microcell cuvette is 1 cm. The temperature ramp rate is 0.5 ^{o}C/min. The samples were prepared using 5 µM RNA with and without 5 µM PNA in 130 µL buffer having 200 mM NaCl, 0.5 mM EDTA, 20 mM NaH_2PO_4, pH 7.5. The samples containing the single-stranded RNA and PNA were annealed by slow cooling from 95 ^{o}C to room temperature, followed by overnight incubation at 4 ^{o}C. Data were normalized at high temperature and the melting temperatures were determined based on the Gaussian fit of the first derivative of the curves.

Cellular uptake and confocal microscopy study

Sf9 (*Spodoptera frugiperda*) insect cells were cultured onto a 4-well chamber slide system (Nunc Lab-Tek) at a cell density of 2×10^5 cells per compartment for 12 h. The cells were washed with PBS and replaced with serum free Sf-900 III medium (Life Technologies Inc., NY, USA) containing 2 µM of R-modified PNA P5-3R-CF or control PNA P1-CF for 12 h at 27 ^{o}C. Then, cells were fixed with 4% paraformaldehyde for 15 min at room temperature, rinsed with PBS. The fixed cells were counterstained with DAPI (Invitrogen, Gaithersburg, USA) for 5 min. Fluorescence signals were observed using an inverted confocal microscope (Leica SP8, Leica Microsystems, Mannheim, Germany).

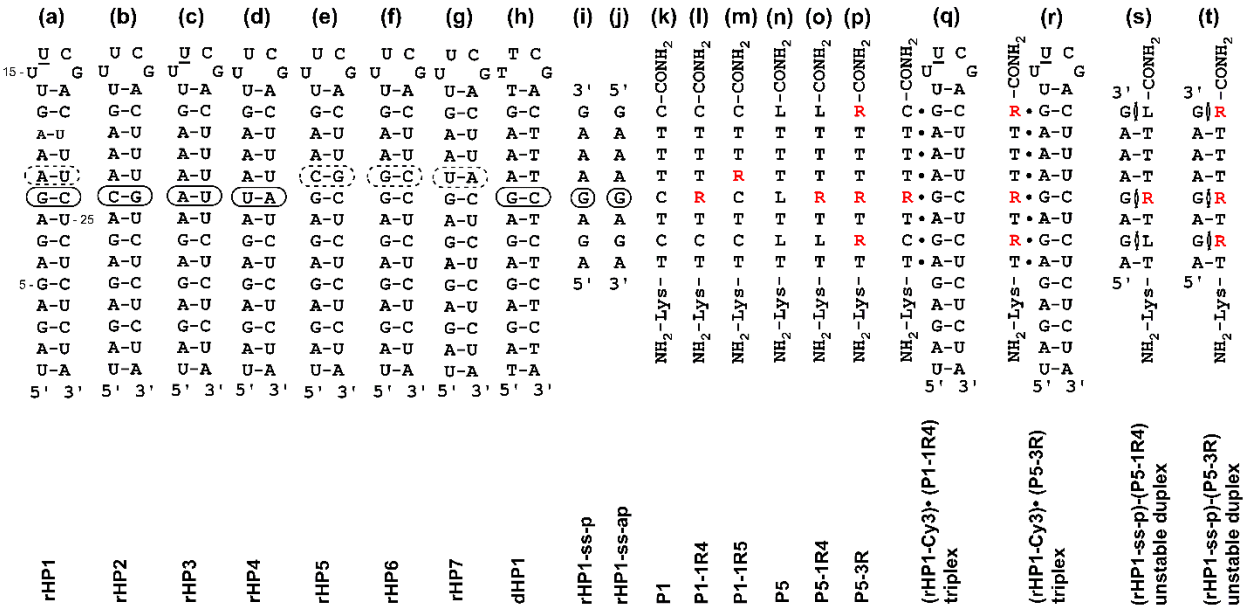


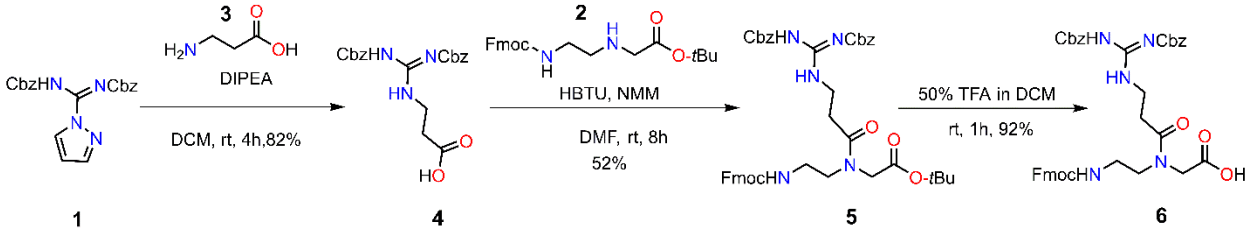
Figure 2. RNA, DNA, and PNA molecules studied in this paper. (**a-h**) Model RNA and DNA hairpins. The underlined U residues in rHP1 and rHP3 are replaced with Cy3 dye to have rHP1-Cy3 and rHP3-Cy3, respectively. (**i-j**) Model ssRNA strands that potentially form parallel and antiparallel duplexes with PNAs. (**k-p**) PNAs studied in this paper. The letter "R" shown in red represents the PNA R residue. (**q,r**) Stable PNA·RNA_2 triplexes formed by PNA P1-1R4 and P5-3R targeting rHP1. (**s,t**) Unstable duplexes formed by rHP1-ss-p with P5-1R4 and P5-3R, respectively.

RESULTS AND DISCUSSION

Recognition of an internal G-C pair in an RNA duplex (but not DNA duplex) by R-modified PNAs

Our modeling data suggest that an R-modified PNA can recognize an RNA duplex through major-groove PNA·dsRNA triplex formation similar to an L-modified PNA (**Figures 1, 2, and S1**). Among the structures with one or more hydrogen bonds formed between R and G or between L and G, we observed two main populations. It seems that L is more optimal for the simultaneous recognition of the N1 and carbonyl group of G on the Hoogsteen edge, probably because L has a relatively large stacking area and more optimized base-backbone linker. However, the R′ residue,

with an additional carbon in the base-backbone linker, shows increased flexibility and may result in the loss of hydrogen bonding and stacking interactions (**Figure S1g-i**). Thus, we focused on the R monomer (**Scheme 1**) in our experimental studies.



Scheme 1. Chemical synthesis of PNA R monomer.

