RNA triplexes : from structural principles to biological and biotech applications

Devi, Gitali; Zhou, Yuan; Zhong, Zhensheng; Toh, Desiree-Faye Kaixin; Chen, Gang

2014

Devi, G., Zhou, Y., Zhong, Z., Toh, D.-F. K., & Chen, G. (2015). RNA triplexes: from structural principles to biological and biotech applications. Wiley interdisciplinary reviews : RNA, 6(1), 111-128.

<https://hdl.handle.net/10356/103792>

<https://doi.org/10.1002/wrna.1261>

© 2014 John Wiley & Sons, Ltd. This is the author created version of a work that has been peer reviewed and accepted for publication by Wiley Interdisciplinary Reviews: RNA, John Wiley & Sons, Ltd. It incorporates referee' s comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<http://dx.doi.org/10.1002/wrna.1261>].

Downloaded on 26 Apr 2023 22:52:10 SGT

RNA Triplexes – From Structural Principles to Biological and Biotech Applications

Gitali Devi[#], Yuan Zhou[#], Zhensheng Zhong[#], Desiree-Faye Kaixin Toh[#], and Gang Chen^{*}

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

^{*}To whom correspondence should be addressed. G.C.: Tel: +65 6592 2549; Fax: +65 6791 1961; Email: RNACHEN@ntu.edu.sg.

[#] These four authors contributed equally to this work.

Abstract

The diverse biological functions of RNA are determined by the complex structures of RNA stabilized by both secondary and tertiary interactions. An RNA triplex is an important tertiary structure motif that is found in many pseudoknots and other structured RNAs. A triplex structure usually forms through tertiary interactions in the major or minor groove of a Watson–Crick base paired stem. A major-groove RNA triplex structure is stable in isolation by forming consecutive major-groove base triples such as U·A-U and C⁺·G-C. Minor-groove RNA triplexes, e.g., A-minor motif triplexes, are found in almost all large structured RNAs. As double-stranded RNA stem regions are often involved in biologically important tertiary triplex structure formation and protein binding, the ability to sequence-specifically target any desired RNA duplexes **by triplex formation** would have great potential for biomedical applications. Programmable chemically-modified triplex-forming oligonucleotides (TFOs) and triplex-forming peptide nucleic acids (PNAs) have been developed to form TFO·RNA₂ and PNA·RNA₂ triplexes, respectively, with enhanced binding affinity and sequence-specificity at physiological conditions. Here we **(i) provide an overview of naturally-occurring RNA triplexes, (ii) summarize the experimental methods for studying triplexes, and (iii) review the development of TFOs and triplex-forming PNAs for targeting an HIV-1 ribosomal frameshift-inducing RNA, a bacterial ribosomal A-site RNA, and a human miRNA hairpin precursor, and for inhibiting the RNA-protein interactions involving human RNA-dependent protein kinase (PKR) and HIV-1 viral protein Rev.**

1. Examples of naturally-occurring RNA triplexes

An RNA triplex is a complex structure stabilized by multiple base triples¹⁻⁷ formed between an RNA duplex and a third strand **of a natural RNA or an artificial nucleic acid**. A pre-formed RNA duplex may accommodate a third strand to form a major-groove and a minor-groove triplex, respectively, without disrupting the pre-formed duplex structure. Naturally-occurring triplexes are important for shaping RNAs into complex three-dimensional architectures and are important for diverse biological functions.^{4,8,9} For example, triplex structures form in biologically important RNAs including telomerase RNAs,¹⁰⁻¹⁵ metabolite-sensing riboswitches,^{7,16-23} -1 ribosomal frameshift-inducing mRNA pseudoknots,²⁴⁻²⁹ long noncoding RNAs (lncRNAs),³⁰⁻³³ group I introns,³⁴⁻³⁶ group II introns,³⁷ and ribosomal RNAs.³⁸⁻⁴⁰

