Published online 4 September 2016 *Nucleic Acids Research, 2016, Vol. 44, No. 19* **9071–9082**
doi: 10.1093/nar/gkw778

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

Desiree-Faye Kaixin Toh†, Gitali Devi†, Kiran M. Patil, Qiuyu Qu, Manikantha Maraswami, Yunyun Xiao, Teck Peng Loh, Yanli Zhao* and Gang Chen*

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

Received July 31, 2015; Revised August 20, 2016; Accepted August 24, 2016

ABSTRACT

RNA duplex regions are often involved in tertiary interactions and protein binding and thus there is great potential in developing ligands that sequence-specifically bind to RNA duplexes. We have developed a convenient synthesis method for a modified peptide nucleic acid (PNA) monomer with a guanidine-modified 5-methyl cytosine base. We demonstrated by gel electrophoresis, fluorescence and thermal melting experiments that short PNAs incorporating the modified residue show high binding affinity and sequence specificity in the recognition of an RNA duplex containing an internal inverted Watson-Crick C-G base pair. Remarkably, the relatively short PNAs show no appreciable binding to DNA duplexes or single-stranded RNAs. The attached guanidine group stabilizes the base triple through hydrogen bonding with the G base in a C-G pair. Selective binding towards an RNA duplex over a single-stranded RNA can be rationalized by the fact that alkylation of the amine of a 5-methyl C base blocks the Watson–Crick edge. PNAs incorporating multiple guanidine-modified cytosine residues are able to enter HeLa cells without any transfection agent.

INTRODUCTION

Ribonucleic acids (RNAs) are crucial to all life, playing important roles in numerous biological functions including protein coding, catalysis, and gene regulation for example. Various RNA-based therapeutic approaches have been investigated to regulate disease-causing interactions and biological functions involving RNA. To this end, many attempts have been reported on the development of small molecule ligands for probing and targeting non-Watson–Crick paired regions in RNA (1,2). Targeting RNA structures with small molecules, however, might still be challenging due to the limited surface contact area and the presence of nonspecific interactions. Short oligonucleotides have seen a wide range of applications including the traditional antisense therapeutics and microarray technologies that act through the formation of Watson–Crick duplexes with the RNA targets (3–6). Chemically modified DNA/RNA oligonucleotides and artificial nucleic acids have been developed to enhance the binding affinity and sequence specificity of duplex formation. The strategy of using oligonucleotides to target RNA through duplex formation, however, may be limited to the unstructured or partially structured regions of RNA. The structured regions in the RNA targets may significantly slow down the binding kinetics and weaken the binding strength of oligonucleotides (7).

Alternatively, through the formation of a highly sequence specific triplex structure, an oligonucleotide may bind to an RNA duplex without disrupting the pre-existing structure of the RNA target (8–10). For example, even though the major groove is relatively deep and narrow, an RNA double helix can accommodate a third strand in the major groove to form a modestly stable triplex (11–13). In a major-groove triplex, the pyrimidine bases in the third strand recognize the sequence of the purine strand of the double helical RNA by Hoogsteen hydrogen bonding, forming U·A-U or T·A-U and C^+·G-C base triples (Figure 1A and B) (8–10). Chemically modified triplex-forming oligonucleotides (TFOs) and peptide nucleic acids (PNAs) are promising in enhancing the recognition of the structured RNAs through TFO·RNA_2 and PNA·RNA_2 triplex formation, respectively (9,10). Compared to TFOs, which typically have a negatively charged phosphate backbone, PNAs have a neu-

*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 Yanli Zhao. Tel: +65 6316 8792; Fax: +65 6791 1961; Email: zhaoyanli@ntu.edu.sg
†These two authors contributed equally to this work as first authors.

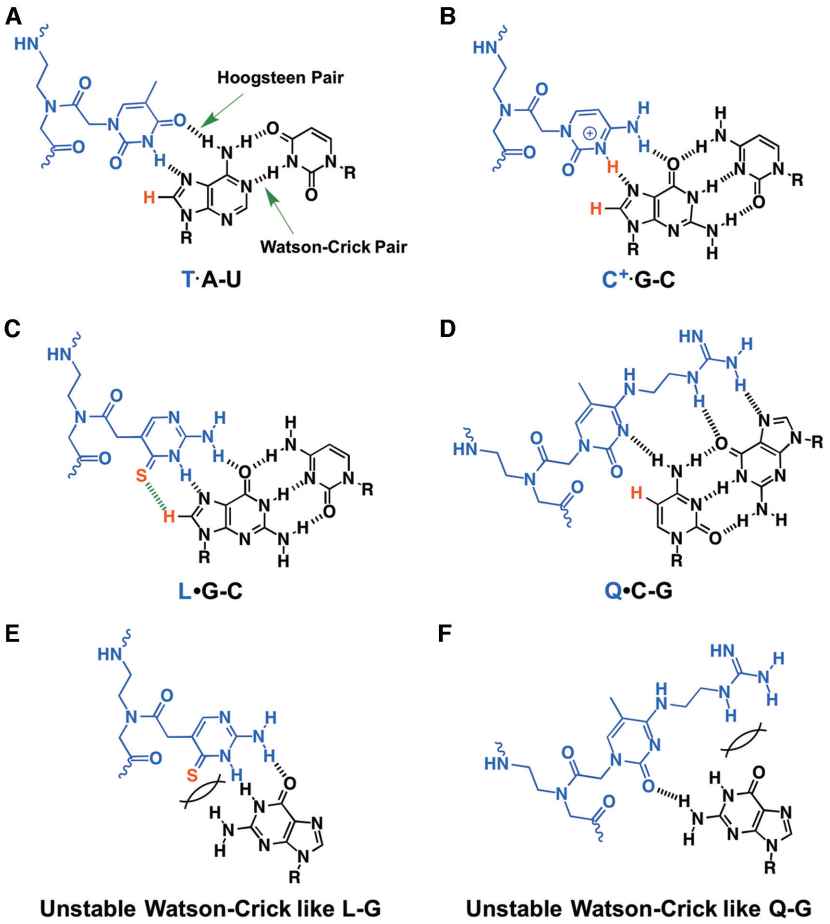


Figure 1. Chemical structures of (**A**) T·A-U, (**B**) C^+·G-C, (**C**) L·G-C and (**D**) Q·C-G major-groove base triples, and (**E**) unstable Watson–Crick-like L-G and (**F**) unstable Watson–Crick-like Q-G pairs. The letter R represents the sugar-phosphate backbone of RNA. The hydrogen bonds are indicated by black dashed lines. A relatively low pH is needed to protonate C to form the C^+·G-C triple (panel B). The enhanced van der Waals interaction between the sulfur atom in L base and H8 in G base (panel C) is indicated by a green dashed line. The 5-methyl group in Q base (panels D and F) may repel the two-carbon linker away and favor the Q·C-G base triple formation, but destabilize a Watson–Crick-like Q-G pair (panel F) due to the steric clash caused by the linker.

tral and chemically stable backbone (14), which facilitates enhanced sequence specific recognition of RNA duplexes.

Through the incorporation of modified bases, it is possible to develop PNAs that selectively bind to the duplex over the single-stranded regions of RNA (9,15–20). For example, we have shown previously (15) that PNAs incorporating a modified neutral base thio-pseudoisocytosine (L) (Figure 1C) can recognize a Watson–Crick G-C pair in an RNA duplex by forming a Hoogsteen L·G pair with improved stability and minimized pH dependence compared to a Hoogsteen C^+·G pair (Figure 1B). The enhanced Hoogsteen L·G pair formation is presumably due to improved van der Waals interaction between the thio group in L and H8 in G, as well as base stacking and hydrogen bonding interactions (9,13,15). PNAs incorporating L-modified residues show selective binding to RNA duplexes over single-stranded RNAs, because a Watson–Crick-like G-L pair (Figure 1E) is destabilized compared to a Watson–Crick G-C pair due to the steric clash between the relatively bulky thio group in L and the amino group in G. Furthermore, PNAs show enhanced binding to RNA duplexes over DNA duplexes (9,15–20). Thus, triplex-forming PNAs may become useful therapeutic agents that not only allow the stabilization of desired RNA duplexes, but also inhibit undesired tertiary interactions and protein binding (8,9,21,22). Compared to small molecule binders, triplex-forming PNAs are programmable and can, in principle, recognize any sequence of RNA duplex structures.

