Recognition of Non-Canonical RNA Base Pairs Using Triplex-Forming Peptide Nucleic Acids

Sara Farshineh Saei,[1] Vladislavs Baskevics,[2] Martins Katkevics,[2] and Eriks Rozners*,[1]

[1] Department of Chemistry, Binghamton University, The State University of New York, Binghamton, New York 13902, United States

[2] Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga, LV-1006, Latvia

ABSTRACT: Non-canonical base pairs play an important role in enabling the structural and functional complexity of RNA. Molecular recognition of such motifs is challenging because of their diversity, significant deviation from the Watson-Crick structures, and dynamic behavior resulting in alternative conformations of similar stability. Triplex-forming peptide nucleic acids (PNAs) have emerged as excellent ligands for the recognition of Watson-Crick base-paired double helical RNA. The present study extends the recognition potential of PNA to RNA helices having non-canonical G∘U, A∘C, and tandem G∘A/A∘G base pairs. The purines of the non-canonical base pairs formed M^+•G∘U, T•A∘C, M^+•G∘A and T•A∘G Hoogsteen triples of similar or slightly reduced stability compared to the canonical M^+•G-C and T•A-U triples. Recognition of pyrimidines was more challenging. While the P•C∘A triple was only slightly less stable than the P•C-G, the E nucleobase did not form a stable triple with U of the U∘G wobble pair. Molecular dynamics simulations suggested the formation of expected Hoogsteen hydrogen bonds for all stable triples. Collectively, these results expand the scope of triple helical recognition to non-canonical structures and sequence motifs common in biologically relevant RNAs.

Introduction

RNA folds in complex double helical conformations featuring a variety of non-Watson-Crick structural elements, such as internal, hairpin and multi-helix junction loops, bulges, and base triples and quadruples. Although the various loops are typically drawn as single stranded in the RNA secondary structure diagrams, in 3D structures most of these segments are well-structured and folded in helices where more than half of the loop nucleotides form non-canonical base pairs.[1, 2] Hence, the RNA motifs built of non-canonical base pairs are abundant and play important roles in structural flexibility of RNA, its recognition by proteins, and long-range RNA-RNA interactions.

Molecular recognition of folded RNA is a highly desirable but daunting task. Extensive efforts to find small molecule ligands that specifically recognize RNA loops and bulges have been successful in specific cases.[3-5] However, the ability to recognize any sequence of folded RNA remains elusive.[6-9] Although it is relatively easy to find strong RNA binders, difficulties still remain in finding sequence-specific binders. RNA helices have a uniform and polar surface that presents little opportunity for hydrophobic shape selective recognition. Consequently, the most common targets for small molecules are RNA bulges and internal loops. However, binding to these structures is frustrated by the conformational flexibility of non-helical RNA. As a result, many strong RNA binders must rely on electrostatic interactions with the negatively charged phosphates to boost the affinity at the expense of selectivity, which is less than ideal for drug development.

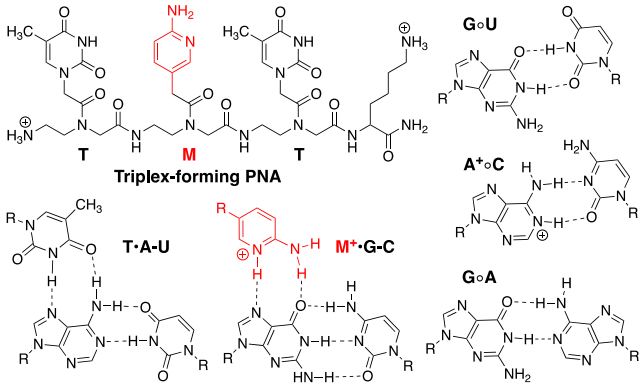


Figure 1. Structures of triplex-forming PNA, Hoogsteen hydrogen-bonded base triples, and representative non-canonical RNA base pairs.

We[10] and other research groups[11-17] have been developing an alternative approach to the recognition of double-stranded RNA (dsRNA) using triplex-forming peptide nucleic acids (PNAs). PNAs are DNA analogues built of neutral and achiral amide backbone (Figure 1). They were initially developed as triplex-forming ligands for DNA in 1991,[18] but have gained a wide variety of applications as duplex-forming probes and diagnostics because of their strong and sequence

specific binding to single-stranded DNA and RNA.[19, 20] Interestingly, PNA-dsRNA triplexes were not explored until the 2010 report by Rozners and co-workers.[21] In the following decade, our research group found that PNA modified with 2-aminopyridine (M in Figure 1) had exceptionally high affinity and sequence specificity for the matched dsRNAs under physiological conditions and formed PNA-dsRNA triplexes that were more than 10-fold more stable than the corresponding PNA-dsDNA triplexes.[22-24] Collaborative work with Sugimoto and Endoh demonstrated that PNA-dsRNA triplex formation was able to inhibit mRNA translation[25] and microRNA maturation[26] in live cells. Most recently, we became interested in using triplex-forming PNAs to control the conformation of non-Watson-Crick motifs of folded RNAs. We discovered that M-modified PNAs formed unusually stable triple helices with dsRNAs having single and double purine nucleotide bulges.[27] Depending on the sequence of PNA, triplex formation shifted the dynamic equilibrium of the bulge from looped-out to stacked-in conformation. Collectively, previous studies by us and others demonstrated that PNAs were excellent ligands for sequence specific recognition of folded double helical RNA and had strong potential to modulate conformation and function of biologically relevant RNAs.[10, 17]