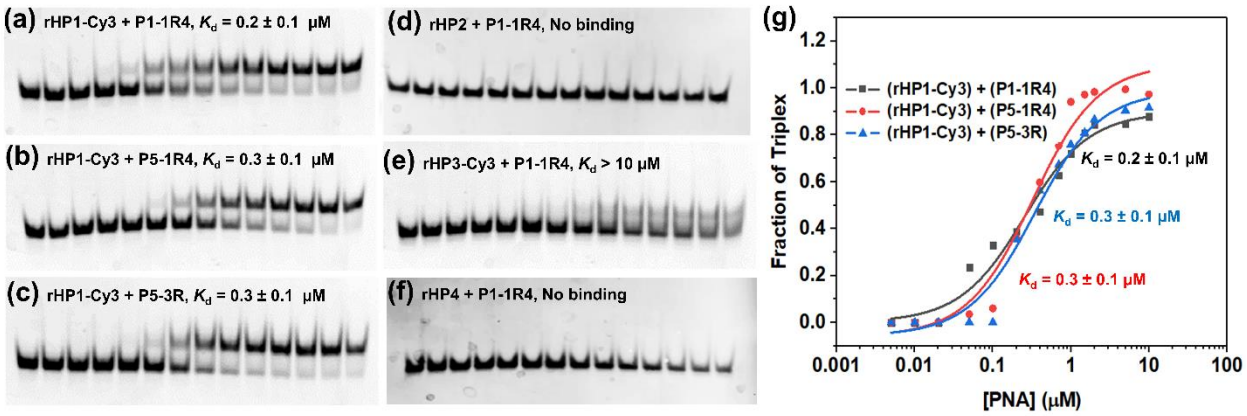


Figure 3. Nondenaturing PAGE characterization of triplex formation for model RNA hairpins. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. (**a-c and g**) PAGE characterization and determination of K_d for the binding of dbPNAs to rHP1 having a Watson-Crick G-C pair opposite to R. (**d-f**) PAGE characterization for the binding of dbPNA P1-1R4 to RNA hairpins with a C-G (rHP2), an A-U (rHP3-Cy3), or a U-A (rHP4) base pair opposite to R. rHP1 shows the strongest binding to P1-R4, suggesting a specific recognition of a G-C pair by an R residue. All the Cy3-labelled samples were prepared at 50 nM with 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 μM of PNA in 20 μL of buffer. rHP2 and rHP4 were at 0.25 μM against 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μM PNA in 25 μL of buffer.

The PNA oligomers were obtained by solid phase peptide synthesis (**Table S1** and **Figure S7)**. Nondenaturing PAGE experiments (**Figures 3** and **S8-S11**) were performed to characterize the binding properties of PNAs incorporated with R residue. Replacement of a single C residue at position 4 in PNA P1 with R (P1-1R4, **Figure 2k,l**) results in a significant enhancement in binding with the K_d value decreased from 4.9 μM[58] to 0.2 μM at 200 mM NaCl, pH 7.5 (**Figure 3, Table 1**). Interestingly, P1-R4 has the same binding affinity as that of PNA P5 (**Figure 2n**) with three L residues (**Figure 3, Table 1**). Notably, PNA P1-1R4 shows significantly weakened binding to rHP3 and no observable binding to rHP2 and rHP4 (**Figure 3, Table 1**), suggesting that the PNA residue R is highly specific for the recognition of a G-C pair in rHP1 but not C-G, A-U, and U-A pairs in rHP2, rHP3, and rHP4, respectively. Furthermore, the fluorescence titration experiments of P1-1R4 into 2-aminopurine-labelled RNAs (**Figures 4 and S12**) reveal a K_d value (0.4 ± 0.2 μM) in binding to dsRNA1-2AP, which has a similar binding site as rHP1. Importantly, no appreciable binding is observed for the duplexes with the base pair opposite to the R residue replaced with a C-G (dsRNA2-2AP), an A-U (dsRNA3-2AP), or a U-A (dsRNA4-2AP) (**Figures 4 and S12**). Clearly, incorporating guanidinium into a dbPNA facilitates specific recognition of a G-C pair in a dsRNA.

Table 1. K_d **values (in μM) for PNAs binding to model RNA and DNA hairpins obtained by nondenaturing PAGE.**[a]

	rHP1-Cy3 (G-C)				rHP2 (C-G)			rHP3-Cy3 (A-U)			rHP4 (U-A)			dHP1 (G-C)	
pH	**6.0**	**7.5**		**8.0**	**6.0**	**7.5**	**8.0**	**6.0**	**7.5**	**8.0**	**6.0**	**7.5**	**8.0**	**6.0**	**7.5**
		No Mg^{2+}	**With Mg^{2+}**												
P1[b]	(0.9±0.2)	(2.1±0.4) 4.9±1.7	(5.5±0.9)	(27.0±5.3)	-	(NB)	-	-	(NB) NB	-	-	(NB)	-	-	(NB)
P5[b]	(0.2±0.1)	(0.2±0.1) 0.2±0.1	(0.2±0.1)	(0.3±0.1)	-	(NB)	-	-	(0.4±0.1) 0.4±0.1	-	-	(1.9±0.4)	-	-	(>50)
P1-1R4	0.1±0.1	0.2±0.1	0.6±0.1	0.8±0.1	(NB)	(NB)	(NB)	>3	>10	>10	(NB)	(NB)	(NB)	(NB)	(NB)
P5-1R4	0.2±0.1	0.3±0.1	1.3±0.2	0.7±0.1	-	-	-	-	-	-	-	-	-	(NB)	(NB)
P5-3R	0.2±0.1	0.3±0.1	0.4±0.1	0.5±0.1	-	(NB)	-	-	>2	-	-	(NB)	-	(NB)	(NB)
	rHP1-Cy3 (A-U)				**rHP5 (C-G)**			**rHP6 (G-C)**			**rHP7 (A-U)**				
P1-1R5	-	NB	-	-	-	(NB)	-	(0.3±0.1)	(1.1±0.2)	(> 3)	-	(NB)	-	-	-

[a]Incubation buffer used: 200 mM NaCl, 0.5 mM EDTA, 20 mM MES at pH 6.0 or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 7.5 with or without 2 mM $MgCl_2$, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES at pH 8.0. The PAGE experiments were carried out at 4 °C. "-" indicates that the data were not obtained in this study. "NB" indicates no binding was observed at this condition. [b]The PNA oligomer sequences were previously studied.[11, 13, 58] The data shown in parentheses were for the gels done with unlabelled hairpins at 0.25 μM and with the gels post stained by EtBr. The errors shown are standard errors.

Both PAGE and 2-aminopurine fluorescence titration data reveal that PNAs P1-1R4 and P5-1R4 show no binding to dHP1 (**Table 1, Figures 4, S10, and S12e**). The dsRNA over dsDNA selectivity of the dbPNAs are consistent with previously reported results.[11-13, 15, 18, 58] It is likely

that a dsDNA has a wider and shallower major groove and a larger twist compared to a dsRNA, resulting in a preferential binding of a PNA to dsRNA over dsDNA.