A purine tract in one of the two strands of an RNA duplex (see Figure 2A and C for example) is needed for a standard pyrimidine motif major-groove triplex. In a pyrimidine motif major-groove PNA·RNA_2 triplex, the PNA strand is parallel to the purine strand in the RNA duplex, i.e. the N-terminus of PNA is aligned with the 5′ end of the purine strand of RNA duplex (Figure 2G). However, in natural RNA duplexes, the purine tract in one of the two strands is often interrupted by pyrimidine residues, resulting in the formation of inverted Watson–Crick C-G and U-A pairs (pyrimidine-purine inversions) (see Figure 2B and D for hairpin structures with the pyrimidine–purine inversions highlighted). Pyrimidine–purine inversions are present in many functional structured RNAs such as pre-microRNAs and the HIV-1 –1 ribosomal frameshift-inducing hairpin (Figure 2R), which are important therapeutic targets (2,23).

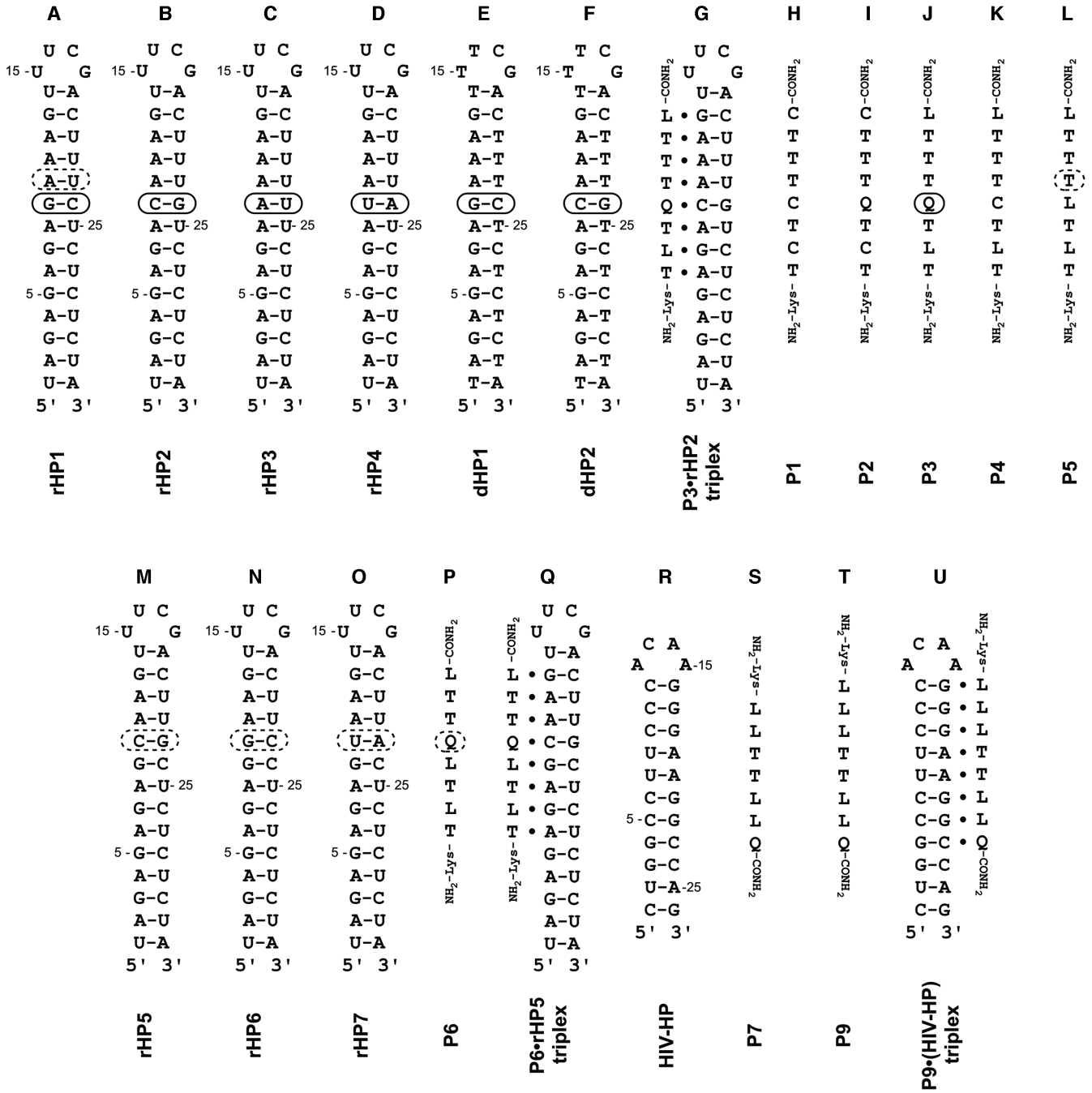


Figure 2. DNA and RNA hairpins and PNAs studied in this paper. Binding specificities were studied between the RNA/DNA base pairs and PNA residues highlighted with boxes. (**A–D**) Model RNA hairpins rHP1, rHP2, rHP3 and rHP4. (**E** and **F**) Model DNA hairpins dHP1 and dHP2. (**G**) A model PNA·RNA_2 triplex formed between PNA P3 and rHP2. (**H**–**L**) PNAs tested for binding to hairpins shown in panels A–F. (**M–Q**) Model RNA hairpins and PNA P6 used for studying triplex formation. (**R–U**) HIV-1 frameshift-inducing hairpin and PNAs for testing the triplex formation.

Previously designed PNAs with modified nucleobases, such as 5-methylisocytosine (iC), 2-pyrimidinone (P), and 3-oxo-2,3-dihydropyridazine (E) show improvement in the binding affinity and selectivity toward the RNA duplexes with pyrimidine–purine inversions (16,17,20). As observed in protein recognition of nucleic acid duplexes (24), the suboptimal binding affinity and selectivity of the reported triplex-forming PNAs are likely due to the fact that the C and U bases, which have only one functional group in the major groove of a duplex, are involved in hydrogen bonding interactions for the triplex formation.

Extensive research has been done to design and synthesize TFOs with modified nucleobases to enhance the recognition of C-G inversions in DNA duplexes (25–28). A guanidine functionalized C base, N^4-(2-guanidoethyl)-5-methylcytosine (27) has been previously incorporated

into relatively long TFOs (17-mer) with 2′-O-methyl (2′-OMe) and 2′-O-aminoethoxy (2′-AE) RNA backbones for TFO·DNA_2 triplex formation. Here, we developed a synthetic route based on previously reported methods (14,27,29,30) for a novel PNA monomer containing the modified C base, N^4-(2-guanidoethyl)-5-methylcytosine (abbreviated as Q in Figure 1D). We studied the binding of Q-modified short (7-mer or 8-mer) PNAs to RNA and DNA duplexes and single-stranded RNAs by gel electrophoresis, fluorescence, and thermal melting methods.

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 MHz, 400 MHz (100 MHz, ^{13}C) and 500 MHz Bruker spectrometers. 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: δ 7.24; dimethyl sulfoxide-d_6: δ 2.50; methanol-d_4: δ 3.31) and ^{13}C (chloroform-d: δ77.0; dimethyl sulfoxide-d_6: δ 39.5; methanol-d_4: δ 49.1) NMR spectra. Mass spectra of the compounds were obtained via liquid-chromatography-mass spectroscopy fitted 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 monomer Q

The detailed synthesis procedures for the PNA monomer Q are shown in Supplementary Material. The PNA monomer Q was synthesized from a series of reactions shown in Scheme 1 based on the previously reported methods (14,27,29,30). Thymine was first reacted with ethyl bromoacetate in the presence of potassium carbonate to yield compound **2**. The C^4 carbonyl group in **2** was activated to form the triazole derivative **3** through the reaction with 1,2,4-triazole and phosphorous oxychloride. Reaction of **3** with *N*-Boc-ethylenediamine subjected to the intermediate **4**. Upon deprotection of the boc protecting group using 50% TFA in DCM, the guanidine functional group was attached to the free amine to yield **5**. Hydrolysis of the ethyl group at N^1 position using aqueous lithium hydroxide and hydrochloric acid yielded **6**, after which it was reacted with PNA backbone ethyl *N*-(2-Boc-aminoethyl)glycinate in the presence of EDC·HCl and DIPEA to give **7**. Finally, hydrolysis of the ester group on the PNA backbone gave the desired PNA monomer **8**.