1.1 RNA major-groove triplexes

An RNA poly(U)·poly(A)-poly(U) major-groove triplex containing many U·A-U base triples (**Figure 1A**) is stable in the presence of Mg²⁺ ions.^{41,42} In the triplex structure, the poly(A) strand forms Watson–Crick A-U and Hoogsteen U·A pairs with two poly(U) strands, respectively (see **Figure 2** for short triplex structures). The Hoogsteen poly(U) strand is parallel to the poly(A) strand. The major-groove triplexes with the third pyrimidine strand parallel to the purine strand of the duplex are known as pyrimidine or parallel motif triplexes. In this review, we will focus on pyrimidine motif RNA triplexes, although other types of major-groove RNA triplexes may also form.⁴³

The natural pyrimidine motif triplex structure **with three consecutive major-groove U·A-U base triples** was first found to form in the human telomerase RNA pseudoknot in the absence of Mg²⁺ ions (**Figure 2A,B**).^{10,13} The human telomerase RNA pseudoknot is a hairpin-type

(H-type) pseudoknot, with two stems (stem 1 and stem 2) and two loops (loop 1 and loop 2) (**Figure 2A,B**).^{10,13} The major-groove triplex structure with U·A-U, C⁺·G-C, or U·G-C base triples (**Figure 1**) formed between stem 2 and loop 1 in telomerase RNA pseudoknot structure is highly conserved between humans and other vertebrates, and is present in yeasts and ciliates, suggesting it is essential for function.¹⁰⁻¹⁵ The pyrimidine motif major-groove triplex is particularly stable in H-type pseudoknots, and is critical for diverse functions including telomerase activity,¹⁰⁻¹⁵ metabolite sensing by riboswitches in mRNA untranslated regions,^{18,23} and stimulating -1 ribosomal frameshifting in mRNA coding regions.⁴⁴

The duplex region of a triplex may not be perfectly Watson–Crick base paired.^{15,18} NMR studies have revealed that a major-groove triplex structure forms in *Kluyveromyces lactis* telomerase RNA pseudoknot even in the presence of bulges in stem 2.¹⁵ The bulge residues may be extruded and not involved in triplex formation.¹⁵ Thus, one may consider the possible formation of bulges in the RNA duplex stems, when conducting computational searching and modelling of RNA major-groove triplexes. The presence of a bulge destabilizes the duplex, the triplex, and thus the pseudoknot structure.^{10,13,15,44,45} For example, the bulge U177 residue in human telomerase RNA pseudoknot (**Figure 2A,B**) was deleted to stabilize the structure for the initial NMR studies.^{10,13}

Bulge nucleotides however, can have important functions. For example, the intracellular sensing of a metabolite S-adenosylmethionine (SAM) by a SAM-II riboswitch present in proteobacterial is facilitated by the molecular recognition of a bulge structure containing U44 residue in stem 2 of an H-type pseudoknot (**Figure 2C,D**).¹⁸ In the ligand-free form of SAM-II riboswitch pseudoknot, stem 2 contains the U44 bulge and a 3 × 3 nucleotide internal loop with the potential to form three consecutive non-Watson–Crick pairs. In the absence of SAM

binding, stem 2 is dynamic and the triplex structure (**Figures 2C,D** and **3B-F**) does not form.^{18,46,47} The crystal structure of the SAM-II pseudoknot in complex with SAM-II reveals that the adenine moiety of SAM intercalates into the bulge region in stem 2 and forms a U·U·A base triple (**Figure 3D**) with the bulge U44 in stem 2 and U10 in loop 1 (**Figure 2C,D**).¹⁸ Evidently, SAM binding induces stabilization of the pseudoknot structure by stabilizing the stem and major-groove triplex structure, which in turn sequesters the ribosome binding site resulting in the repression of protein translation.¹⁸