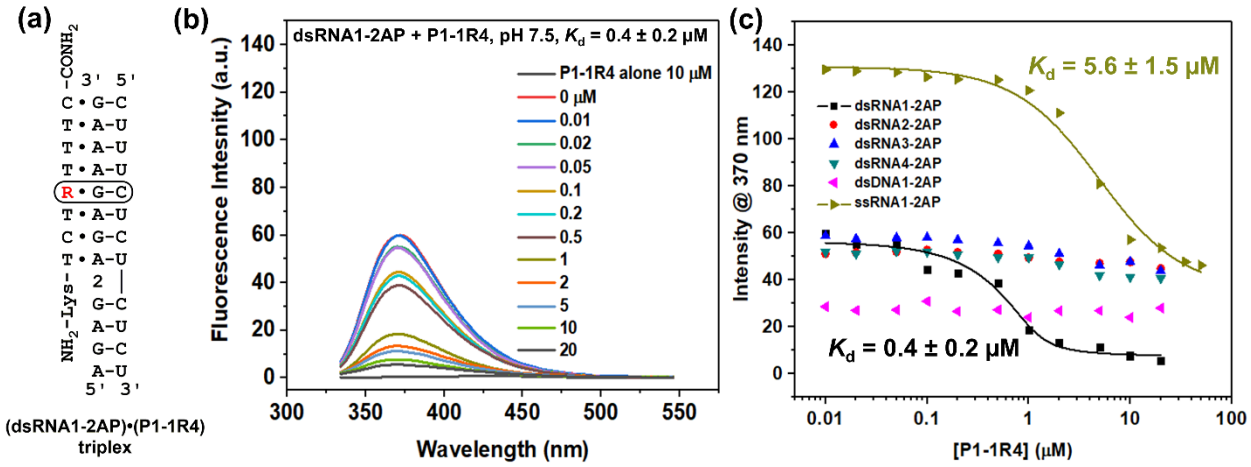


Figure 4. Fluorescence titration of R-modified P1-1R4 into 2-aminopurine (2AP) labelled dsRNAs, a dsDNA, and an ssRNA. (**a**) Triplex structure of (dsRNA1-2AP)·(P1-1R4). 2AP is shown as "2" in the triplex. (**b-c**) Fluorescence titration data. P1-1R4 binding to RNA duplex dsRNA1-2AP results in fluorescence quenching. The G-C pair highlighted in the box of dsRNA1-2AP (panel **a**) is replaced with a C-G, an A-U, and a U-A pair, respectively, in dsRNA2-2AP, dsRNA3-2AP, dsRNA4-2AP. dsDNA1-2AP is a DNA homolog of dsRNA1-2AP. Titration of P1-1R4 into the dsRNAs with (ds-RNA2-2AP, ds-RNA3-2AP, and ds-RNA4-2AP) and the dsDNA1-2AP results in no significant change in fluorescence intensity. ssRNA1-2AP fluorescence intensity is quenched upon duplex formation with P1-1R4 and indicates a 14-fold weaker binding than dsRNA1-2AP. All the samples were prepared with 1 µM RNA/DNA in the presence of 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 µM of PNA P1-1R4. For binding to ssRNA1-2AP, two more high-concentration points (35 and 50 µM) were added. The buffer contains 70 µL of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. The excitation wavelength is 303 nm. The measurements were done at room temperature.

Next, we incorporated the R residue at position 5 (P1-1R5) to further characterize the binding affinity and specificity (**Figure 2m**). PNA P1-1R5 contains three cytosine residues and may be expected to have a weakened binding with a relatively strong pH dependence due to the requirement of the formation of three C^+·G-C base triples (**Figures 1b and 2m**). Indeed, the PAGE

data show that PNA P1-R5 binds to the target rHP6 (**Figure 2f**) with a K_d of 1.1 ± 0.2 μM at 200 mM NaCl, pH 7.5 (**Figures S9c and S11o**). Again, P1-1R5 shows no observable binding to the hairpins with the corresponding G-C pair in rHP6 replaced with a C-G (rHP5), an A-U (rHP1-Cy3), or a U-A (rHP7) (**Figure S9**). Taken together, our data show that dbPNAs having a single R residue bind to dsRNAs highly sequence specifically.

Effects of sequence context, pH, and salt on PNA·RNA$_2$ triplex formation

PNA P1 contains three C residues and shows a significantly weakened binding (> 20 μM, **Table 1, Figure S8**) at 200 mM NaCl, pH 8.0. Surprisingly, both P1-1R4 (containing two C residues) and P1-1R5 (containing three C residues) show a reduced pH dependence compared to P1 in binding to target dsRNAs (**Table 1, Figures S8** and **S9**). Thus, incorporating R immediately next to C as present in P1-1R5 may be compatible for triplex formation. It is likely that C^+ and R^+ (guanidinium) have coupled stacking and hydrogen bonding interactions within a triplex structure, compensating for the expected charge-charge repulsion. The delocalized positive charges in R^+ and C^+ may also reduce charge-charge repulsion.

Interestingly, P5-1R4 and P5-3R show similar binding strength compared to P1-1R4 and P5 at various pH's (**Table 1**, **Figures 3**, **S8 and S11)**. It is possible that L and R residues at or near the terminal ends of a PNA show a relatively reduced stabilization in binding to a dsRNA.[11] Addition of 2 mM $MgCl_2$ moderately weakens the binding of P5-1R4 and P5-3R, suggesting that R residues stabilize the binding mainly through hydrogen bonding and stacking interactions (**Figures 1f,g,k,l and S1f**), but not nonspecific electrostatic interactions.

Incorporation of PNA R residues stabilizes and destabilizes PNA·RNA$_2$ triplex and PNA-RNA duplex, respectively

We further tested the binding of an R-modified PNA to a ssRNA. Our 2-aminopurine fluorescence titration data indicate that P1-1R4 shows binding to ssRNA1-2AP with a K_d of 5.6 ± 1.5 μM, which is significantly weaker than that for binding to dsRNA1-2AP (0.4 ± 0.2 μM) (**Figures 4 and S12a,f**). The dsRNA over ssRNA binding selectivity revealed by fluorescence titration is further corroborated with thermal melting experiments. The UV-absorbance detected thermal melting data show that P1-1R4 can form parallel and antiparallel duplexes with complementary ssRNAs (rHP1-ss-p and rHP1-ss-ap) with melting temperatures of 39.9 and 34.8 ^{o}C, respectively, which are comparatively lower than the duplexes formed by PNA P1 (44.2 and 48.5 ^{o}C) (**Figures 2 and 7**).[13] Importantly, P5-1R4 shows no melting transitions for both rHP1-ss-p and rHP1-ss-ap as compared to the melting temperature for P5 with rHP1-ss-p (31.5 ^{o}C) and rHP1-ss-ap (34.4 ^{o}C) (**Figures 2 and Figure 7**).[13] The data suggest that, compared to an L-modified PNA, an R-modified PNA is more unfavorable in binding to ssRNAs. Presumably, destabilization of PNA-RNA duplex formation results from an unfavorable Watson-Crick-like R-G pair (**Figure 1h**), similar to an unstable Watson-Crick-like L-G pair (**Figure S1c**).[11, 69] Taken together, our data clear show that incorporating R residues into PNAs facilitates structure specific binding to dsRNAs over ssRNAs.