Synthesis of PNA oligomers

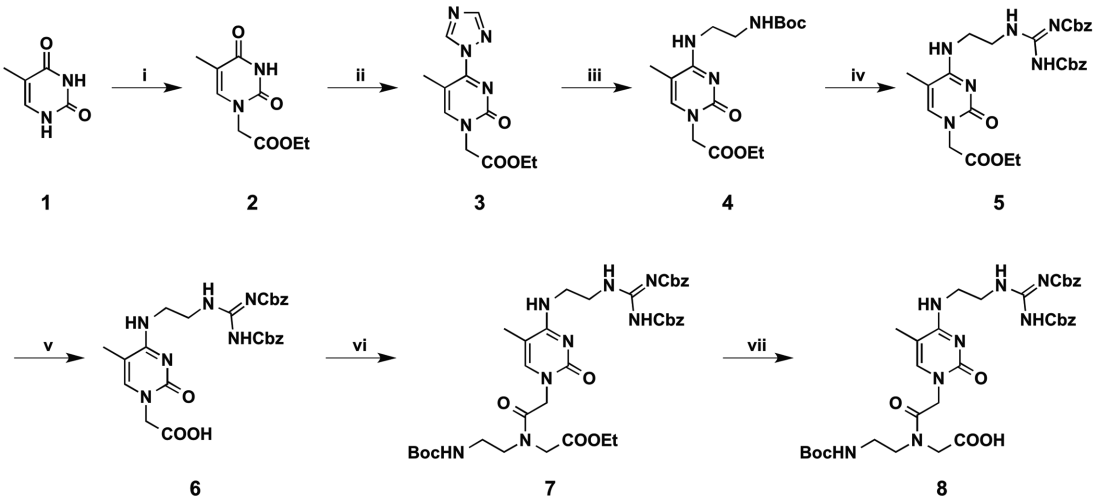
The PNA (*aeg*PNA) monomers were purchased from ASM Research Chemicals. PNA monomer L was synthesized following the reported method (15). PNA oligomers were synthesized manually using Boc chemistry via a Solid-Phase Peptide Synthesis protocol. 4-Methylbenzhydrylamine hydrochloride (MBHA·HCl) polystyrene resins were used. The loading value used for the synthesis of the oligomers was 0.35 mmol/g and acetic anhydride was used as the capping reagent. (Benzotriazol-1-yloxy)tripyrrolidinophosphonium 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 water and purified by reverse phase-high-performance liquid chromatography (RP-HPLC) using water–ACN–0.1% TFA as the mobile phase. For the PNAs containing Cy3 dye, the PNAs were first functionalized with an alkyne group through the addition of a propargyl glycine unit at the N-terminus. The PNAs were then subsequently reacted with an azide-containing Cy3 fluorescent dye via copper-catalyzed click chemistry (31). Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) analysis was used to characterize the oligomers, with the use of α-cyano-4-hydroxycinnamic acid (CHCA) as the sample crystallization matrix.

Non-denaturing polyacrylamide gel electrophoresis

Non-denaturing (12%) polyacrylamide gel electrophoresis (PAGE) experiments were conducted with 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), with and without 2 mM $MgCl_2$, or 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES (pH 8.0). The loading volume for each sample was 10 or 20 μl. 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 gel was run at 4°C at 250 V for 5 h. The gels were then stained with ethidium bromide and imaged by the Typhoon Trio Variable Mode Imager.

Fluorescence binding study

Fluorescence emission spectra were recorded on a Varian Cary Eclipse fluorescence spectrophotometer at room temperature using a 1 cm square cuvette. All samples contain 1 μM 2-aminopurine-labeled single-stranded RNA (ssRNA) or double-stranded RNA (dsRNA) in 70 μl of incubation buffer. The PNA concentration ranges from 0.2 to 50 μM. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5 or pH 8.0. The samples containing the RNA duplex and PNA were prepared by slow cooling of the RNA duplex from 95°C to room temperature, followed



Scheme 1. Synthesis of PNA Q monomer: (i) ethyl bromoacetate, K_2CO_3, dry DMF, rt., overnight. (ii) 1,2,4-triazole, $POCl_3$, TEA, dry ACN/DCM, 0°C to rt., 21 h (30). (iii) *t*-Butyl (2-aminoethyl)carbamate, K_2CO_3, ACN, rt., 20 h (27). (iv) 50% TFA in DCM, 1 h, TEA, diCbz-protected 1-guanyl pyrazole (27,29). (v) aq. 1 M LiOH, THF, 2 M HCl, 0°C to rt., 0.5 h. (vi) Ethyl *N*-(2-Boc-aminoethyl)glycinate, EDC·HCl, DIPEA, rt., 12 h. vii) aq 1 M LiOH, THF, 2 M HCl, 0°C to rt., 0.5 h.

by annealing with PNA oligomers at room temperature for 1–2 h and incubation at 4°C overnight. The samples containing the single-stranded RNA and PNA were annealed by slow cooling from 95°C to room temperature, followed by incubation at 4°C overnight. The emission spectra of 2-aminopurine were measured over a wavelength range of 330–550 nm with an excitation wavelength of 303 nm.

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. All samples contain 5 μM RNA with and without 5 μM PNA in 130 μl buffer. The buffer is 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°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 fit of the first derivative of the curves.

Confocal microscopy studies

HeLa cells were seeded on cover slides in 6-well plates at a density of 1×10^5 per well and grown in DMEM medium at 37°C for 24 h. Then, the cells were treated with different PNA samples, respectively, for 24 h. The medium was removed and the slides were washed with PBS buffer for three times. Then the cells were fixed with 4.0% formaldehyde at 4°C for 15 min. Finally, the cells on cover slides were fixed and observed.

RESULTS AND DISCUSSION

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

The PNA monomer Q was synthesized based on previously reported methods (Scheme 1) (14,27,29,30) with the detailed procedures shown in Supplementary Material (Supplementary Figures S1–S7). All the PNAs were purified by RP-HPLC and characterized by MALDI-TOF (Supplementary Table S1 and Figure S8). We carried out 12% non-denaturing PAGE experiments (Figure 3, Supplementary Figures S9–S15) to determine the binding affinity of a series of short (7-mer or 8-mer) PNAs towards various RNA and DNA hairpins (Figure 2, Table 1).

The guanidine group in the Q monomer, incorporated via a two-carbon linker, is predicted to form two hydrogen bonds with the G base in a C-G pair (Figure 1D) (27,28). Thus, the incorporation of Q residues into triplex-forming PNAs may improve the binding affinity and sequence specificity to an RNA duplex with an inverted Watson–Crick C-G pair (Figure 2B and G). To enhance the PNA·RNA_2 triplex formation at near-physiological conditions, we incorporated a previously developed PNA monomer thiopseudoisocytosine (L) (Figure 1C) (15) into a series of short PNAs (Figure 2J–L).

The non-denaturing PAGE result suggests that PNA P3 (NH_2-Lys-TLTQTTTL-$CONH_2$, see Figure 2J), which has both Q and L modifications, binds to rHP2 ($K_d = 4.7 \pm 0.6$ μM) to form a PNA·RNA_2 triplex (Figure 2B and G) in a near physiological buffer (200 mM NaCl, pH 7.5) (Figure 3B, Table 1). Remarkably, PNA P3 shows no or weak binding to other RNA hairpins with one C-G pair replaced by G-C (rHP1), A-U (rHP3) or U-A (rHP4) (Figures 2 and 3, Table 1, Supplementary Figures S9 and S13). In addition, PNA P3 does not show binding to the DNA hairpins with the sequence homologous to rHP2 (dHP2) or rHP1 (dHP1) (Figures 2 and 3, Table 1, Supplementary Figure S13), pre-