A pyrimidine motif major-groove triplex structure may also form in an internal loop-type (I-type) pseudoknot structure (**Figure 2E,F**).⁴⁸ For example, in several **viral polyadenylated nuclear (PAN)** long noncoding RNAs (lncRNAs), the poly(A) tail and an upstream U-rich internal loop within a stem-loop structure may form a relatively large I-type pseudoknot structure **as evidenced by biochemical and X-ray crystallographic studies**.^{30,31} The upstream stem-loop structure containing the U-rich internal loop is called the expression and nuclear retention element (ENE) (**Figure 2E**), serving as a *cis* element for stabilizing the lncRNAs by sequestering the poly(A) tail from degradation.^{30,31} The crystal structure of the **truncated ENE core from Kaposi's sarcoma-associated herpesvirus (KSHV) PAN** lncRNA in complex with a 9-nucleotide oligo(A) sequence shows that the “*trans* pseudoknot” structure is stabilized by Watson–Crick base-paired stems as well as extensive tertiary interactions including a pyrimidine motif major-groove triplex (**Figures 1A and 2E,F**) and an A-minor motif triplex (**Figures 2E,F and 4A,B**, see **Section 1.2**).³⁰ Thus, the formation of both major- and minor-groove triplex structures in the I-type pseudoknot may contribute to the repression of degradation of the poly(A) tail and thus protection of viral polyadenylated lncRNAs.^{30,31}

The ENE with a stem-loop structure containing a U-rich internal loop is also found in non-polyadenylated cellular lncRNAs, i.e., MALAT1 (metastasis-associated lung adenocarcinoma transcript 1) and MEN β (multiple endocrine neoplasia- β).^{32,33} Similar to viral lncRNAs, the U-rich internal loop of an ENE and a downstream A-rich tract of a cellular lncRNA may form an I-type pseudoknot with extensive pyrimidine motif major-groove triplex structures and A-minor motif interactions.³² A similar triplex structure was independently proposed for these two cellular lncRNAs with slightly different sequences; the only difference in the structure is that the stem away from hairpin loop (bottom stem) (see **Figure 2E**) is absent.³³ Thus, lncRNAs without poly(A) may also be protected from degradation by forming the pseudoknot and triplex structures.

Clearly, a pyrimidine motif major-groove RNA triplex structure is stable in isolation or within a pseudoknot or other large structures. Note that non-standard major-groove base triples other than pyrimidine motif (see **Figure 3B,C,F** for example) may also form in a pseudoknot or other structures, facilitated by topological constraints and specific structural environments.^{1,2,5-7,18,34-37}

1.2 RNA minor-groove triplexes

Minor-groove triplexes, usually not stable in isolation, often form within large RNAs and RNA-protein complexes. The A-minor motif triplex (**Figures 2 and 4**) is a recurrent motif present in almost all large RNAs, including group I intron,³⁴⁻³⁶ group II intron,³⁷ riboswitches,^{7,16-23} ribosomal RNAs,³⁸⁻⁴⁰ and the mRNA-tRNA-ribosomal RNA complex structure at the decoding site of ribosome.^{40,49-51} A standard A-minor motif^{39,40,49-53} involves tertiary base-base and base-backbone interactions between the “smooth” sugar edge (including the nucleobase atoms C2-N3 and ribose 2'-OH group, see **Figure 1A**) and

Watson–Crick edge (including the nucleobase atoms N1–C2, see **Figure 1A**) of adenine and the minor groove of a Watson–Crick pair in a duplex (**Figure 4A,B**).

A wobble G·U pair (**Figure 4C**) or a face-to-face imino G·A (cis Watson–Crick/Watson–Crick) pair in a duplex has the G amino group exposed in the minor groove and may thus allow ideal A-minor motif interactions.^{34,38,40,54–58} In large RNAs, the third strand of a triplex and a fourth strand may form internal loops with consecutive non-Watson–Crick pairs, such as sheared A·G or A·A (trans Hoogsteen/sugar) pairs (**Figure 4C**).^{2,34,40,56–59} Thus, an RNA triplex in a large RNA may be part of a quadruplex structure with two strands forming a Watson–Crick canonical stem and the other two forming a non-Watson–Crick base-paired loop (**Figure 4C**). Presumably, formation of local non-Watson–Crick loop structures^{34,40,52,60–64} and topological (steric) constraints of secondary structures^{4,65,66} contribute to the (de)stabilization of A-minor motif and other triplex structures in large RNAs.