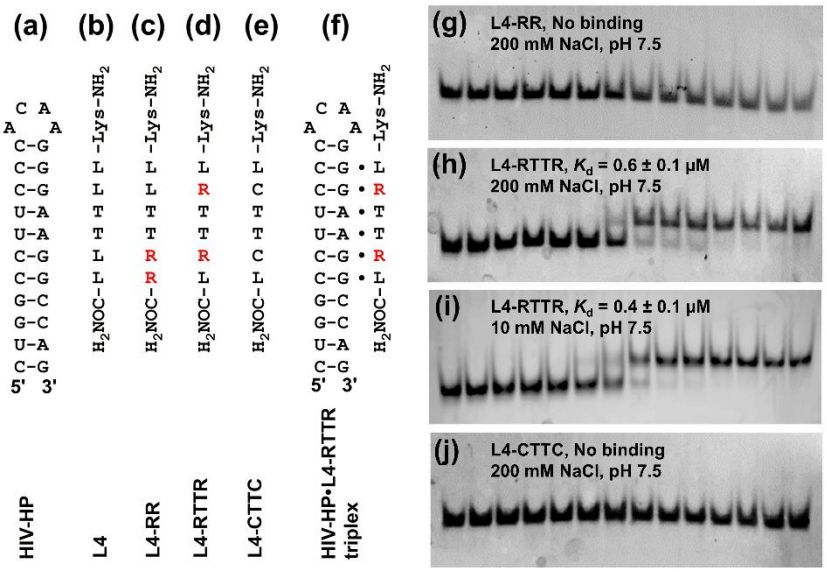


Figure 5. Nondenaturing PAGE characterization of PNAs binding to HIV-1 frameshift-inducing RNA (HIV-HP). The incubation buffer is 0.5 mM EDTA, 20 mM HEPES, pH 7.5, with 200 mM NaCl or 10 mM NaCl. (**a**-**f**) Sequences and structures of HIV-HP, PNAs, and HIV-HP·L4-RTTR triplex. (**g**-**j**) PAGE characterization for the binding of PNAs L4-RR, L4-RTTR, and L4-CTTC to HIV-HP. All the samples were prepared with 0.25 μM RNA (25 μL) with varied concentration of PNA (0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, and 20 μM, in lanes from left to right). Reducing the NaCl concentration from 200 mM to 10 mM results in a slight enhancement in binding for L4-RTTR (panels **h** and **i**), suggesting that the two R residues recognize G-C pairs with specific hydrogen bonding and stacking but not nonspecific ionic interactions.

Binding to HIV-1 minus-one ribosomal frameshift inducing RNA hairpin and influenza viral panhandle structure by dbPNAs containing multiple R residues

We tested the application of R- and L-modified dbPNAs for targeting HIV-1 frameshift site RNA hairpin[70-73] and influenza viral panhandle structure (PH-v).[55, 74-79] We made a series of R-containing PNAs targeting HIV-1 frameshift stimulatory RNA hairpin (HIV-HP, **Figure 5**). Our previous work shows that a 6-mer PNA L4 binds to HIV-HP with a K_d of 0.3 ± 0.1 μM (at 200 mM NaCl, pH 7.5).[58] Incorporating two non-consecutive R residues (L4-RTTR) into L4 does not affect the binding significantly (K_d = 0.6 ± 0.1 μM). Reducing the NaCl concentration from 200

mM to 10 mM results in a K_d of 0.4 μM for L4-RTTR (**Figure 5**), further suggesting that an R-modified PNA binds to a dsRNA through hydrogen bonding and stacking but not nonspecific ionic interactions. However, replacing the two R residues with C residues (L4-CTTC) results in no binding at both 200 mM and 10 mM NaCl (**Figures 5 and S11)**, suggesting the stabilizing effect of R residues. Interestingly, incorporating two consecutive R residues into L4 (L4-RR) causes no binding to HIV-HP at both 200 mM and 10 mM NaCl (**Figures 5 and S11)**, which suggests that consecutive R residues are not favorable for PNA·dsRNA triplex formation, presumably due to the unfavorable stacking of consecutive R residues. We speculate that the stacking pattern restricted by the helical structure of a PNA·dsRNA triplex does not allow the favorable self-stacking geometry of two guanidinium as observed in other systems.[23] Thus, R residues may be incorporated into PNAs immediately adjacent to T, C, and L but not another R for enhanced recognition of G-C pairs in dsRNAs.

Influenza A viruses contain eight viral RNA (vRNA) segments with highly conserved panhandle duplex structure formed between the first 13 nucleotides at the 5′ end and the last 12 nucleotides at the 3′ end.[55, 74-79] A 10-mer dbPNA IR-1 binds to PH-v with a K_d of 0.3 μM at 200 mM NaCl, pH 7.5.[55, 69] We replaced one and three L residues in IR-1 with R to obtain IR-1R and IR-3R, respectively (**Figure 6**). The PAGE data show that IR-1R and IR-3R bind to a Cy3-labelled PH-v (PH-v-Cy3) with K_d values of 0.2 ± 0.1 and 0.3 ± 0.1 μM, respectively (**Figure 6j,k)**. Notably, the K_d values of the dbPNAs are similar to that of an 11-mer antisense PNA (asPNA) (IR-ASP, K_d = 0.5 ± 0.1 μM, **Figure 6l**). However, asPNA IR-ASP binds to the single-stranded segment PH-v-ss with a thermal melting temperature of 59.0 ºC (**Figure 7**). Importantly, the thermal melting data show that IR-1R and IR-3R do not form a duplex with PH-v-ss (**Figure 7**). Thus, compared to

asPNAs, dbPNAs are more advantageous in sequence- and structure-specific recognition of dsRNAs over ssRNAs. Clearly, the newly developed PNA R residue can be compatible with T, C, L, and Q residues in dbPNAs for selectively targeting dsRNAs over ssRNAs.