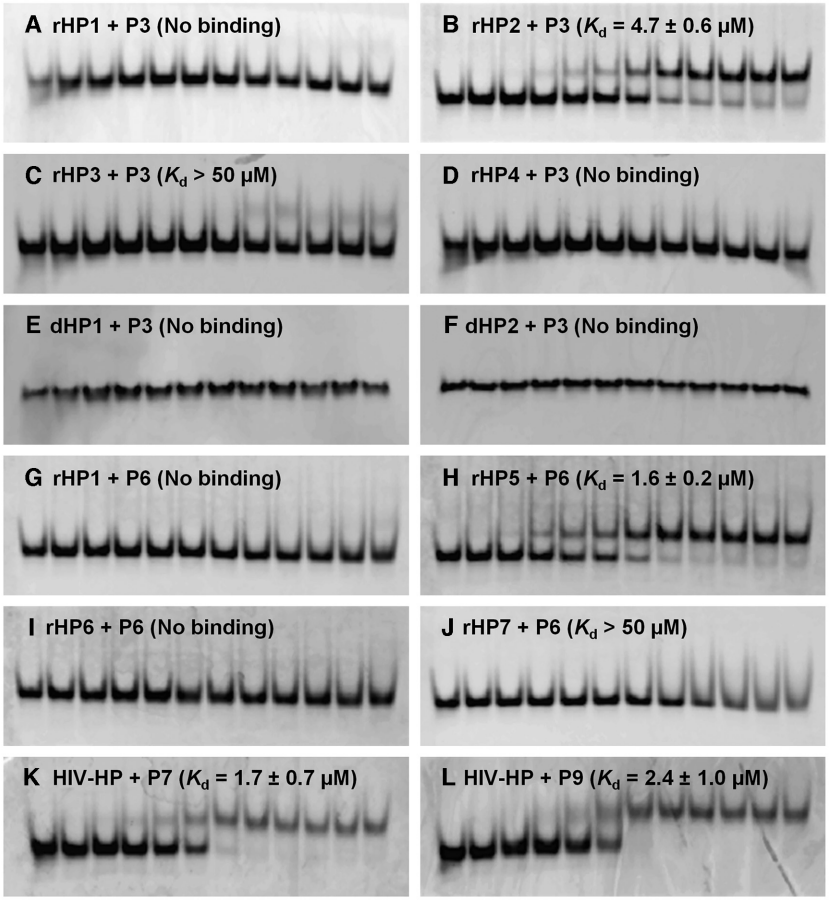


Figure 3. Non-denaturing PAGE (12%) with running buffer of 1× TBE, pH 8.3 for 5 h at 250 V. The incubation buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5. The loaded RNA hairpins (rHP1, rHP2, rHP3, rHP4, rHP5, rHP6, rHP7 and HIV-HP) and DNA hairpins (dHP1 and dHP2) are at 1 μM in 20 μl and 10 μl, respectively. The PNA concentrations in lanes from left to right are 0, 0.2, 0.4, 1, 1.6, 2, 4, 10, 16, 20, 28 and 50 μM, respectively. (**A–D**) PNA P3 shows binding to rHP2 with a K_d value of (4.7 ± 0.6) μM but no or weak binding to rHP1, rHP3, and rHP4. (**E** and **F**) PNA P3 shows no binding to dHP1 or dHP2. (**G–J**) PNA P6 shows binding to rHP5 with a K_d value of (1.6 ± 0.2) μM but no or weak binding to rHP1, rHP6, and rHP7. (**K** and **L**) P7 (NH_2-Lys-LLTTLLQ-$CONH_2$) and P9 (NH_2-Lys-LLLTTLLQ-$CONH_2$) bind to HIV-HP with K_d values of 1.7 ± 0.7 and 2.4 ± 1.0 μM, respectively.

Table 1. K_d (μM) values for PNA binding to complementary and non-complementary RNA and DNA hairpins obtained by non-denaturing PAGE[a]

	rHP1 (G-C)				rHP2 (C-G)				rHP3 (A-U)	rHP4 (U-A)	dHP1 (G-C)	dHP2 (C-G)
	pH7.5	pH6.0	pH8.0	Mg^{2+}	pH7.5	pH6.0	pH8.0	Mg^{2+}	pH7.5	pH7.5	pH7.5	pH7.5
P1	5.1 ± 1.1 5.3 ± 1.3[b]	0.9 ± 0.2	>20[b]	5.5 ± 0.9	NB[b]	–	–	–	NB[b]	NB[b]	NB[b]	NB
P2	NB	NB	–	NB	NB	NB	–	NB	NB	NB	NB	NB
P3 (**Q**)	NB	NB	–	NB	**4.7 ± 0.6**	1.2 ± 0.2	8.8 ± 1.4	7.1 ± 2.2	>50	NB	NB	NB
P4	3.2 ± 0.7	–	–	–	>50	>50	–	>50	NB	NB	NB	NB
P5 (**L**)	**0.2 ± 0.1** 1.7 ± 0.6[b]	0.2 ± 0.1	0.3 ± 0.1 3.8 ± 1.9[b]	0.2 ± 0.1	NB NB[b]	–	–	–	0.4 ± 0.1	1.9 ± 0.4 12.2 ± 3.8[b]	>50 >50[b]	NB >50[b]

[a]A total of four incubation buffers used: pH7.5 (200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5), pH 6.0 (200 mM NaCl, 0.5 mM EDTA, 20 mM MES, pH 6.0), pH8.0 (200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 8.0), and Mg^{2+} (200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, 2 mM $MgCl_2$, pH 7.5). '–' represents that the data were not measured. 'NB' indicates that no binding was observed.

[b]Data from a previous study (15). In this study, for the gels with P5, the RNA/DNA hairpin concentrations are 0.25 μM (instead of 1 μM as done previously (15)) and the PNA concentrations are 0, 0.02, 0.05, 0.1, 0.15, 0.2, 0.4, 1, 2, 4, 10, 20 and 50 μM. The conditions for the other gels can be found in Figure 3 caption. We observed tighter binding for P5 in this study compared to the previous study (15).

sumably because the major groove of a DNA duplex is not structurally compatible for short PNA binding. The selective binding of PNA P3 to dsRNA over dsDNA observed here is consistent with previous non-denaturing PAGE and isothermal titration calorimetry (ITC) studies (15,18,19). Thus, the Q residue in a PNA is highly sequence specific in forming an internal Q·C-G base triple in a PNA·RNA_2 triplex.

Our non-denaturing PAGE data reveal that substitution of the single Q residue in P3 with a C residue (P4, NH_2-Lys-TLTCTTTL-$CONH_2$, see Figure 2K) results in weak binding (K_d > 50 μM) to rHP2 (Table 1, Supplementary Figure S11), suggesting the guanidine group in the Q residue is critical for the recognition of the C-G pair in rHP2. We note that the guanidine group of arginine is often utilized in proteins for the sequence specific recognition of the Hoogsteen

edge of a G base with two hydrogen bond acceptors (carbonyl and N7, Figure 1D) (24,32), and have been utilized in modified nucleic acids to enhance the recognition of the Hoogsteen edge of a G base (25,26).

We carried fluorescence titration studies to confirm the binding affinities for the PNA·RNA_2 triplex formation. We incorporated a fluorescent 2-aminopurine residue into dsRNAs (Figures 4 and 5) for the fluorescence measurements. The fluorescence intensity of 2-aminopurine is known to be sensitive to its local structural environment (33–35). For example, base stacking interactions cause 2-aminopurine fluorescence to be quenched. We observed that the fluorescence of 2-aminopurine of the dsRNA (dsRNA2-2AP, Figure 4B) is quenched upon the addition of the complementary PNA P3 (Figure 5C and D). The quenching is presumably due to the fact, upon the triplex formation, the 2-aminopurine base experiences enhanced base stacking interactions with the adjacent base triple.

The fluorescence signal facilitates the extraction of binding affinities in solution. Consistent with our PAGE studies, the fluorescence titration studies reveal that PNA P3 (NH_2-Lys-TLTQTTTL-$CONH_2$) binds to the complementary dsRNA2-2AP (Figure 4B, H and I) with a binding affinity ($K_d = 0.8 \pm 0.1$ μM) in 200 mM NaCl, pH 7.5 buffer (Figure 5C and D, Table 2). The binding affinity obtained by the fluorescence method is tighter than that obtained by the PAGE method ($K_d = 4.7 \pm 0.6$ μM), presumably because of the running buffer of the PAGE experiment with a relatively high pH (pH 8.3) and/or the difference in the constructs used for PAGE and fluorescence studies. Consistently, the fluorescence titration experiments reveal that PNA P3 binds to non-complementary dsRNAs (dsRNA1-2AP ($K_d = 35.1 \pm 7.3$ μM), dsRNA3-2AP ($K_d = 12.4 \pm 2.0$ μM), and dsRNA4-2AP ($K_d = 15.7 \pm 5.0$ μM)) relatively weakly (Figure 4A, C and D, Table 2, Supplementary Figures S16 and S17). The fluorescence titration data suggest no binding between PNA P3 and dsDNAs (dsDNA1-2AP and dsDNA2-2AP, Figure 4E and F, Table 2, Supplementary Figure S18), consistent with our PAGE results (Table 1).