In the present study, we expand the triple helical recognition of dsRNA to include non-canonical G◦U, A◦C, and tandem G◦A/A◦G base pairs. In the context of the present study, we consider only the Watson-Crick U-A and C-G as the canonical base pairs, while all other base pairing possibilities are called non-canonical whether they engage in Watson-Crick-like or Hoogsteen or sugar edge hydrogen bonding.[1, 2] Previously, our group[22, 26, 28, 29] and others[14, 30-33] have reported recognition of biologically relevant RNA targets that presumably involved Hoogsteen hydrogen binding to purines of non-canonical G◦U and A◦C base pairs. The formation of stable triple helices was observed in these studies; however, the effect of non-canonical base pairs on the complex formation was not specifically explored. Recognition of pyrimidines of non-canonical base pairs is less studied. A notable exception is the report by Chen and co-workers that U of U◦G was recognized by an extended heterocyclic nucleobase S but not by nucleobase E designed to bind the canonical U-A base pairs.[34] Herein, we report systematic study of triple helical recognition of double helical RNA featuring the non-canonical G◦U, A◦C, and tandem G◦A/A◦G base pairs. We found that triplex-forming PNAs recognized the purine nucleosides of non-canonical and Watson-Crick base pairs with similar or slightly reduced affinity. Recognition of pyrimidines was notably less efficient, especially that of uridine in the U◦G wobble pair. These results further demonstrate the effectiveness of triplex-forming PNAs in molecular recognition of complex structures of folded RNA.

Results and Discussion

In our previous studies on PNA-dsRNA triplexes, we used a model system of four hairpins (HRP1-HRP4) having a variable base pair in the middle of the stem.[22-24] In the present study, we modified our model system by placing the non-canonical base pairs at the same position (Figures 2, 3, and 6). The original sequences formed highly stable hairpins HRP1-HRP4 having melting temperatures (T_ms) > 90 °C (Figure S12A, Table S13). Modification of sequences to incorporate the non-canonical base pairs resulted in a slight loss of thermal stability; however, all hairpins used in this study had T_ms > 84 °C (Figures S12B and S12C, Tables S14 and S15), which was ~20 °C higher than T_ms of the corresponding triplexes.

PNAs carrying modified nucleobases (M, E, P, etc.) were synthesized as previously described[35] and their binding affinity for the RNA hairpins having the non-canonical base pairs was measured using isothermal titration calorimetry and UV thermal melting, as in our previous studies.[24] The UV melting experiments were done at 300 nm, a wavelength that allows selective detection of triplex melting without interference from the melting of RNA hairpins.[36] This was possible because at 300 nm the RNA nucleobases have negligible absorbance while the M nucleobases have weak but easily detectable absorbance.

The G◦U wobble is the most prevalent and functionally important non-canonical base pair in complex RNAs.[37] The geometry of G◦U differs significantly from the canonical Watson-Crick base pairs. In the G◦U wobble, U is displaced towards the deep (major) groove, while G is rotated towards the shallow (minor) groove.[37] The Hoogsteen face of G is available for hydrogen bonding in a similar fashion as in the G-C base pair. Consistent with these considerations, PNA1 formed a triplex with HRP1ug that was only slightly less stable than the canonical PNA1-HRP1 triplex (Figure 2, c.f. K_a 22 vs. 38 × 10^6). In contrast, the stability of the PNA2-HRP4gu triplex that involved a putative E•U◦G triple was significantly lower compared to PNA2-HRP4 having the E•U-A triple (Figure 2, c.f. K_a 0.7 vs. 11 × 10^6). In fact, the stability of the PNA2-HRP4gu triplex was as low as those of the least stable mismatched triplexes in our earlier study.[24] These results suggested that the conformational changes in G◦U wobble pair (compared to Watson-Crick pairs) allowed the formation of a stable M^+•G◦U triple; however, the shift of U towards the major groove prevented the formation of a stable E•U◦G triple.

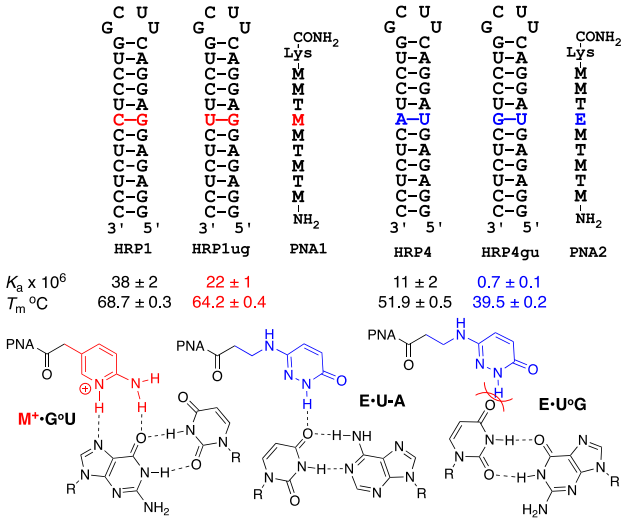


Figure 2. Structures of RNA hairpins and triplex-forming PNAs to study recognition of the U°G wobble base pairs. All binding experiments were done in 50 mM potassium phosphate buffer (pH 7.4) containing 2 mM $MgCl_2$, 90 mM KCl, 10 mM NaCl. Association constants (K_a × 10^6 M^{-1}) obtained by ITC are averages of three experiments ± standard deviation at 25 °C. UV thermal melting temperatures, T_m °C are averages of five experiments ± standard deviation measured at 300 nm and 18 μM of each dsRNA and PNA.