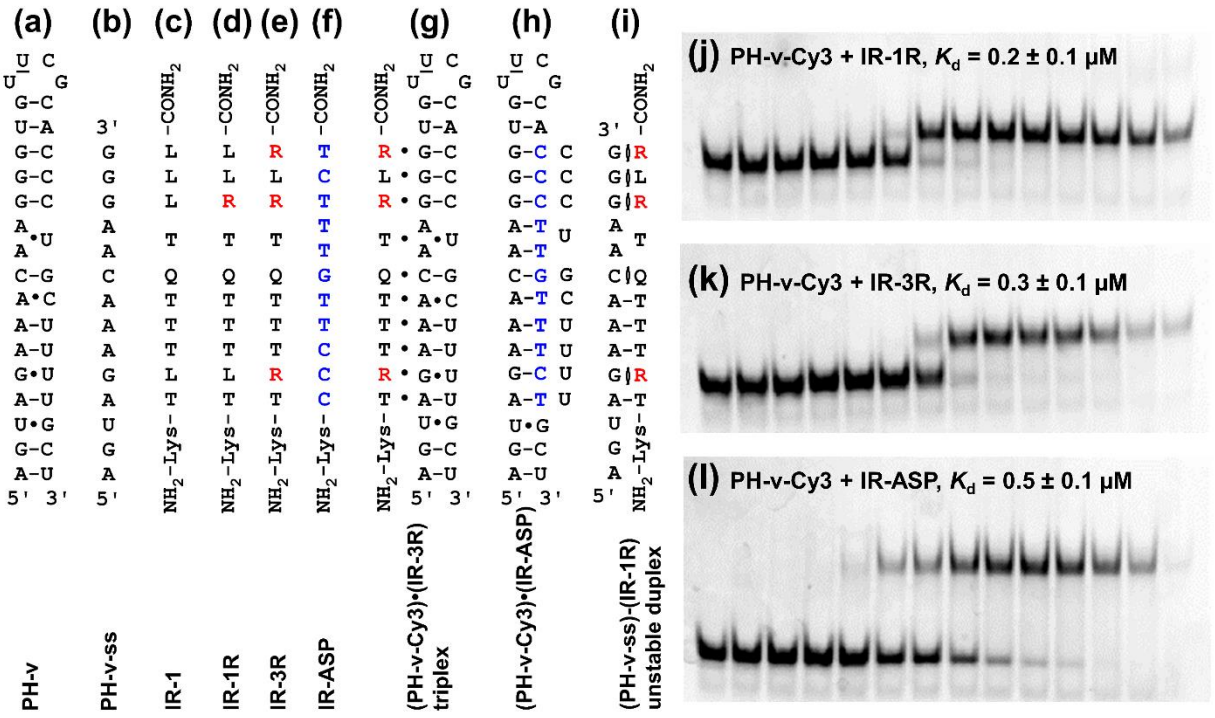


Figure 6. Nondenaturing PAGE characterization of the binding of dbPNAs and an asPNA to influenza viral panhandle structure. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. (**a,b**) Influenza panhandle viral RNA duplex (PH-v) and ssRNA segment of PH-v (PH-v-ss). The underlined U residue in PH-v is replaced with Cy3 dye to have PH-v-Cy3. (**c-f**) dbPNAs IR-1, IR-1R, IR-3R and asPNA IR-ASP designed for targeting PH-v. (**g** and **i**) A stable PH-v·IR-3R triplex and an unstable (PH-v-ss)-(IR-3R) duplex. PH-v but not the ssRNA segment of PH-v (PH-v-ss) is targeted by dbPNAs IR-1R and IR-3R. (**h**) A duplex invasion complex formed between PH-v and IR-ASP. (**j-l**) nondenaturing PAGE characterization. The dbPNAs IR-1R and IR-3R with L residues replaced with R at one or more positions show a similar binding affinity to PH-v-Cy3.[55, 69] The 10-mer dbPNAs IR-1R and IR-3R show a stronger binding affinity than the unmodified 11-mer asPNA IR-ASP. The Cy3-labelled RNA is at 50 nM (20 μL) with varied concentration of PNA (0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 5, and 10 μM of PNA in lanes from left to right).

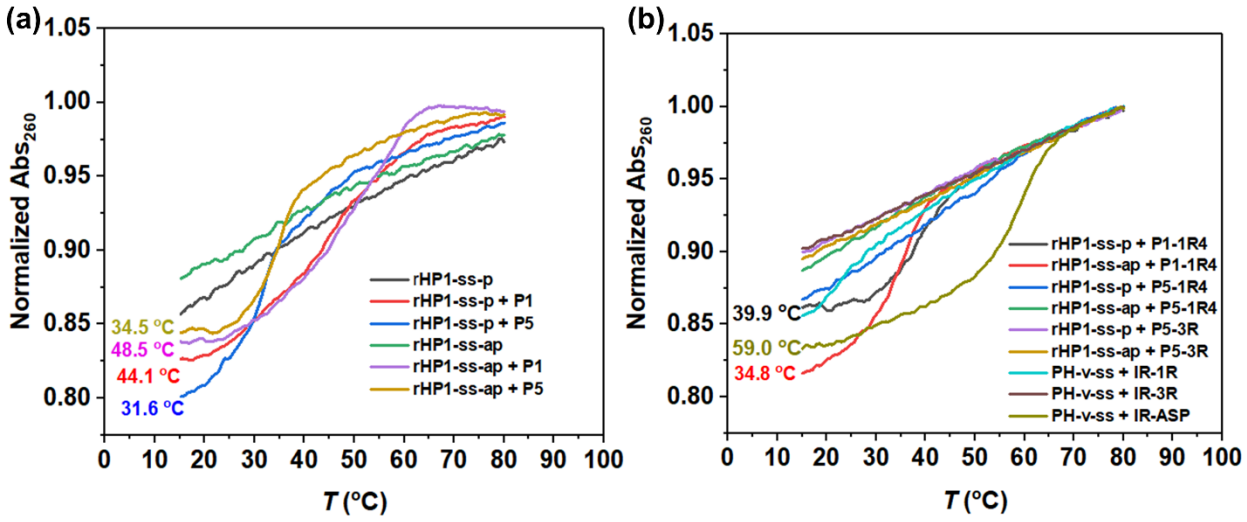


Figure 7. Thermal melting results for potential PNA-RNA duplexes. The heating curves are shown. The melting temperatures are shown for the curves with melting transitions. The ssRNAs and PNAs were at 5 µM (130 µL). The buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM NaH_2PO_4, pH 7.5. (**a**) Thermal melting results for control PNAs P1 and P5 with rHP1-ss-p and rHP1-ss-ap. Compared to unmodified P1, P5 contains L residues and shows reduced binding to ssRNAs, consistent with our previous work.[11] (**b**) Thermal melting results for R-modified PNAs and an asPNA IR-ASP. Incorporation of PNA R residues causes destabilization of PNA-RNA duplexes, due to unfavorable Watson-Crick-like base pairing, which is similar to PNA L and Q residues (see **Figures 1h and S1c,d**).

Cellular uptake of an R-modified PNA

It is known that incorporation of the guanidinium moiety facilitates enhanced cellular uptake of various cargos including PNAs.[13, 23, 38-49] We attached carboxyfluorescein to PNAs P1 (H_2N-Lys-TCTCTTTC-$CONH_2$,) and P5-3R (H_2N-Lys-TRTRTTTR-$CONH_2$) to obtain P1-CF and P5-3R-CF, respectively (**Table S1**), to assess how incorporation of R residues may affect cellular permeability. We incubated Sf9 cells with carboxyfluorescein-labelled P5-3R-CF and P1-CF for 12 hours. Interestingly, we observed fluorescence signal for P5-3R-CF mainly in cytoplasm (**Figure 8a,b,e,f**), although it is known that the fixation of the cells may perturb the cellular localization of the PNA. However, no significant fluorescence signal was observed for the cells

treated with P1-CF (**Figure 8c,d,g,h**), consistent with the fact that unmodified PNAs are not taken up by cells. P5-3R-CF is a derivative of P1-CF with the C residues replaced with R. Thus, incorporating R residues improves cell permeability.