The fluorescence titration study shows that PNA P2 (NH_2-Lys-TCTQTTTC-$CONH_2$, Figure 2I) has no binding to the complementary dsRNA2-2AP (Figure 4B, Table 2 and Supplementary Figure S19) even in the presence of complementary T·A-U, C^+·G-C, and Q·C-G base triples. Our results suggest the importance of the L·G-C base triples in stabilizing the PNA·RNA_2 triplex (15), which is in agreement with the non-denaturing PAGE results (Table 1 and Supplementary Figure S11).

Effects of pH and salt on PNA·RNA_2 triplex formation

As revealed by the non-denaturing PAGE results, varying pH and magnesium ion concentration causes no significant changes for the binding of P5 (NH_2-Lys-TLTLTTTL-$CONH_2$, see Figure 2L) to rHP1 (Table 1 and Supplementary Figures S14 and S15). The results are consistent with the fact that L base (with an increased pK_a compared to C base) can form stable L·G-C base triple with minimal pH dependence (15) and that P5 has only two positive charges due to the presence of an N-terminal lysine residue. Consequently, upon lowering the pH from 7.5 to 6.0, the binding affinity of P3 to rHP2 was moderately enhanced (with the K_d value decreased from 4.7 ± 0.6 to 1.2 ± 0.2 μM, Table 1 and Supplementary Figures S11 and S12). The addition of 2 mM magnesium ion slightly weakens the binding affinity of P3 to rHP2 ($K_d = 7.1 \pm 2.2$ μM, Table 1 and Supplementary Figures S11 and S12).

In agreement with our PAGE studies, our fluorescence titration experiments show that with increased pH from 7.5 to 8.0, the binding affinity of P3 to the complementary dsRNA2-2AP (Figure 4I) was slightly weakened (with the K_d changing from 0.8 ± 0.1 to 2.7 ± 0.4 μM, Figure 5A, C–F, Table 2). Thus, the attached guanidine group in P3 stabilizes the PNA·RNA_2 triplex formation mainly through sequence specific Q·C-G base triple formation, but not non-specific ionic interaction as the case when the guanidine group is attached to the PNA backbone (16,36–38).

Effect of position of Q modification on PNA·RNA_2 triplex formation

In PNA P3 (NH_2-Lys-TLTQTTTL-$CONH_2$), the Q residue is flanked by two T residues. To determine the effects of position and sequence context on the recognition of an RNA C-G pair by PNA Q base, we synthesized PNA P6 (NH_2-Lys-TLTLQTTL-$CONH_2$, see Figure 2P), with Q residue flanked by L and T residues. We studied the binding of P6 to complementary and non-complementary RNA hairpins by non-denaturing PAGE (Figures 2A, M–O and 3G–J, Table 3). P6 binds to the complementary RNA hairpin rHP5 ($K_d = 1.6 \pm 0.2$ μM, Figure 3H, Supplementary Figure S15i) to form a triplex containing a Q·C-G base triple (Figure 2Q). Our PAGE results reveal no binding of P6 to the non-complementary RNA hairpins with the C-G pair in rHP5 (Figure 2M) replaced by an A-U (rHP1), G-C (rHP6), and U-A (rHP7), respectively (Figures 2A, n and o and 3G, I and J). Thus, Q-modified PNAs are useful in recognizing dsRNAs containing an internal C-G pair with different sequence context and positions.

We also tested the recognition of the HIV-1 frameshift inducing RNA hairpin structure by PNAs with the Q residue incorporated at the terminal end (Figure 2S and T). As revealed by the non-denaturing PAGE data, at 200 mM NaCl, pH 7.5, PNA P7 (NH_2-Lys-LLTTLLQ-$CONH_2$) binds to the HIV-1 frameshift inducing RNA hairpin ($K_d = 1.7 \pm 0.7$ μM, Figure 3K, Supplementary Figure S12e). The binding affinity of PNA P7 is comparable to the PNA without the Q residue (NH_2-Lys-LLTTLL-$CONH_2$, $K_d = 1.1 \pm 0.3$ μM) (15), and PNA P9 (NH_2-Lys-LLLTTLLQ-$CONH_2$, $K_d = 2.4 \pm 1.0$ μM) (Figure 3L, Supplementary Figure S12f). Thus, the stabilization effect of a Q residue at the terminal end is not as significant when it is in the middle of a PNA, which is similar to the L residue (15).

In this study, we found that an RNA hairpin concentration of 1 μM is relatively too high for measuring the relatively tight binding between P5 (NH_2-Lys-TLTLTTTL-$CONH_2$, Figure 2L) and some of the RNA hairpins. Thus, we used a relatively lower concentration of rHP1 (0.25 μM instead of 1 μM (15)) to measure by PAGE the relatively tight binding between P5 and rHP1 and other hairpins more accurately. The new gel data suggest that P5 can bind to

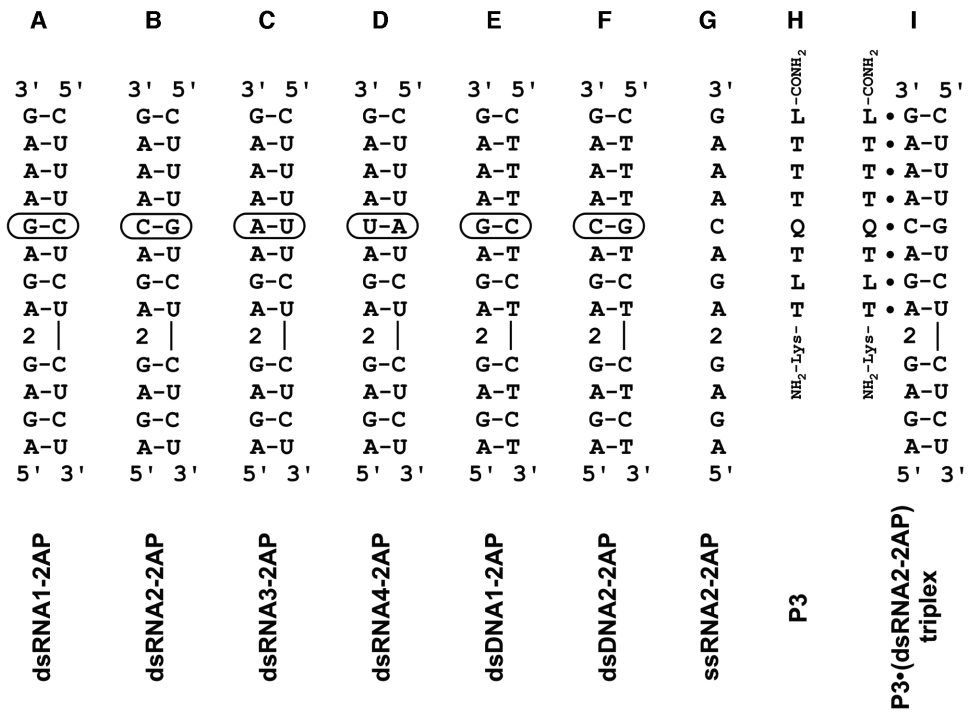


Figure 4. 2-Amimopurine-labeled RNAs, DNAs, and PNA P3 used for fluorescence titration study. (**A–F**) The 2-aminopurine residue (designated as '2') is inserted into the duplexes as a bulge. The RNA duplex construct dsRNA2-2AP (panel **B**) contains the complementary base pairs for forming eight base triples with P3 (**H and I**).

Table 2. K_d (μM) values for PNA binding to 2AP-labeled dsRNAs, dsDNAs, and an ssRNA obtained by fluorescence experiments[a]

	dsRNA1-2AP (G-C)	dsRNA2-2AP (C-G)	dsRNA3-2AP (A-U)	dsRNA4-2AP (U-A)	dsDNA1-2AP (G-C)	dsDNA2-2AP (C-G)	ssRNA2-2AP
P2	–	NB	–	–	–	–	NB
P3 (**Q**)	35.1 ± 7.3	**0.8 ± 0.1** 2.7 ± 0.4[b]	12.4 ± 2.0	15.7 ± 5.0	NB	NB	NB

[a]Incubation buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5 was used, unless otherwise noted.
[b]Incubation buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 8.0 was used. '–' represents that the data were not measured. 'NB' indicates that no binding was observed.