The minor-groove triplexes have intrinsic structural plasticity, and all three edges of the residues in the third strand may interact with the minor groove of a Watson–Crick duplex. For example, in a “twisted A-minor motif” triplex, successive adenosine residues in an RNA strand may rotate along the minor groove of a duplex such that the tertiary contacts are formed through the Hoogsteen, Watson–Crick, and sugar edges (see **Figure 1**) of the adenosine residues, respectively (**Figure 4D–H**).¹⁸ The “twisted A-minor motif” triplex is often found in an H-type pseudoknot, with the base triples formed between stem 1 and loop 2 (**Figures 2A–D** and **4D–H**). H-type pseudoknots containing “twisted A-minor motif” triplexes are critical for the functions of biologically important RNAs such as telomerase RNA,^{10,13} metabolite-sensing riboswitches,^{7,18–22} –1 ribosomal frameshift inducing mRNA

pseudoknots,²⁴⁻²⁹ tRNA like structures in some plant viral RNA genomes,⁶⁷ and transfer-messenger RNA (tmRNA) pseudoknots essential for bacterial trans-translation.⁶⁸

Non-pseudoknot structures may also contain non-standard A-minor motif triplexes.^{1-3,5-7,16,17,34-37,53} A-minor motif interactions often coexist with coaxial stacking interactions in both pseudoknot and non-pseudoknot structures.² Intrinsic flexibility of minor-groove triplexes allows for the adenosine residues in the third strand to be replaced with guanosine or other residues (see **Figure 4I** for example).^{1-3,5-7} Adenosine is more prevalent than guanosine in forming minor-groove triplexes in large RNAs, probably because adenosine has a smooth sugar/Watson–Crick edge (N1-C2-N3, **Figure 1A**) and guanosine tends to be involved in misfolding by forming relatively stable locally base-paired structures (i.e., Watson–Crick G-C, and non-Watson–Crick G·U, G·A, and G·G pairs).^{69,70}

2. Assessing triplex formation

As discussed above, X-ray crystallography and NMR provide the direct evidence for the RNA triplex structure formation. We will summarize the reported methods for assessing the thermodynamics, kinetics, and dynamics of triplex formation.

2.1 Ensemble methods

In a UV-absorbance detected thermal melting experiment,^{10,71-76} a triplex structure is often disrupted in two steps by increasing the solution temperature. The third strand usually dissociates before the “melting” of the duplex structure and both of the two-state structural disruption reactions can be monitored by UV absorbance increase (hyperchromicity). The two transitions may merge into one if the binding of the third strand is tight. Thermodynamic parameters of triplex formation can be obtained from the melting curves if the transitions are

in equilibrium. However, triplex formation is usually slow, and the thermal melting of triplexes may not always be in equilibrium.

In a non-denaturing polyacrylamide gel electrophoresis (PAGE) or Electrophoretic Mobility Shift Assay (EMSA) experiment,^{71,72,76} a titration curve may be obtained by having the duplex or a hairpin at a constant concentration below the dissociation constant (K_D) and varying concentrations of the third strand. The duplex and the third strand may be annealed and pre-incubated to make sure an equilibrium state is reached. Hairpins and triplexes can be resolved in the gel because they have different electrophoretic mobilities. The gel may be imaged by post-staining with ethidium bromide or by labelling the RNA duplex with ^{32}P or a dye molecule. Standard binding curves between a duplex and a third strand can be obtained to extract the K_D values. UV thermal melting and EMSA are relatively fast and inexpensive in providing useful thermodynamic information, but quantitative kinetic information is relatively hard to obtain.

Isothermal Titration Calorimetry (ITC) has been utilized for the thermodynamic characterization of RNA structures including triplexes.⁷⁷⁻⁸³ The advantage of ITC is that the characterization is done in solution and no labelling is needed. The disadvantage is that often a large amount of sample is needed and kinetic information is relatively hard to obtain.

Surface Plasmon Resonance (SPR),^{76,84,85} is useful in detecting bimolecular interactions in real time, and thus for measuring the kinetics of the triplex structure formation. To facilitate SPR measurement, the duplex or the third strand needs to be immobilized on a surface. The third strand or the duplex may then be injected into the flow cell and the binding and

dissociation can be monitored in real time. SPR measurement usually requires small amount of samples.