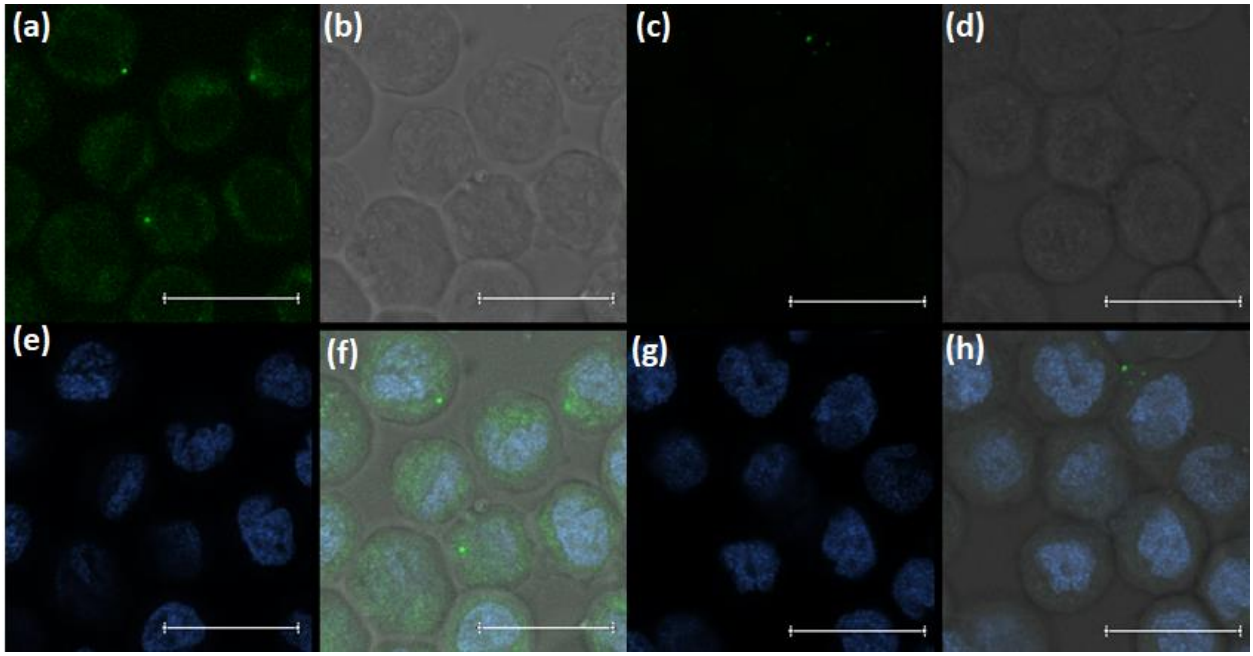


Figure 8. Cell uptake study of an R-modified PNA. Confocal microscopy images shown are for Sf9 cells treated with 2 μM (**a, b, e, f**) R-modified PNA P5-3R-CF (CF-H_2N-Lys-TCTCTTTC-$CONH_2$) and (**c, d, g, h**) control PNA P1-CF (CF-H_2N-Lys-TRTRTTTR-$CONH_2$). CF: caryboxyfluorescein. (**a, c**) Carboxyfluorescein imaging. Ex, 492nm; Em, 517nm. (**b, d**) Bright field imaging. (**e, g**) DAPI imaging. Ex, 358nm; Em, 461nm. (**f, h**) Merged images. Scale bar: 20 μm.

Our previous work shows that incorporating multiple Q but not L residues may not improve cellular permeability,[13, 55] although both Q and L have a guanidine moiety (**Figure 1**). It is possible that, compared to PNA residue L, which has the guanidine moiety within an aromatic base, Q and R have a positive charge and more optimal flexibility.[13, 39, 41, 47, 49] Both positive charges and optimal flexibility are critical for facilitating tight binding with the cell surface counter anions

(such as phosphate and sulfate groups present on cell membrane) for initiating cellular uptake.[43, 44, 49] Taken together, our results suggest that one may incorporate multiple R and/or Q residues into dbPNAs for an improved cellular permeability.

CONCLUSIONS

In summary, we have developed a simple guanidinium-based PNA residue R for the specific recognition of Watson-Crick G-C base pairs in RNA duplexes at physiological pH conditions. The synthetic scheme for the PNA R monomer involves a few facile and simple steps, which is advantageous compared to our previously developed PNA L monomer.[11] Our study exemplifies the potential broad scope of the newly developed R-modified dbPNAs in the sequence and structure specific targeting of dsRNAs over ssRNAs and dsDNAs. To the best of our knowledge, it is the first report on incorporating a simple guanidinium into a PNA scaffold as a pairing and stacking partner for the recognition of RNA G-C pairs. In addition, our dbPNAs are uniquely positioned on the major groove for molecular recognition. Our work shows that incorporating nonconsecutive R residues into our dbPNA platform containing other modified L and Q residues facilitates the pairing of R with G on the Hoogsteen edge (**Figure 1f,g,k,l**) and avoids other potential pairing geometries with any other natural bases including C on the Watson-Crick face with two hydrogen bond acceptors.[27, 28, 80] The presence of guanidinium group in single-stranded PNAs also serves to improve the cellular uptake (**Figure 8**), with potentially improved solubility and bioactivity. It is possible that one may combine multiple R and/or Q[13] residues together with other cell penetrating moiety such as neamine[55] for further improved cellular uptake and activities of dbPNAs.

AUTHOR INFORMATION

Corresponding Author

RNACHEN@ntu.edu.sg

ORCID:

Gang Chen: 0000-0002-8772-9755

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

This work was supported by NTU-A*STAR Seed Funding Research Award (2018), Singapore Ministry of Education (MOE) Tier 1 grant (RG152/17), and MOE Tier 2 grant (MOE2015-T2-1-028) to G.C.. The work was also supported by MOE Tier 1 grant (RG146/17) to Y.M.

ASSOCIATED CONTENT

Supporting Information

Additional information as noted in the text are supplied as Supporting Information. The Supporting Information is available free of charge on the ACS Publications website.

Synthetic methods of the PNA monomer and oligomers, additional modeling and experimental figures and tables, and NMR spectra of the compounds (PDF).

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] Disney, M. D., Yildirim, I., and Childs-Disney, J. L. (2014) Methods to enable the design of bioactive small molecules targeting RNA, *Org Biomol Chem 12*, 1029-1039.

[3] Kole, R., Krainer, A. R., and Altman, S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides, *Nat Rev Drug Discov 11*, 125-140.

[4] Kaczmarek, J. C., Kowalski, P. S., and Anderson, D. G. (2017) Advances in the delivery of RNA therapeutics: from concept to clinical reality, *Genome Med 9*, 60.

[5] Han, H., and Dervan, P. B. (1993) Sequence-specific recognition of double helical RNA and RNA.DNA by triple helix formation, *Proc Natl Acad Sci USA 90*, 3806-3810.

[6] Roberts, R. W., and Crothers, D. M. (1992) Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition, *Science 258*, 1463-1466.

[7] Escude, C., Francois, J. C., Sun, J. S., Ott, G., Sprinzl, M., Garestier, T., and Helene, C. (1993) Stability of triple helices containing RNA and DNA strands: experimental and molecular modeling studies, *Nucleic Acids Res 21*, 5547-5553.

[8] 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.

[9] Patil, K. M., and Chen, G. (2016) Recognition of RNA sequence and structure by duplex and triplex formation: targeting miRNA and pre-miRNA, In *Modified Nucleic Acids in Biology and Medicine,* (Erdmann, V. A., Jurga, S., and Barciszewski, J., Eds.), pp 299-317, Springer.

[10] 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 Visualized Exp 127*, e56221.

[11] 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.

[12] 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.

[13] Toh, D.-F. 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.

[14] Puah, R. Y., Jia, H., Maraswami, M., Toh, D.-F. K., Ero, R., Yang, L., Patil, K. M., Ong, A. A. L., Krishna, M. S., Sun, R., Tong, C., Huang, M., Chen, X., Loh, T. P., Gao, Y. G., Liu, D. X., and Chen, G. (2018) Selective Binding to mRNA Duplex Regions by Chemically Modified Peptide Nucleic Acids Stimulates Ribosomal Frameshifting, *Biochemistry 57*, 149-159.