In structures of large folded RNA, the non-canonical A°C base pairs are dynamic and may adopt various conformations, such as, a protonated A^+°C wobble pair, neutral A°C pairs with one hydrogen bond either between exocyclic amine of C and N1of A,[38] or between exocyclic amine of A and the O2 of C,[39] water bridged pairs[40] with no direct hydrogen bonds between A and C, and may even form Watson-Crick-like base pairs involving minor high energy imino tautomers.[41] The protonated A^+°C is structurally similar to the G°U wobble pair.[37] Similar to the recognition of G°U wobble pair, PNA 3 formed a stable triplex with HRP2ca comparable with the canonical PNA3-HRP2 (Figure 3, K_a 12 × 10^6 for both triplexes). Next, we studied the recognition of C in C°A using nucleobases P and CR3,[42] and V[43] that were developed in our earlier studies on the recognition of C in C-G. In contrast to the recognition of U in G°U wobble pair, the stability of PNA4-HRP3ac triplex involving a P•C°A triple was only slightly lower compared to PNA4-HRP3 having the P•C-G triple (Figure 2, c.f. K_a 4.2 vs. 5.4 × 10^6). Similar results were obtained with PNA5 and PNA6 that formed CR_3•C°A and V•C°A triples. These results suggested that the conformation of the A°C base pair in our model triplexes was different than the conformation of the G°U wobble pair.

To further test if A may be protonated, as required for the formation of the A^+°C wobble pair, we compared the stability of PNA3-HRP2ca with PNA3-HRP2 and PNA1-HRP1ug at pH 6.9, 7.4 and 7.9 (Figure S14). As expected, all triplexes were significantly more stable at lower pH because of the more favorable protonation of M to form the M^+•G-C triples. The pK_a of A ~3.5 while the pK_a of M ~6.5. If a protonated T•A^+°C triple was involved in PNA3-HRP2ca, we expected that the stability of the PNA3-HRP2ca triplex would be more enhanced at pH 6.9 and more disfavored at pH 7.9 compared to triplexes involving M^+•G-C triples only. However, all three triplexes were similarly affected by changes in pH (Figure S14). These results suggested that the PNA3-HRP2ca triplex most likely involved a neutral T•A°C triple.

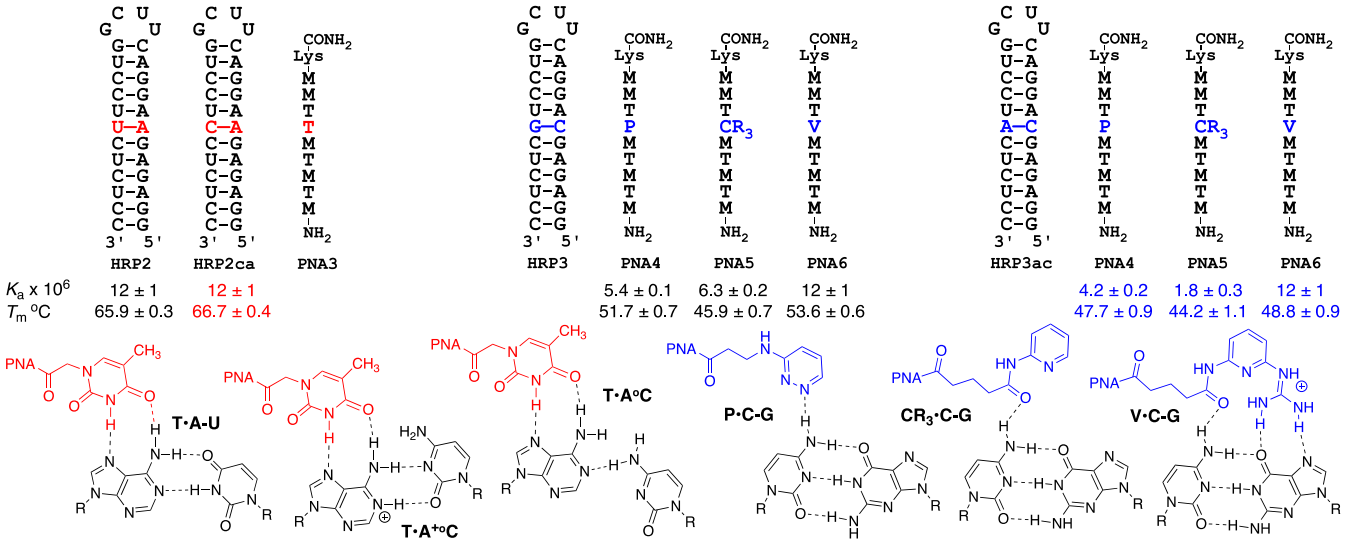


Figure 3. RNA hairpins and PNAs to study recognition of the C°A and A°C base pairs. Experimental conditions as in Figure 2.