Kinetics and thermodynamics of triplex formation may also be obtained by Fluorescence Resonance Energy Transfer (FRET).⁸⁶ For the FRET measurement, the duplex and the third strand need to be labelled with donor and acceptor dyes, respectively.

Chemical and enzymatic mapping^{18,87-89} may also be utilized for assessing the thermodynamics and kinetics of triplex formation at single-nucleotide resolution. The mapping techniques are based on the fact that the triplex structure formation causes the protection of functional groups, and thus inhibits the reaction with small molecule chemicals, and binding and activity of ribonucleases, respectively. Triplex formation can be inferred based on the changes in the reactivity of individual residues in the duplex or the single strand.

2.2 Single-molecule methods

The single-molecule FRET (smFRET) method⁹⁰ has been used to detect RNA triplex formation in a pseudoknot (**Figure 2A,B**) at various solution conditions.^{45,46} smFRET technique also facilitates the direct revelation of (un)binding kinetics of the third strand to a duplex.⁹¹ To facilitate smFRET characterization of dynamics and kinetics of triplex formation, the duplex or the third strand needs to be immobilized on a surface.

Single-molecule nanomanipulation by optical tweezers⁹⁰ allows the direct measurement of the effect of triplex formation on the mechanical stability of a pseudoknot (**Figure 2A,B**).⁴⁴ In an optical tweezers single-molecule pulling experiment, mechanical stretching force is applied on the terminal ends of a pseudoknot **or other molecules**.^{44,90,92} To facilitate the

pulling experiment, a single molecule may be immobilized between two micrometer-sized beads through two long double-stranded nucleic acid “molecular handles”. A correlation was observed among triplex structures, pseudoknot mechanical stability **with the force applied at the two terminal ends of the pseudoknots**, and the efficiency of -1 ribosomal frameshifting in vitro.⁴⁴ The results suggest that triplex structures are critical for stimulating -1 ribosomal frameshifting, when the other factors are kept largely constant.⁴⁴ Pseudoknot-induced -1 ribosomal frameshifting may also be affected by many other factors including the slippery sequence, spacer length, pseudoknot folding kinetics, and interactions with ribosome.^{28,44,92}

3. How to design RNA triplexes to specifically target RNA duplexes

Recognition of RNA duplexes through intermolecular major-groove triplex formation with unmodified third strands (Hoogsteen strands) is typically limited to a length of more than 10 base triples and low pH condition.^{71,72,75,76,87} We will be focused on reviewing the development of replacing the unmodified Hoogsteen strand with relatively short chemically modified triplex-forming oligonucleotides (TFOs) (**Figure 5A-D**) and triplex-forming peptide nucleic acids (PNAs) (**Figures 5E and 6**) for targeting RNA duplexes to form major-groove TFO·RNA₂ and PNA·RNA₂ triplexes, respectively.

3.1 Targeting RNA duplexes through major-groove TFO·RNA₂ triplex formation

Extensive studies have been carried out to specifically target any sequence of DNA duplexes by triplex formation.⁹³ The rules for the DNA triplex formation may not always apply for RNA triplex formation, and the disadvantage of targeting RNA duplexes by TFOs is that TFOs often bind more tightly to DNA than to RNA duplexes.^{71,72,75,76,87,94} For example, UV-absorbance detected thermal melting and EMSA studies reveal that DNA and 2'-OMe RNA (**Figure 5C**) strands can only bind to DNA but not RNA duplexes.^{71,72,75,76,87,94} In addition to

base-base hydrogen bonds (**Figures 1 and 3A**), a backbone-backbone hydrogen bond **formed between the 2'-OH from the TFO and a non-bridging phosphate oxygen in the purine strand** stabilizes an RNA triplex (**Figure 3E**).^{3,30,72,74-76} The 2'-OH-phosphate hydrogen bond is **missing** with a DNA or 2'-OMe RNA as the third strand. In addition, the 2'-OMe group may cause steric clashes with the RNA duplex. Taken together, DNA and 2'-OMe RNA strands are unable to bind to RNA duplexes to form triplexes due to the loss of a hydrogen bond and steric clash. However, enhanced formation of a TFO·DNA₂ triplex was observed with 2'-OMe modification in a TFO, probably due to the geometrical compatibility with the major groove of a DNA duplex.^{76,94,95}