[15] 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.

[16] 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.

[17] Muse, O., Zengeya, T., Mwaura, J., Hnedzko, D., McGee, D. W., Grewer, C. T., and Rozners, E. (2013) Sequence selective recognition of double-stranded RNA at physiologically relevant conditions using PNA-peptide conjugates, *ACS Chem Biol 8*, 1683-1686.
[18] 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.
[19] Wojciechowski, F., and Hudson, R. H. (2007) Nucleobase modifications in peptide nucleic acids, *Curr Top Med Chem 7*, 667-679.
[20] 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.
[21] Kotikam, V., Kennedy, S. D., MacKay, J. A., and Rozners, E. (2019) Synthetic, Structural, and RNA Binding Studies on 2-Aminopyridine-Modified Triplex-Forming Peptide Nucleic Acids, *Chem Eur J 25*, 4367-4372.
[22] 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.
[23] Vazdar, M., Heyda, J., Mason, P. E., Tesei, G., Allolio, C., Lund, M., and Jungwirth, P. (2018) Arginine "Magic": Guanidinium Like-Charge Ion Pairing from Aqueous Salts to Cell Penetrating Peptides, *Acc Chem Res 51*, 1455-1464.
[24] Gund, P. (1972) Guanidine, trimethylenemethane, and "Y-delocalization." Can acyclic compounds have "aromatic" stability?, *J. Chem. Educ. 49*, 100-103.
[25] Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy, *Science 257*, 76-80.
[26] Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins, *Proc Natl Acad Sci USA 73*, 804-808.
[27] Cheng, A. C., Chen, W. W., Fuhrmann, C. N., and Frankel, A. D. (2003) Recognition of nucleic acid bases and base-pairs by hydrogen bonding to amino acid side-chains, *J Mol Biol 327*, 781-796.
[28] Kondo, J., and Westhof, E. (2011) Classification of pseudo pairs between nucleotide bases and amino acids by analysis of nucleotide-protein complexes, *Nucleic Acids Res 39*, 8628-8637.
[29] Kim, J. H., Resende, R., Wennekes, T., Chen, H. M., Bance, N., Buchini, S., Watts, A. G., Pilling, P., Streltsov, V. A., Petric, M., Liggins, R., Barrett, S., McKimm-Breschkin, J. L., Niikura, M., and Withers, S. G. (2013) Mechanism-based covalent neuraminidase inhibitors with broad-spectrum influenza antiviral activity, *Science 340*, 71-75.
[30] Bottcher, T., Kolodkin-Gal, I., Kolter, R., Losick, R., and Clardy, J. (2013) Synthesis and activity of biomimetic biofilm disruptors, *J Am Chem Soc 135*, 2927-2930.
[31] McKeever, C., Kaiser, M., and Rozas, I. (2013) Aminoalkyl derivatives of guanidine diaromatic minor groove binders with antiprotozoal activity, *J Med Chem 56*, 700-711.
[32] Takayama, K., Mori, K., Taketa, K., Taguchi, A., Yakushiji, F., Minamino, N., Miyazato, M., Kangawa, K., and Hayashi, Y. (2014) Discovery of selective hexapeptide agonists to human neuromedin U receptors types 1 and 2, *J Med Chem 57*, 6583-6593.
[33] Wu, H., Esteve, E., Tremaroli, V., Khan, M. T., Caesar, R., Manneras-Holm, L., Stahlman, M., Olsson, L. M., Serino, M., Planas-Felix, M., Xifra, G., Mercader, J. M., Torrents, D., Burcelin, R., Ricart, W., Perkins, R., Fernandez-Real, J. M., and Backhed, F. (2017) Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug, *Nat Med 23*, 850-858.
[34] Heinzl, G. A., Huang, W., Yu, W., Giardina, B. J., Zhou, Y., MacKerell, A. D., Jr., Wilks, A., and Xue, F. (2016) Iminoguanidines as Allosteric Inhibitors of the Iron-Regulated Heme Oxygenase (HemO) of Pseudomonas aeruginosa, *J Med Chem 59*, 6929-6942.
[35] Saczewski, F., and Balewski, L. (2009) Biological activities of guanidine compounds, *Expert Opin Ther Pat 19*, 1417-1448.

[36] Holmes, T. C., May, A. E., Zaleta-Rivera, K., Ruby, J. G., Skewes-Cox, P., Fischbach, M. A., DeRisi, J. L., Iwatsuki, M., Ōmura, S., and Khosla, C. (2012) Molecular Insights into the Biosynthesis of Guadinomine: A Type III Secretion System Inhibitor, *J Am Chem Soc 134*, 17797-17806.

[37] von Itzstein, M., Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Phan, T. V., Smythe, M. L., White, H. F., Oliver, S. W., Colman, P. M., Varghese, J. N., Michael Ryan, D., Woods, J. M., Bethell, R. C., Hotham, V. J., Cameron, J. M., and Penn, C. R. (1993) Rational design of potent sialidase-based inhibitors of influenza virus replication, *Nature 363*, 418-423.

[38] Gupta, P., Muse, O., and Rozners, E. (2012) Recognition of double-stranded RNA by guanidine-modified peptide nucleic acids, *Biochemistry 51*, 63-73.

[39] Thomas, S. M., Sahu, B., Rapireddy, S., Bahal, R., Wheeler, S. E., Procopio, E. M., Kim, J., Joyce, S. C., Contrucci, S., Wang, Y., Chiosea, S. I., Lathrop, K. L., Watkins, S., Grandis, J. R., Armitage, B. A., and Ly, D. H. (2013) Antitumor effects of EGFR antisense guanidine-based peptide nucleic acids in cancer models, *ACS Chem Biol 8*, 345-352.

[40] 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.

[41] Lu, X. W., Zeng, Y., and Liu, C. F. (2009) Modulating the hybridization property of PNA with a peptoid-like side chain, *Org Lett 11*, 2329-2332.

[42] 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.

[43] Patil, K. M., Naik, R. J., Rajpal, Fernandes, M., Ganguli, M., and Kumar, V. A. (2012) Highly efficient (R-X-R)-type carbamates as molecular transporters for cellular delivery, *J Am Chem Soc 134*, 7196-7199.

[44] 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.

[45] 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.

[46] Du, S., Liew, S. S., Li, L., and Yao, S. Q. (2018) Bypassing Endocytosis: Direct Cytosolic Delivery of Proteins, *J Am Chem Soc 140*, 15986-15996.

[47] 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.

[48] Gait, M. J., Arzumanov, A. A., McClorey, G., Godfrey, C., Betts, C., Hammond, S., and Wood, M. J. A. (2019) Cell-Penetrating Peptide Conjugates of Steric Blocking Oligonucleotides as Therapeutics for Neuromuscular Diseases from a Historical Perspective to Current Prospects of Treatment, *Nucleic Acid Ther 29*, 1-12.

[49] Wexselblatt, E., Esko, J. D., and Tor, Y. (2014) On guanidinium and cellular uptake, *J Org Chem 79*, 6766-6774.

[50] Huang, L., Wang, J., Wilson, T. J., and Lilley, D. M. J. (2017) Structure of the Guanidine III Riboswitch, *Cell Chem Biol 24*, 1407-1415 e1402.