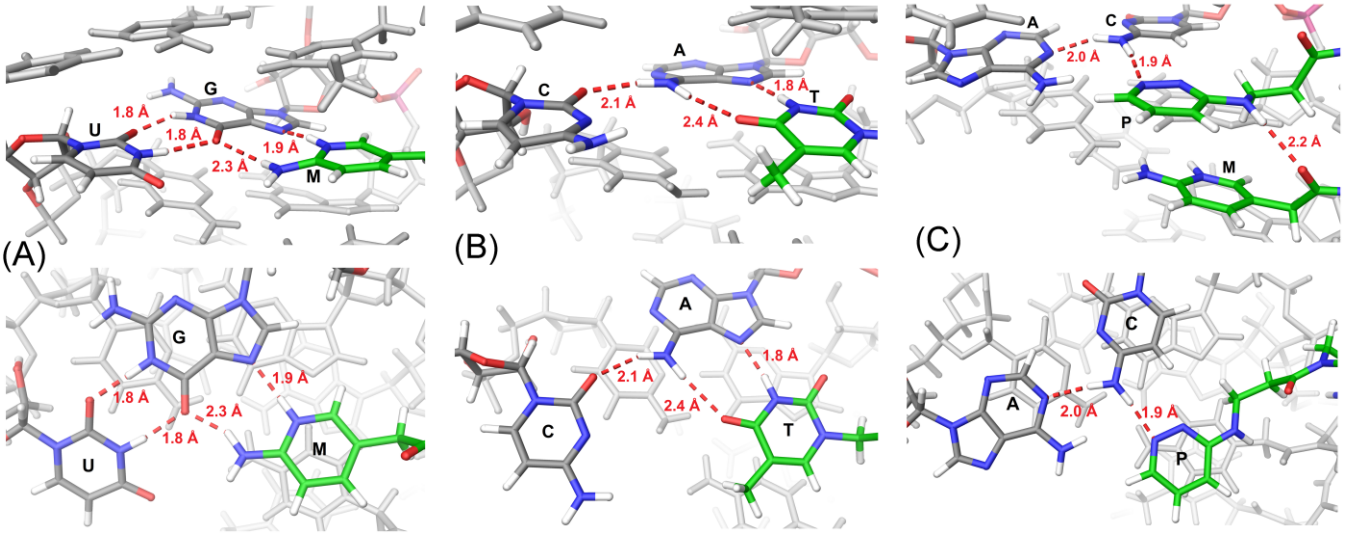


Figure 4. Major groove (upper images) and top (lower images) views of hydrogen-bonding interactions in base triples from molecular dynamics simulations of (A) PNA1-HRP1ug; (B) PNA3-HRP2ca; and (C) PNA4-HRP2ac triplex models. The hydrogen-bonding interactions and average distances observed during molecular dynamics simulations are highlighted in red.

To obtain additional insights into structures of non-canonical base triples, we used molecular modeling and dynamics of model PNA-dsRNA hairpins built in our previous studies.[42, 43] We modeled the conformation of all triples by running 150 ns unrestricted Desmond molecular dynamics (MD) and analyzing the last 50 ns of simulation, when the system had stabilized. To evaluate triple geometries, we performed Desmond trajectory clustering and analyzed clusters with the highest RMSD probability. The simulation of the PNA1-HRP1ug triplex (Figure 4A) showed the expected M^+•G○U triple with M forming two hydrogen bonds to the Hoogsteen face of G. Consistent with the low stability of PNA2-HRP4gu triplex, molecular dynamics simulations did not result in a stable E•U○G triple. The E nucleobase lost hydrogen bonding contacts with U within the first 30 ns of all simulations and moved in the major groove outside the helical stack.

The simulation of the PNA3-HRP2ca triplex (Figure 4B) suggested that in the neutral T•A○C triple, C6-NH_2 of A formed a single hydrogen bond to C2=O of C, as observed in crystal structure of lead-dependent ribozyme.[39] The T•A○C triple was formed by the standard Hoogsteen hydrogen bonding between T and A (Figure 4B). In contrast, the simulation of the PNA4-HRP2ac triplexes (Figure 4C) suggested that in the neutral P•C○A triple, the C○A base pair adopted a different geometry than in the T•A○C triple. In the P•C○A triple, C4-NH_2 of C formed a single hydrogen bond to N1 of A (Figure 4C), as observed in the crystal structures of large RNAs.[38] The P•C○A triple was formed by the expected single hydrogen bond between endocyclic N of P and C4-NH_2 of C. Interestingly, molecular dynamics showed that in the most stable PNA6-HRP2ac triplex, the V•C○A triple formed unusual hydrogen bonding interactions (Figure S38). Because the guanidine group of V base could not engage the Hoogsteen face of A in the same manner as G (see, Figure 3) the V base tilted and hydrogen bonded to U of a neighboring A-U base pair.

The different geometries of the wobble U○G and the neutral C○A base pairs explain the higher stability of the P•C○A triple compared to the E•U○G triple. In the former, C is not pushed out in the major groove, which allows favorable positioning of the P nucleobase. The P nucleobase is stabilized by an additional hydrogen bond between the N-H of the linker connecting P to PNA backbone and the C=O of the linker of adjacent M (2.2 Å in Figure 4C). These results illustrate the dynamic nature of the C○A base pair in our model system. It is conceivable that binding of the PNA ligand shifts the conformation of C○A base pair to optimize the overall stability of PNA-RNA triplexes, as was observed for RNAs having purine bulges in our previous study.[27]

Tandem G○A/A○G base pairs are common motifs of the non-canonical structure of functional RNAs that adopt either a Watson-Crick-like imino or an A Hoogsteen to G sugar edge hydrogen bonded sheared conformations (Figures 5 and 6).[44] Because the imino conformation widens the major groove of RNA leaving both A and G Hoogsteen faces available for triple formation, we expected that the triplex-forming PNAs would be able to bind tandem G○A/A○G base pairs in the imino conformation. In contrast, in the sheared conformation, the Hoogsteen face of A is already engaged in hydrogen bonding with the sugar edge of G. The sheared conformation contracts the major groove and pushes A and G outwards in the minor and major grooves, respectively. While the Hoogsteen face of G is formally available for hydrogen bonding, we expected that the overall distortion of the major groove and lack of hydrogen bonding sites on A would be detrimental to triplex formation at the sheared G○A/A○G conformation.