A Locked Nucleic Acid (LNA) modification (**Figure 5D**) in a TFO, however, stabilizes a TFO·RNA₂ triplex as well as a TFO·DNA₂ triplex, as evidenced by thermal melting and EMSA studies.⁷⁶ The stabilization effect of an LNA-modified TFO is probably due to the highly constrained methylene group, resulting in TFO backbone pre-organization and minimization of steric clash, despite the loss of 2'-OH-phosphate hydrogen bond.⁷⁶

Thermal melting and EMSA studies show that 2-thio U base modifications (**Figure 1B**) in a TFO enhances the formation of both TFO·RNA₂ and TFO·DNA₂ triplexes.⁷⁶ Possible stabilizing effects of 2-thio U modification include the enhanced van der Waals contacts, base stacking, hydrogen bonding, preorganized C3'-endo sugar pucker, and reduced dehydration energy. In addition, it has been reported that shifting the pK_a of the hydrogen bond donors in a monomer from far above 7 towards near 7 may enhance the hydrogen bonding strength.⁹⁶ Thus, tuning the pK_a by an atomic mutation **from U to 2-thio U** in a TFO may be useful for enhancing the hydrogen bonding with the Hoogsteen edge of an A base (**Figures 1A,B**) in an RNA duplex for the TFO·RNA₂ triplex formation.⁷⁶

3.2 Targeting RNA duplexes through major-groove PNA·RNA₂ triplex formation

Compared to DNA and RNA, PNAs (**Figures 5E and 6**) are chemically stable and have flexible and neutral peptide-like backbone, and thus show significantly enhanced binding to natural nucleic acids. These characteristics make PNAs attractive for applications in antisense and other nucleic acid-based gene regulation approaches.⁹⁷⁻¹⁰¹ PNAs were originally designed to enhance binding affinity with DNA duplexes to form major-groove PNA·DNA₂ triplexes. However, strand invasion often occurs resulting in the formation of PNA·DNA-PNA complexes with two PNA strands forming Watson–Crick and Hoogsteen base pairs with the purine DNA strand, respectively.^{97,98}

Formation of PNA·RNA₂ triplexes without strand invasion has been reported recently, and thus PNAs show great potential in targeting RNA duplex structures.^{78-83,102} In a pyrimidine motif major-groove PNA·RNA₂ triplex, the PNA strand is parallel to the purine tract in the RNA duplex, i.e., the N-terminus of PNA is aligned with the 5' end of the purine tract of RNA duplex. As evidenced by thermal melting, EMSA, and ITC studies, a PNA·RNA₂ triplex is significantly more stable than a PNA·DNA₂ triplex,^{78,82,102} probably because the RNA duplex major groove provides geometry compatibility and favorable backbone-backbone interactions with PNA. X-ray crystallography and NMR studies are needed to rationalize why a triplex-forming PNA binds to an RNA duplex more tightly than a DNA duplex.

3.2.1 Recognition of Watson–Crick A-U and G-C pairs

As discussed above, a standard major-groove U·A-U base triple (**Figures 1A, 2, and 3E**) is often found in natural RNAs. Thus, a PNA thymidine (T) residue is often used to recognize a

Watson–Crick A-U pair to form a T·A-U PNA·RNA₂ base triple (**Figure 6A**). Due to the relatively low pK_a (about 4.5) of monomer C, a low pH is needed to fully protonate C residues in a triplex-forming PNA to form a C⁺·G-C base triple (**Figures 1C and 6B**). A neutral PNA nucleobase pseudoisocytosine (J) has been developed to form a J·G-C base triple (**Figure 6C**);⁹⁸ the pK_a of the J monomer is about 9.4.^{103,104} Thermal melting, EMSA, and ITC studies show that incorporation of J residues in a PNA significantly alleviates the pH dependence for PNA·RNA₂ triplex formation.^{82,102} A PNA nucleobase 2-aminopyridine (M) with a monomer pK_a of 6.7^{105,106} is promising for the recognition of a Watson–Crick G-C pair to form an M⁺·G-C base triple (**Figure 6D**).⁸² The replacement of cytosine residues by M residues allows the formation of stable and sequence-specific PNA·RNA₂ triplexes at near-physiological pH conditions.