Table 3. K_d (μM) values for P5 and P6 binding to complementary and non-complementary RNA hairpins obtained by non-denaturing PAGE[a]

	rHP1 (A-U)	rHP5 (C-G)	rHP6 (G-C)	rHP7 (U-A)
P5 (**T**)	**0.2 ± 0.1**	NB	1.3 ± 0.1	0.4 ± 0.2
P6 (**Q**)	NB	**1.6 ± 0.2**	NB	>50

[a]Incubation buffer of 200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5 was used. 'NB' indicates that no binding was observed. For the gels with P5, the RNA hairpin concentrations are 0.25 μM and the PNA concentrations are 0, 0.02, 0.05, 0.1, 0.15, 0.2, 0.4, 1, 2, 4, 10, 20 and 50 μM. The conditions for the gels with P6 can be found in Figure 3 caption.

rHP1, rHP3, and rHP4 more tightly compared to our previous gel measurement (Table 1, Supplementary Figures S14 and S15) (15). PNA P1 (NH_2-Lys-TCTCTTTC-$CONH_2$, Figure 2H) shows a similar binding affinity (with the K_d around 5 μM, Table 1, Supplementary Figures S9 and S10) to rHP1 in both the current and our previous gel measurements (15). P5 shows no or weak binding to rHP2 and dHP1 (Table 1, Supplementary Figure S14), consistent with our previous study (15).

We studied the binding of P5 to RNA hairpins rHP5, rHP6, and rHP7 (Figure 2L–O, Table 3) to test the selectivity of the PNA T residue in the recognition of RNA A-U pair over C-G, G-C, and U-A pairs. The PAGE data suggest that the T residue in P5 (Figure 2L) can selectively recognize an A-U pair over U-A, G-C, and C-G pairs although the selectivity is moderate (Table 3).

Q-modified PNAs show no binding to ssRNA

Previous computational modeling studies have indicated that the two-carbon linker and the guanidine group are orientated to form a total of three hydrogen bonds with the major-groove edge of a C-G pair (see Figure 1D) (27). Consistently, previous experimental studies show that, alkyla-

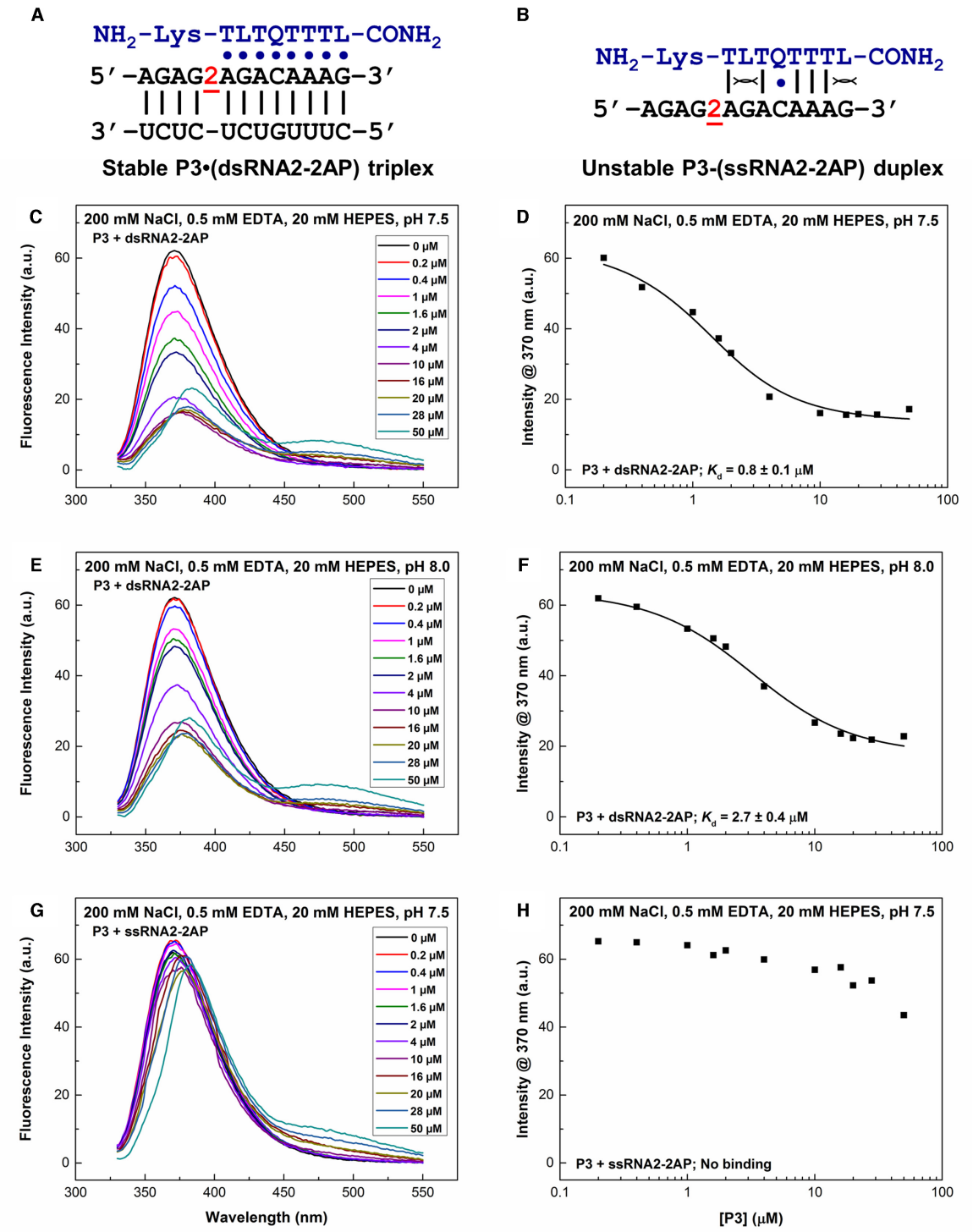


Figure 5. Fluorescence titration study of PNA P3 binding to 2-aminopurine-labeled RNAs. The 2-aminopurine residue is designated as '2' in the RNA sequence. (**A**) A PNA·RNA_2 triplex formed between P3 and a 2-aminopurine-labeled dsRNA. The RNA duplex construct dsRNA2-2AP contains the complementary base pairs for forming eight base triples with P3. (**B**) A hypothetical PNA–RNA duplex formed between P3 and a 2-aminopurine-labeled single-stranded RNA (ssRNA2-2AP). (**C** and **E**) Fluorescence emission spectra for the 2-aminopurine-labeled RNA duplex (1 μM) with varied P3 concentration at pH 7.5 (**C**) and pH 8.0 (**E**), respectively. The peak at around 475 nm is due to the weak fluorescence emission of L base in the PNA. PNA P2 (with no L base incorporated) shows no fluorescence emission at around 475 nm (see Supplementary Figure S19). (**D** and **F**) K_d determination based on the plots of 2-aminopurine fluorescence intensity at 370 nm versus PNA P3 concentration. Consistent with the PAGE studies (Table 1), the PNA·RNA_2 triplex is slightly destabilized upon increasing pH from 7.5 to 8.0. (**G**) Fluorescence emission spectrum for the 2-aminopurine-labeled single-stranded RNA (1 μM) with varied P3 concentration at pH 7.5. (**H**) The fluorescence intensity at 370 nm for the 2-aminopurine-labeled single-stranded RNA remains relatively constant with varied P3 concentration. The presence of destabilizing Q·C mismatch and Watson–Crick-like G-L pairs results in no binding of PNA P3 to the 2-aminopurine-labeled single-stranded RNA.