Expecting that two adjacent non-canonical base pairs may lower the overall stability of triplexes, we extended our test sequence of PNA1 by two nucleobases in PNA7 (Figure 6). Compared to PNA1, PNA7 was binding to the canonical HRP1 notably stronger (Figure 6, c.f. K_a 38 vs. 150 × 10^6). Interestingly, both PNA1 and PNA7 were binding to HRP5 featuring the tandem G○A/A○G sandwiched between purine neighbors (5′-GGAG-3′ recognition sequence) with similar albeit slightly decreased affinity compared to the canonical triplexes. This

result was consistent with tandem GºA/AºG base pairs being in the imino conformation in the 5′-GGAG-3′ sequence of HRP5. We do not have a compelling explanation for the lack of increased stability of the PNA7-HRP5 triplex, formed by two nucleobases longer PNA7, compared to the PNA1-HRP5. Molecular dynamics simulations showed that PNA1-HRP5 and PNA7-HRP5 triplexes were able to form the expected M^+•GºA and T•AºG triples when GºA/AºG base pairs were in the imino conformation (Figure 5). As expected, molecular dynamics simulations did not reveal reasonable PNA-RNA hydrogen bonding arrangements when the GºA/AºG base pairs were arranged in the sheared conformation.

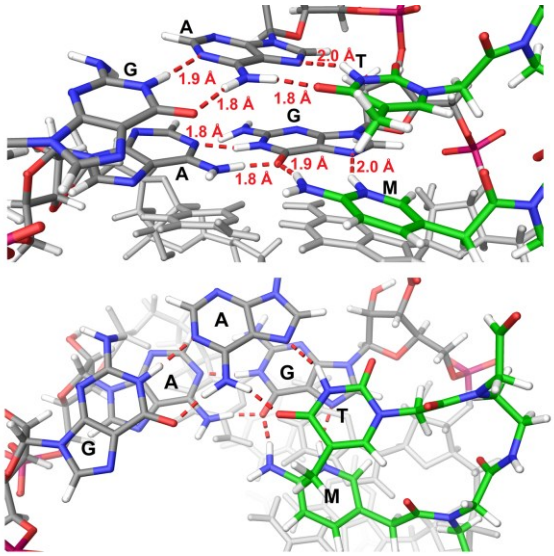


Figure 5. Major groove (upper image) and top (lower image) views of hydrogen-bonding interactions in base triples from molecular dynamics simulations of the PNA7-HRP5 triplex model having the GºA/AºG base pairs in imino conformation. The hydrogen-bonding interactions and average distances observed during molecular dynamics simulations are highlighted in red.

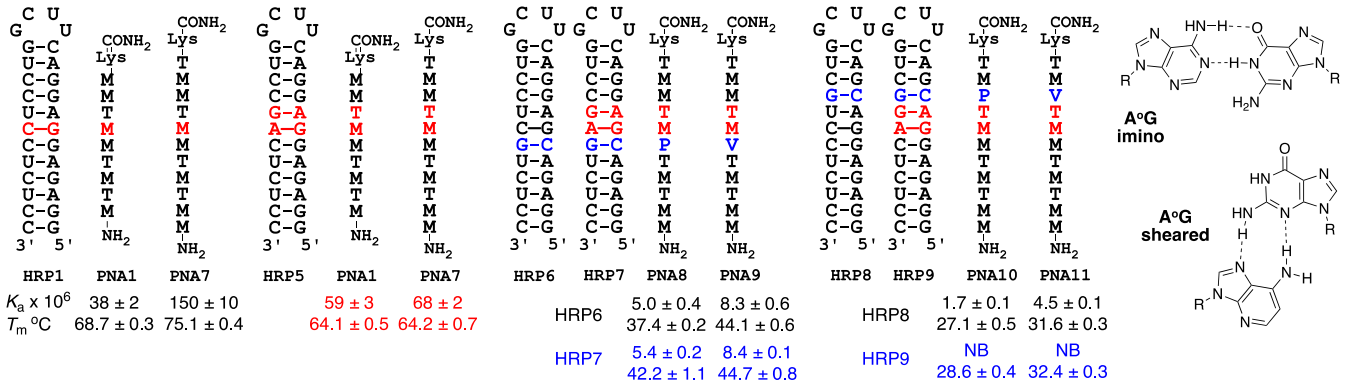


Figure 6. Structures of RNA hairpins and triplex-forming PNAs to study recognition of the tandem GºA/AºG base pairs. Experimental conditions as in Figure 2.

Pioneering work by Turner and co-workers demonstrated that in short RNA duplexes GºA/AºG base pairs preferred sheared conformation in 5′-CGAG-3′ sequences[45] and imino conformation in 5′-GGAC-3′ sequences.[46] To test if this preference would affect triple-helix formation in our model hairpins, we studied the binding of PNAs to HRP6 featuring the 5′-CGAG-3′ sequence and HRP7 featuring the 5′-GGAC-3′ sequence. Each hairpin presented two challenges for triplex-forming PNAs, the tandem GºA/AºG base pairs and an adjacent C-G base pair forming a pyrimidine interruption of the purine strand. Based on previous studies by Turner and co-workers, we expected that the tandem GºA/AºG base pairs would adopt the sheared conformation in HRP7 (Figure 6),[45] which would prevent triplex formation, and imino conformation in HRP9,[46] which would allow triplex formation. Surprisingly, the results showed the opposite trend. PNA8 and PNA9 formed weak but detectable triplexes with HRP7. The stabilities of PNA8-HRP7 and PNA9-HRP7 were similar to the stabilities of control triplexes PNA8-HRP6 and PNA9-HRP6 having all Watson-Crick base pairs. The stabilities of PNA8-HRP7 and PNA9-HRP7 were also significantly lower than that of PNA7-HRP5 (Figure 6) and less but notably lower than that observed for the recognition of a single C-G base pair in PNA4-HRP3 and PNA6-HRP3 (Figure 3). This result was consistent with the less-than-ideal P•C-G and V•C-G triples being the main driving force for lowering the stability of PNA8-HRP7 and PNA9-HRP7 triplexes. Contrary to the expected preference for sheared conformation in HRP7, our results suggested that the tandem GºA/AºG base pairs in HRP7 triplexes might have adopted the imino conformation. However, the relatively low stability of these triplexes does not exclude a possibility of weak interactions between PNAs and sheared GºA/AºG base pairs.