[51] Sherlock, M. E., Malkowski, S. N., and Breaker, R. R. (2017) Biochemical Validation of a Second Guanidine Riboswitch Class in Bacteria, *Biochemistry 56*, 352-358.

[52] Wilds, C. J., Maier, M. A., Manoharan, M., and Egli, M. (2003) Structural basis for recognition of guanosine by a synthetic tricyclic cytosine analogue: Guanidinium G-clamp, *Helv Chim Acta 86*, 966-978.

[53] Fox, K. R., and Brown, T. (2011) Formation of stable DNA triplexes, *Biochem Soc Trans 39*, 629-634.

[54] 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.

[55] Kesy, J., Patil, K. M., Kumar, S. M., Shu, Z., Yee, Y. H., Zimmermann, L., Ong, A. A. L., Toh, D.-F. K., Krishna, M. S., Yang, L., Decout, J. L., Luo, D., Prabakaran, M., Chen, G., and Kierzek, E. (2019) A short chemically modified dsRNA-binding PNA (dbPNA) inhibits influenza viral replication by targeting viral RNA panhandle structure, *Bioconjugate Chem 30*, 931-943.

[56] 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.

[57] Ong, A. A. L., Toh, D. F. K., Patil, K. M., Meng, Z., Yuan, Z., Krishna, M. S., Devi, G., Haruehanroengra, P., Lu, Y., Xia, K., Okamura, K., Sheng, J., and Chen, G. (2019) General recognition of U-G, U-A, and C-G pairs by double-stranded RNA-binding PNAs incorporated with an artificial nucleobase, *Biochemistry 58*, 1319-1331.

[58] Patil, K. M., Toh, D.-F. K., Yuan, Z., Meng, Z., Shu, Z., Zhang, H., Ong, A. A. L., Krishna, M. S., Lu, L., Lu, Y., and Chen, G. (2018) Incorporating uracil and 5-halouracils into short peptide nucleic acids for enhanced recognition of A-U pairs in dsRNAs, *Nucleic Acids Res 46*, 7506-7521.

[59] Zgarbova, M., Otyepka, M., Sponer, J., Mladek, A., Banas, P., Cheatham, T. E., III, 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.

[60] Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham, T. E., III, Laughton, C. A., and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers, *Biophys J 92*, 3817-3829.

[61] 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.

[62] Jakalian, A., Jack, D. B., and Bayly, C. I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *J Comput Chem 23*, 1623-1641.

[63] Mark, P., and Nilsson, L. (2001) Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K, *J Phys Chem A 105*, 9954-9960.

[64] Kresse, G., and Hafner, J. (1994) Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium, *Phys Rev B 49*, 14251-14269.

[65] Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L. P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., and Pande, V. S. (2017) OpenMM 7: Rapid development of high performance algorithms for molecular dynamics, *PLoS Comput Biol 13*, e1005659.

[66] McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernandez, C. X., Schwantes, C. R., Wang, L. P., Lane, T. J., and Pande, V. S. (2015) MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories, *Biophys J 109*, 1528-1532.

[67] Thomson, S. A., Josey, J. A., Cadilla, R., Gaul, M. D., Fred Hassman, C., Luzzio, M. J., Pipe, A. J., Reed, K. L., Ricca, D. J., Wiethe, R. W., and Noble, S. A. (1995) Fmoc mediated synthesis of Peptide Nucleic Acids, *Tetrahedron 51*, 6179-6194.

[68] Bernatowicz, M. S., Wu, Y., and Matsueda, G. R. (1993) Urethane protected derivatives of 1-guanylpyrazole for the mild and efficient preparation of guanidines, *Tetrahedron Lett 34*, 3389-3392.

[69] Krishna, M. S., Toh, D. F. K., Meng, Z., Ong, A. A. L., Wang, Z., Lu, Y., Xia, K., Prabakaran, M., and Chen, G. (2019) Sequence- and structure-specific probing of RNAs by short nucleobase-modified dsRNA-binding PNAs incorporating a fluorescent light-up uracil analog, *Anal Chem 91*, 5331-5338.

[70] Baril, M., Dulude, D., Gendron, K., Lemay, G., and Brakier-Gingras, L. (2003) Efficiency of a programmed –1 ribosomal frameshift in the different subtypes of the human immunodeficiency virus type 1 group M, *RNA 9*, 1246-1253.

[71] Garcia-Miranda, P., Becker, J. T., Benner, B. E., Blume, A., Sherer, N. M., and Butcher, S. E. (2016) Stability of HIV Frameshift Site RNA Correlates with Frameshift Efficiency and Decreased Virus Infectivity, *J Virol 90*, 6906-6917.

[72] Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Jr., Swanstrom, R., Burch, C. L., and Weeks, K. M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome, *Nature 460*, 711-716.

[73] Gareiss, P. C., and Miller, B. L. (2009) Ribosomal frameshifting: an emerging drug target for HIV, *Curr Opin Investig Drugs 10*, 121-128.

[74] Lenartowicz, E., Kesy, J., Ruszkowska, A., Soszynska-Jozwiak, M., Michalak, P., Moss, W. N., Turner, D. H., Kierzek, R., and Kierzek, E. (2016) Self-Folding of Naked Segment 8 Genomic RNA of Influenza A Virus, *PLoS One 11*, e0148281.

[75] Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M., and Palese, P. (1987) Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle, *Proc Natl Acad Sci USA 84*, 8140-8144.

[76] Bae, S. H., Cheong, H. K., Lee, J. H., Cheong, C., Kainosho, M., and Choi, B. S. (2001) Structural features of an influenza virus promoter and their implications for viral RNA synthesis, *Proc Natl Acad Sci USA 98*, 10602-10607.

[77] Gultyaev, A. P., Fouchier, R. A. M., and Olsthoorn, R. C. L. (2010) Influenza Virus RNA Structure: Unique and Common Features, *Int Rev Immunol 29*, 533-556.

[78] Medina, R. A., Rojas, M., Tuin, A., Huff, S., Ferres, M., Martinez-Valdebenito, C., Godoy, P., Garcia-Sastre, A., Fofanov, Y., and SantaLucia, J., Jr. (2011) Development and characterization of a highly specific and sensitive SYBR green reverse transcriptase PCR assay for detection of the 2009 pandemic H1N1 influenza virus on the basis of sequence signatures, *J Clin Microbiol 49*, 335-344.

[79] Te Velthuis, A. J. W., Long, J. C., Bauer, D. L. V., Fan, R. L. Y., Yen, H. L., Sharps, J., Siegers, J. Y., Killip, M. J., French, H., Oliva-Martin, M. J., Randall, R. E., de Wit, E., van Riel, D., Poon, L. L. M., and Fodor, E. (2018) Mini viral RNAs act as innate immune agonists during influenza virus infection, *Nat Microbiol 3*, 1234-1242.

[80] Chen, Y., Yang, F., Zubovic, L., Pavelitz, T., Yang, W., Godin, K., Walker, M., Zheng, S., Macchi, P., and Varani, G. (2016) Targeted inhibition of oncogenic miR-21 maturation with designed RNA-binding proteins, *Nat Chem Biol 12*, 717-723.

TOC figure