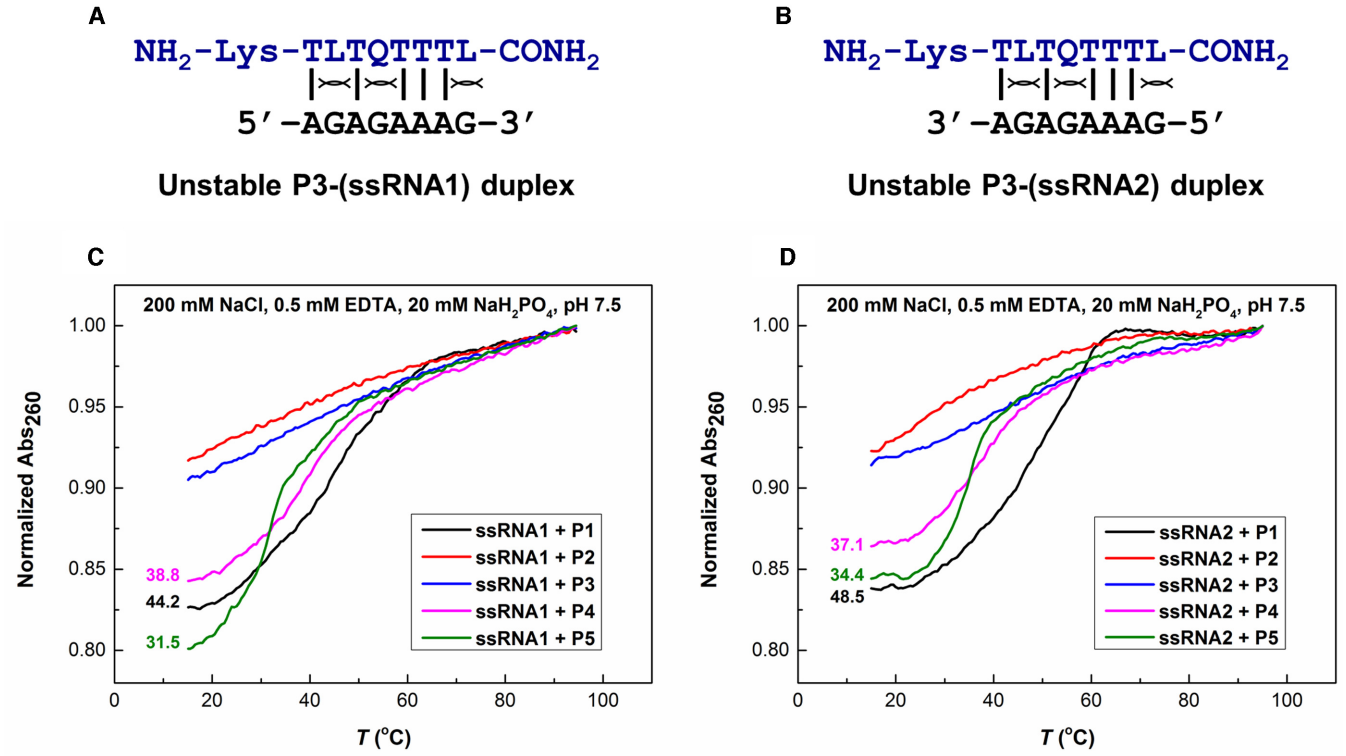


Figure 6. Thermal melting results for RNA-PNA duplexes. The buffer is 200 mM NaCl, 0.5 mM EDTA, 20 mM NaH_2PO_4, pH 7.5. For the curves with transitions, hysteresis between heating and cooling was observed (see Supplementary Figure S20). Only the heating curves are shown. (**A** and **B**) Hypothetical PNA-RNA duplexes formed between PNA P3 and ssRNA1 with PNA P3 binding to ssRNA1 (5′-AGAGAAAG-3′) in a parallel orientation (**A**), and between PNA P3 and ssRNA2 with PNA P3 binding to ssRNA2 (5′-GAAAGAGA-3′) in an anti-parallel orientation (**B**). The steric clash is indicated for the Watson–Crick-like Q-G and L-G pairs. (**C** and **D**) Melting curves for the samples containing ssRNA1 (**C**) and samples containing ssRNA2 (**D**). The melting temperatures are shown for the curves with melting transitions. Due to the steric clash present in a Watson–Crick-like Q-G pair, the PNAs containing Q residues (P2 and P3) show no melting transitions suggesting no binding to the single-stranded RNAs. Compared to the unmodified PNA P1, PNAs P4 and P5 contain modified L residues but no Q residues, and show decreased melting temperature for the corresponding RNA-PNA duplexes due to the steric clash present in a Watson–Crick-like L-G pair.

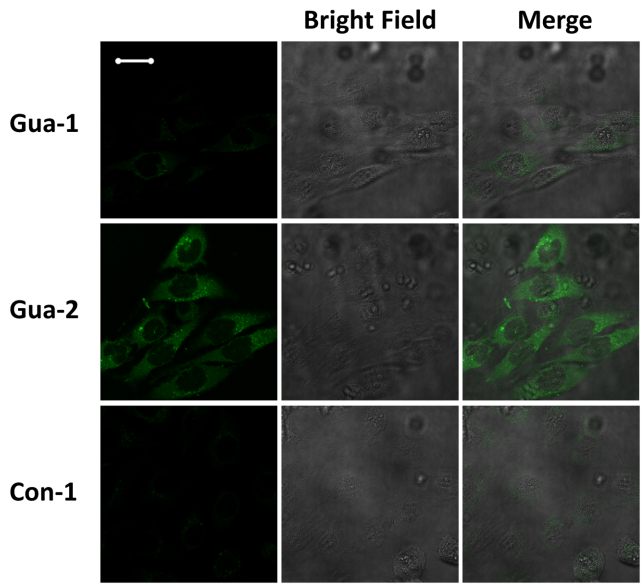


Figure 7. Confocal microscope images of HeLa cells after being treated with 0.5 μM of Cy3-labeled Gua-1 (Cy3-Lys-TCTQTTTC-$CONH_2$), Gua-2 (Cy3-Lys-TQTQTTTQ-$CONH_2$) and Con-1 (Cy3-Lys-TCTCTTTC-$CONH_2$) for 24 hours. Images were captured under the same laser intensity (Ex.: 543 nm, Em: 570–630 nm). Scale bar: 20 μm.

tion of the amine of a C base causes destabilization of a Watson–Crick C-G pair, due to the blocking of its own Watson–Crick edge by the alkyl group (39,40). Presumably, the 5-methyl group in Q base may repel the carbon linker away and favor the orientation with the linker towards its own Watson–Crick edge (27), and thus destabilizes a Watson-Crick-like Q-G pair due to the steric clash caused by the linker (Figure 1F), but favors Q·C-G base triple formation (Figure 1D).

Consistently, the UV-absorbance-detected thermal melting results reveal no appreciable binding between the PNAs containing a Q residue (PNAs P2 and P3) and single-stranded RNAs (Figure 6 and Supplementary Figure S20). For the curves with transitions, hysteresis between heating and cooling was observed (see Supplementary Figure S20), presumably due to the slow binding and/or dissociation rates compared to the temperature ramp rate (0.5°C/min). Consistent with our previous studies (15), the PNAs incorporating L residues (P4 and P5) show significantly weakened binding to single-stranded RNAs (Figure 6 and Supplementary Figure S20).

Consistently, the 2-aminopurine-labeled single-stranded RNA shows no fluorescence intensity change upon the addition of PNAs P2 and P3 (Figure 5B, G, H and Supplementary Figure S19), suggesting no appreciable binding

between the single-stranded RNA and PNAs P2 and P3. Thus, the presence of a destabilizing Q·C mismatch and/or Watson–Crick-like G-L pairs avoids the formation of a stable PNA-RNA duplex.

Taken together, our results show that L and Q modifications in PNAs facilitate enhanced recognition of internal G-C and C-G pairs, respectively, in RNA duplexes with minimal binding to single-stranded RNAs or DNA duplexes, at near-physiological conditions. We note that the RNA duplex-binding triplex-forming PNAs (9,10) are complementary to the pyrrol-imidazole polyamides, which selectively and sequence specifically bind to the minor groove of DNA duplexes (41).

Cellular uptake of Q-modified PNAs

Without the aid of transfection agents, oligonucleotides and PNAs usually have low cell permeability (42–47). The cellular uptake of a DNA duplex-binding TFO with a 2′-OMe or 2′-AE RNA backbone containing a Q base is facilitated by electroporation (27). We reason that combining the guanidine group and the flexibility of PNA backbone may facilitate the penetration (37,38,46,48–53) of the Q modified PNAs through the cell membrane. Our preliminary confocal microscopy studies (Figure 7) show that a PNA incorporating three Q residues and labeled with Cy3 dye (Supplementary Table S1) is taken by HeLa cells without any transfection agent.