In contrast, using ITC, we were not able to detect the formation of PNA10-HRP9 and PNA11-HRP9 triplexes under our experimental conditions. UV melting of these complexes

showed transitions at very low temperatures confirming the notion that PNA10-HRP9 and PNA11-HRP9 triplexes were severely destabilized. The overall low stability of these triplexes prevented any conclusions about the preferred conformation of tandem G○A/A○G base pairs in HRP9.

Control experiments on triplexes PNA10-HRP8 and PNA11-HRP8 (Figure 6) having all Watson-Crick base pairs and PNA7-HRP7 and PNA7-HRP9 triplexes having a single M•C-G mismatch (~39 and ~32 °C, respectively, Table S10) showed similarly low affinity. Taken together, these results suggested that unstable P•C-G and V•C-G triples were the main driving forces for the low stability. Interestingly, control experiments binding PNA8-PNA11 to HRP5 (Table S12) showed that the mismatched triplexes formed by PNA8 and PNA9 were notably more stable (T_m ~44 and 48°C, respectively) than the mismatched triplexes formed by PNA10 and PNA11 (T_m ~33 and 35°C, respectively). These results suggested that a mismatched or less than optimal base triple was less tolerated at the C-G pyrimidine interruption in the sequence of HRP8 and HRP9, than at a similar C-G interruption in HRP6 and HRP7.

Our attempts at using molecular dynamics simulations to obtain more insights into conformational preferences of triple helices formed with HRP7 featuring the 5´-CGAG-3´ sequence and HRP9 featuring the 5´-GGAC-3´ sequence were not successful. The simulations showed various alternative structures and did not converge on conformations that could provide explanations of the observed differences. Taken together with the low stability of triplexes formed with HRP7 and HPP9, the molecular dynamics suggested that the combination of non-canonical G○A/A○G base pairs adjacent to the P•C-G and V•C-G triples strongly destabilized the complexes. Thus, G○A/A○G base pairs embedded in a polypurine sequence could be recognized with adequate affinity but pyrimidine interruptions directly adjacent to these motifs were not tolerated.

Conclusions

Our study demonstrates the potential of triple-helical recognition of RNA helices having non-canonical G○U, A○C, and tandem G○A/A○G base pairs. The triplex-forming PNAs engaged purines of the non-canonical base pairs and formed M⁺•G○U, T•A⁺○C, M⁺•G○A and T•A○G Hoogsteen triples of similar or slightly reduced stability compared to the canonical M⁺•G-C and T•A-U triples (Figures 2, 3, and 6). Recognition of pyrimidines was more sensitive to conformational preferences of the non-canonical base pairs. The flexible C○A base pair appeared to adopt a conformation that was favorable for the formation of a relatively stable P•C○A triple (Figure 3). On the other hand, the conformation of the rigid U○G wobble pair that places U further out in the major grove was not allowing E nucleobase to form a stable E•U○G triple (Figure 2). For all stable triples, molecular dynamics simulations showed the formation of expected Hoogsteen hydrogen bonds (Figures 4 and 5). Collectively, these results expand the scope of triple helical recognition to non-canonical structures and sequence motifs common in biologically relevant RNAs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Synthesis, HPLC purification, and LC-MS characterization of PNA oligomers; UV melting results; ITC results and representative titration images; experimental details of molecular dynamics simulations (PDF).

AUTHOR INFORMATION

Corresponding Author

* To whom correspondence should be addressed. Tel: +1 607 777 2441; Email: erozners@binghamton.edu.

Author Contributions

The manuscript was written through contributions of all authors. / All authors have approved the final version of the manuscript.

Funding Sources

This work was supported by the National Institutes of Health (R35 GM130207 to ER); the National Science Foundation (CHE-2107900 to ER); and the Recovery and resilience facility (RRF) (5.2.1.1.i.) grant Nr. 16/OSI/DG to VB.

ACKNOWLEDGMENT

We thank I. Kumpina for assistance with PNA synthesis and ITC experiments, C. A. Ryan for the synthesis of PNA5 and PNA6, and B. R. Tessier for help with the ITC data analysis.

REFERENCES

(1) Sweeney, B. A.; Roy, P.; Leontis, N. B. An introduction to recurrent nucleotide interactions in RNA. *WIREs RNA* **2015**, *6*, 17-45.

(2) Leontis, N. B.; Westhof, E. Geometric nomenclature and classification of RNA base pairs. *RNA* **2001**, *7*, 499-512.

(3) Rzuczek, S. G.; Colgan, L. A.; Nakai, Y.; Cameron, M. D.; Furling, D.; Yasuda, R.; Disney, M. D. Precise small-molecule recognition of a toxic CUG RNA repeat expansion. *Nat. Chem. Biol.* **2017**, *13*, 188-193.

(4) Angelbello, A. J.; Rzuczek, S. G.; McKee, K. K.; Chen, J. L.; Olafson, H.; Cameron, M. D.; Moss, W. N.; Wang, E. T.; Disney, M. D. Precise small-molecule cleavage of an r(CUG) repeat expansion in a myotonic dystrophy mouse model. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 7799-7804.

(5) Velagapudi, S. P.; Cameron, M. D.; Haga, C. L.; Rosenberg, L. H.; Lafitte, M.; Duckett, D. R.; Phinney, D. G.; Disney, M. D. Design of a small molecule against an oncogenic noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5898-5903.

(6) Ursu, A.; Childs-Disney, J. L.; Andrews, R. J.; O'Leary, C. A.; Meyer, S. M.; Angelbello, A. J.; Moss, W. N.; Disney, M. D. Design of small molecules targeting RNA structure from sequence. *Chem. Soc. Rev.* **2020**, *49*, 7252-7270.

(7) Childs-Disney, J. L.; Yang, X.; Gibaut, Q. M. R.; Tong, Y.; Batey, R. T.; Disney, M. D. Targeting RNA structures with small molecules. *Nat. Rev. Drug Discovery* **2022**, *21*, 736-762.

(8) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discovery* **2018**, *17*, 547-558.

(9) Zafferani, M.; Hargrove, A. E. Small molecule targeting of biologically relevant RNA tertiary and quaternary structures. *Cell Chem. Biol.* **2021**, *28*, 594-609.

(10) Katkevics, M.; MacKay, J. A.; Rozners, E. Triplex-forming peptide nucleic acids as emerging ligands to modulate structure and function of complex RNAs. *Chem. Commun.* **2024**, *60*, 1999-2008.

(11) Zhou, Y.; Kierzek, E.; Loo, Z. P.; Antonio, M.; Yau, Y. H.; Chuah, Y. W.; Geifman-Shochat, S.; Kierzek, R.; Chen, G. Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. *Nucleic Acids Res.* **2013**, *41*, 6664.

(12) Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. Incorporation of thio-pseudoisocytosine into triplex-forming peptide nucleic acids for enhanced recognition of RNA duplexes. *Nucleic Acids Res.* **2014**, *42*, 4008-4018.

(13) Sato, T.; Sakamoto, N.; Nishizawa, S. Kinetic and thermodynamic analysis of triplex formation between peptide nucleic acid and double-stranded RNA. *Org. Biomol. Chem.* **2018**, *16*, 1178-1187.

(14) Lee, E. T. T.; Sato, Y.; Nishizawa, S. Small molecule–PNA oligomer conjugates for rRNA A-site at neutral pH for FID assays. *Chem. Commun.* **2020**, *56*, 14976-14979.

(15) Kim, K. T.; Chang, D.; Winssinger, N. Double-Stranded RNA-Specific Templated Reaction with Triplex Forming PNA. *Helv. Chim. Acta* **2018**, *101*, e1700295.

(16) Tähtinen, V.; Granqvist, L.; Murtola, M.; Strömberg, R.; Virta, P. 19F NMR Spectroscopic Analysis of the Binding Modes in Triple-Helical Peptide Nucleic Acid (PNA)/MicroRNA Complexes. *Chem. Eur. J.* **2017**, *23*, 7113-7124.

(17) Zhan, X.; Deng, L.; Chen, G. Mechanisms and applications of peptide nucleic acids selectively binding to double-stranded RNA. *Biopolymers* **2022**, *113* (2), e23476.

(18) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*, 1497-1500.

(19) Brodyagin, N.; Katkevics, M.; Kotikam, V.; Ryan, C. A.; Rozners, E. Chemical approaches to discover the full potential of peptide nucleic acids in biomedical applications. *Beilstein J. Org. Chem.* **2021**, *17*, 1641-1688.

(20) Saarbach, J.; Sabale, P. M.; Winssinger, N. Peptide nucleic acid (PNA) and its applications in chemical biology, diagnostics, and therapeutics. *Curr. Opin. Chem. Biol.* **2019**, *52*, 112-124.

(21) Li, M.; Zengeya, T.; Rozners, E. Short Peptide Nucleic Acids Bind Strongly to Homopurine Tract of Double Helical RNA at pH 5.5. *J. Am. Chem. Soc.* **2010**, *132*, 8676-8681.

(22) Zengeya, T.; Gupta, P.; Rozners, E. Triple Helical Recognition of RNA Using 2-Aminopyridine-Modified PNA at Physiologically Relevant Conditions. *Angew. Chem., Int. Ed.* **2012**, *51*, 12593-12596.

(23) Hnedzko, D.; McGee, D. W.; Karamitas, Y. A.; Rozners, E. Sequence-selective recognition of double-stranded RNA and enhanced cellular uptake of cationic nucleobase and backbone-modified peptide nucleic acids. *RNA* **2017**, *23*, 58-69.

(24) Ryan, C. A.; Brodyagin, N.; Lok, J.; Rozners, E. 2-Aminopyridine Nucleobase Improves Triple Helical Recognition of RNA and DNA when Used Instead of Pseudoisocytosine in Peptide Nucleic Acids. *Biochemistry* **2021**, *60*, 1919-1925.

(25) Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. Nucleobase-Modified PNA Suppresses Translation by Forming a Triple Helix with a Hairpin Structure in mRNA In Vitro and in Cells. *Angew. Chem., Int. Ed.* **2016**, *55*, 899-903.

(26) Endoh, T.; Brodyagin, N.; Hnedzko, D.; Sugimoto, N.; Rozners, E. Triple-Helical Binding of Peptide Nucleic Acid Inhibits Maturation of Endogenous MicroRNA-197. *ACS Chem. Biol.* **2021**, *16*, 1147-1151.

(27) Ryan, C. A.; Rahman, M., Md; Kumar, V.; Rozners, E. Triplex-Forming Peptide Nucleic Acid Controls Dynamic Conformations of RNA Bulges. *J. Am. Chem. Soc.* **2023**, *145*, 10497–10504.

(28) Gupta, P.; Zengeya, T.; Rozners, E. Triple helical recognition of pyrimidine inversions in polypurine tracts of RNA by nucleobase-modified PNA. *Chem. Commun.* **2011**, *47*, 11125-11127.

(29) Muse, O.; Zengeya, T.; Mwaura, J.; Hnedzko, D.; McGee, D. W.; Grewer, C. T.; Rozners, E. Sequence Selective Recognition of Double-Stranded RNA at Physiologically Relevant Conditions Using PNA-Peptide Conjugates. *ACS Chem. Biol.* **2013**, *8*, 1683-1686.