CONCLUSION

In summary, we have developed a convenient synthesis method for the PNA monomer Q. Utilizing native polyacrylamide gel electrophoresis and fluorescence binding studies, we have demonstrated that short PNAs containing Q residue and previously reported thio-pseudoisocytosine (L) residue (15) show enhanced recognition of internal C-G and G-C pairs, respectively, in an RNA duplex at near-physiological conditions. Importantly, the relatively short PNAs show no appreciable binding to DNA duplexes or single-stranded RNAs. PNAs incorporating multiple Q residues pass into the HeLa cells without any transfection agent. Our work provides a foundation for the development of next-generation RNA duplex-binding PNAs with high sequence-specificity and bioavailability for biotechnological applications.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We thank Dr Walter N. Moss for critically reading the manuscript.

FUNDING

Singapore Ministry of Education (MOE) Tier 1 [RGT3/13 to G.C.]; MOE Tier 2 [MOE2013-T2-2-024 to G.C.]; MOE Tier 2 [MOE2015-T2-1-028 to G.C.]; Singapore National Research Foundation Fellowship [NRF2009NRF-RF001-015 to Y.Z.]. A patent application (PAT/179/14/15/PCT) based on the work reported here has been filed. Funding for open access charge: MOE Tier 2 [MOE2015-T2-1-028 to G.C.].

Conflict of interest statement. None declared.

REFERENCES

1. 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.
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. Testa,S.M., Disney,M.D., Turner,D.H. and Kierzek,R. (1999) Thermodynamics of RNA-RNA duplexes with 2- or 4-thiouridines: implications for antisense design and targeting a group I intron. *Biochemistry*, **38**, 16655–16662.
4. Kierzek,E., Kierzek,R., Moss,W.N., Christensen,S.M., Eickbush,T.H. and Turner,D.H. (2008) Isoenergetic penta- and hexanucleotide microarray probing and chemical mapping provide a secondary structure model for an RNA element orchestrating R2 retrotransposon protein function. *Nucleic Acids Res.*, **36**, 1770–1782.
5. Gude,L., Berkovitch,S.S., Santos,W.L., Kutchukian,P.S., Pawloski,A.R., Kuimelis,R., McGall,G. and Verdine,G.L. (2012) Mapping targetable sites on human telomerase RNA pseudoknot/template domain using 2′-OMe RNA-interacting polynucleotide (RIPtide) microarrays. *J. Biol. Chem.*, **287**, 18843–18853.
6. Kole,R., Krainer,A.R. and Altman,S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.*, **11**, 125–140.
7. Armitage,B.A. (2003) The impact of nucleic acid secondary structure on PNA hybridization. *Drug Discov. Today*, **8**, 222–228.
8. Conrad,N.K. (2014) The emerging role of triple helices in RNA biology. *Wiley Interdiscip. Rev. RNA*, **5**, 15–29.
9. 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.
10. 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.
11. 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. U.S.A.*, **90**, 3806–3810.
12. 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.
13. 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.
14. 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.
15. 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.
16. Gupta,P., Muse,O. and Rozners,E. (2012) Recognition of double-stranded RNA by guanidine-modified peptide nucleic acids. *Biochemistry*, **51**, 63–73.
17. 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.
18. 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.
19. 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.

20. Zengeya,T., Li,M. and Rozners,E. (2011) PNA containing isocytidine nucleobase: synthesis and recognition of double helical RNA. *Bioorg. Med. Chem. Lett.*, **21**, 2121–2124.
21. Masliah,G., Barraud,P. and Allain,F.H. (2013) RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell. Mol. Life Sci.*, **70**, 1875–1895.
22. Vukovic,L., Koh,H.R., Myong,S. and Schulten,K. (2014) Substrate recognition and specificity of double-stranded RNA binding proteins. *Biochemistry*, **53**, 3457–3466.
23. Brakier-Gingras,L., Charbonneau,J. and Butcher,S.E. (2012) Targeting frameshifting in the human immunodeficiency virus. *Expert Opin. Ther. Targets*, **16**, 249–258.
24. Seeman,N.C., Rosenberg,J.M. and Rich,A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 804–808.
25. Fox,K.R. and Brown,T. (2011) Formation of stable DNA triplexes. *Biochem. Soc. Trans.*, **39**, 629–634.
26. 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.
27. 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.
28. 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.
29. Bernatowicz,M.S., Wu,Y.L. and Matsueda,G.R. (1993) Urethane protected derivatives of 1-guanylpyrazole for the mild and efficient preparation of guanidines. *Tetrahedron Lett.*, **34**, 3389–3392.
30. Robles,J., Grandas,A. and Pedroso,E. (2001) Synthesis of modified oligonucleotides containing 4-guanidino-2-pyrimidinone nucleobases. *Tetrahedron*, **57**, 179–194.
31. Rostovtsev,V.V., Green,L.G., Fokin,V.V. and Sharpless,K.B. (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chem. Int. Ed.*, **41**, 2596–2599.
32. 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.
33. Hochstrasser,R.A., Carver,T.E., Sowers,L.C. and Millar,D.P. (1994) Melting of a DNA helix terminus within the active-site of a DNA-polymerase. *Biochemistry*, **33**, 11971–11979.
34. Sakakibara,Y., Abeysirigunawardena,S.C., Duc,A.C.E., Dremann,D.N. and Chow,C.S. (2012) Ligand- and pH-induced conformational changes of RNA domain helix 69 revealed by 2-aminopurine fluorescence. *Angew. Chem. Int. Ed.*, **51**, 12095–12098.
35. Wilcox,J.L. and Bevilacqua,P.C. (2013) A simple fluorescence method for pK_a determination in RNA and DNA reveals highly shifted pK_a's. *J. Am. Chem. Soc.*, **135**, 7390–7393.
36. 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.
37. 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.
38. Jain,D.R., Anandi,V.L., Lahiri,M. and Ganesh,K.N. (2014) Influence of pendant chiral C(gamma)-(alkylideneamino/guanidino) cationic side-chains of PNA backbone on hybridization with complementary DNA/RNA and cell permeability. *J. Org. Chem.*, **79**, 9567–9577.
39. Engel,J.D. and von Hippel,P.H. (1974) Effects of methylation on the stability of nucleic acid conformations: studies at the monomer level. *Biochemistry*, **13**, 4143–4158.
40. Micura,R., Pils,W., Hobartner,C., Grubmayr,K., Ebert,M.O. and Jaun,B. (2001) Methylation of the nucleobases in RNA oligonucleotides mediates duplex-hairpin conversion. *Nucleic Acids Res.*, **29**, 3997–4005.
41. Chenoweth,D.M., Meier,J.L. and Dervan,P.B. (2013) Pyrrole-imidazole polyamides distinguish between double-helical DNA and RNA. *Angew. Chem. Int. Ed.*, **52**, 415–418.
42. Stein,C.A., Hansen,J.B., Lai,J., Wu,S., Voskresenskiy,A., Hog,A., Worm,J., Hedtjarn,M., Souleimanian,N., Miller,P. *et al.* (2010) Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic Acids Res.*, **38**, e3.
43. Wojtkowiak-Szlachcic,A., Taylor,K., Stepniak-Konieczna,E., Sznajder,L.J., Mykowska,A., Sroka,J., Thornton,C.A. and Sobczak,K. (2015) Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res.*, **43**, 3318–3331.
44. Fabani,M.M., Abreu-Goodger,C., Williams,D., Lyons,P.A., Torres,A.G., Smith,K.G., Enright,A.J., Gait,M.J. and Vigorito,E. (2010) Efficient inhibition of miR-155 function in vivo by peptide nucleic acids. *Nucleic Acids Res.*, **38**, 4466–4475.
45. 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.
46. 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.
47. 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.
48. 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.
49. 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.
50. 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.
51. Herce,H.D., Garcia,A.E. and Cardoso,M.C. (2014) Fundamental molecular mechanism for the cellular uptake of guanidinium-rich molecules. *J. Am. Chem. Soc.*, **136**, 17459–17467.
52. Wexselblatt,E., Esko,J.D. and Tor,Y. (2014) On guanidinium and cellular uptake. *J. Org. Chem.*, **79**, 6766–6774.
53. 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.