(30) Okeke, C. U.; Miura, H.; Sato, Y.; Nishizawa, S. Kinetic analysis of highly effective triplex formation between a small molecule–peptide nucleic acid conjugate probe and the influenza A virus RNA promoter region at neutral pH. *Org. Biomol. Chem.* **2023**, *21*, 3402-3410.

(31) Krishna, M. S.; Wang, Z.; Zheng, L.; Bowry, J.; Ong, A. A. L.; Mu, Y.; Prabakaran, M.; Chen, G. Incorporating G-C Pair-Recognizing Guanidinium into PNAs for Sequence and Structure Specific Recognition of dsRNAs over dsDNAs and ssRNAs. *Biochemistry* **2019**, *58*, 3777-3788.

(32) Kesy, J.; Patil, K. M.; Kumar, S. R.; Shu, Z.; Yong, H. Y.; Zimmermann, L.; Ong, A. A. L.; Toh, D.-F. K.; Krishna, M. S.; Yang, L.; et al. A Short Chemically Modified dsRNA-Binding PNA (dbPNA) Inhibits Influenza Viral Replication by Targeting Viral RNA Panhandle Structure. *Bioconjugate Chem.* **2019**, *30*, 931-943.

(33) Krishna, M. S.; Toh, D.-F. K.; Meng, Z.; Ong, A. A. L.; Wang, Z.; Lu, Y.; Xia, K.; Prabakaran, M.; Chen, G. Sequence- And Structure-Specific Probing of RNAs by Short Nucleobase-Modified dsRNA-Binding PNAs Incorporating a Fluorescent Light-up Uracil Analog. *Anal. Chem.* **2019**, *91*, 5331-5338.

(34) 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.; et al. General recognition of U-G, U-A, and C-G pairs by double-stranded RNA-binding PNAs incorporated with an artificial nucleobase. *Biochemistry* **2019**, *58*, 1319-1331.

(35) Kumpina, I.; Baskevics, V.; Walby, G. D.; Tessier, B. R.; Farshineh Saei, S.; Ryan, C. A.; MacKay, J. A.; Katkevics, M.; Rozners, E. Synthesis of 2-Aminopyridine-Modified Peptide Nucleic Acids. *Synlett* **2024**, *35*, 649-653.

(36) Kumar, V.; Rozners, E. Fluorobenzene Nucleobase Analogues for Triplex-Forming Peptide Nucleic Acids. *ChemBioChem* **2022**, *23*, e202100560.

(37) Masquida, B.; Westhof, E. On the wobble GoU and related pairs. *RNA* **2000**, *6* (1), 9-15.

(38) Wild, K.; Weichenrieder, O.; Leonard, G. A.; Cusack, S. The 2 Å structure of helix 6 of the human signal recognition particle RNA. *Structure* **1999**, *7*, 1345-1352.

(39) Wedekind, J. E.; McKay, D. B. Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nat. Struct. Biol.* **1999**, *6*, 261-268.

(40) Correll, C. C.; Wool, I. G.; Munishkin, A. The Two Faces of the Escherichia coli 23 S rRNA Sarcin/Ricin Domain: The Structure at 1.11 Å Resolution. *J. Mol. Biol.* **1999**, *292*, 275-287.

(41) Kimsey, I.; Al-Hashimi, H. M. Increasing occurrences and functional roles for high energy purine-pyrimidine base-pairs in nucleic acids. *Curr. Opin. Struct. Biol.* **2014**, *24*, 72-80.

(42) Kumpina, I.; Baskevics, V.; Nguyen, K. D.; Katkevics, M.; Rozners, E. Nucleobase and Linker Modification for Triple-Helical Recognition of Pyrimidines in RNA Using Peptide Nucleic Acids. *ChemBioChem* **2023**, *24*, e202300291.

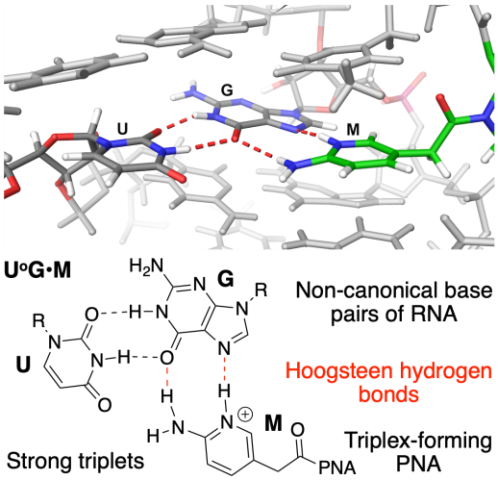
(43) Ryan, C. A.; Baskevics, V.; Katkevics, M.; Rozners, E. 2-Guanidyl pyridine PNA nucleobase for triple-helical Hoogsteen recognition of cytosine in double-stranded RNA. *Chem. Commun.* **2022**, *58*, 7148-7151.

(44) Gautheret, D.; Konings, D.; Gutell, R. R. A major family of motifs involving G•A mismatches in ribosomal RNA. *J. Mol. Biol.* **1994**, *242*, 1-8.

(45) Santa Lucia, J., Jr.; Turner, D. H. Structure of (rGGCGAGCC)2 in solution from NMR and restrained molecular dynamics. *Biochemistry* **1993**, *32*, 12612-12623.

(46) Wu, M.; Turner, D. H. Solution Structure of (rGCGGACGC)2 by Two-Dimensional NMR and the Iterative Relaxation Matrix Approach. *Biochemistry* **1996**, *35*, 9677-9689.

TOC Image


G
U
M
U°G·M
G
U
M
Non-canonical base pairs of RNA
Hoogsteen hydrogen bonds
Strong triplets
Triplex-forming PNA
PNA