A PNA L nucleobase, a thiolated derivative of the J nucleobase, has been reported for forming an L·G-C base triple (**Figure 6E**).¹⁰² Thermal melting and EMSA studies reveal that, compared to J-modified PNAs, L-incorporated 8-mer PNAs show enhanced affinity and specificity in recognizing the duplex region of a model RNA hairpin to form a PNA·RNA₂ triplex at near-physiological buffer condition (200 mM NaCl, pH 7.5). Importantly, an L-incorporated 8-mer PNA shows relatively weak binding to single-stranded RNAs, single-stranded DNAs, or double-stranded DNAs.¹⁰² Similar to 2-thio U modification (**Figure 1B**) in a TFO as discussed above,⁷⁶ the stabilization of a PNA·RNA₂ triplex by L modification may be due to the combined effects of enhanced van der Waals contacts, base stacking, hydrogen bonding, and reduced dehydration energy. Compared to a standard Watson–Crick G-C pair, the steric clash and loss of two hydrogen bonds in a Watson–Crick-like G-L pair may result in the destabilization of Watson–Crick RNA-PNA and DNA-PNA duplexes.¹⁰² Thus, it is possible to develop PNAs with base and backbone modifications to enhance the

stability of PNA·RNA₂ triplexes and minimize the formation of RNA-PNA duplexes or PNA·RNA-PNA triplexes. Based on the geometric compatibility and hydrogen bond patterns, we expect that **M** and other modified PNA residues (**Figure 6**) may also facilitate the selective binding to RNA duplexes in the major groove rather than single strands.

Taken together, we can conclude that at the current stage, the recognition of Watson-Crick A-U and G-C pairs in an RNA duplex at physiological condition can be facilitated by the formation of major-groove T·A-U, M⁺·G-C or L·G-C PNA·RNA₂ base triples, respectively (**Figure 6A,D,E**).

3.2.2 Recognition of Watson–Crick C-G and U-A pairs

A purine tract in **one of the two strands of** an RNA duplex is needed for a standard pyrimidine motif major-groove triplex formation (**Figures 1, 2, and 7A-C**). However, in natural RNA duplexes, a purine tract **in one strand is often interrupted by pyrimidine residues, resulting in the formation of** inverted Watson–Crick C-G and U-A pairs (pyrimidine inversions, see **Figures 6F-I and 7D-F**). PNA residues with modified bases such as 5-methylisocytosine (iC)^{107,108} and 2-pyrimidinone (P)^{109,110} have been incorporated into PNAs to target Watson–Crick C-G pairs to form iC·C-G and P·C-G base triples, respectively (**Figures 6F,G and 7E**).⁷⁹⁻⁸¹ **ITC measurements show that an** iC-incorporated 6-mer PNA (CTCiCTC) binds to **the target** RNA duplex **containing one inverted** C-G pair with a K_D of 1 μ M in a buffer of 100 mM sodium acetate, pH 5.5.⁸⁰ The binding affinity of the iC-incorporated PNA to the target RNA duplex is slightly enhanced compared to PNAs with the iC replaced by unmodified bases (C, T, G and A), which have the K_D values ranging from 1.4 to 5 μ M. The sequence selectivity is low as the iC-incorporated 6-mer PNA also binds to

other RNA duplexes with the C-G pair replaced by G-C, A-U, or U-A with the K_D values of 1.3 to 2 μM .

To improve the sequence selectivity for the recognition of a Watson–Crick C-G pair, the Rozners group developed a modified PNA P monomer to form a P·C-G base triple (**Figure 6G**).⁷⁹ As evidenced by ITC studies, a 9-mer PNA incorporating a P residue (CTCTCPTCC) binds to the target RNA duplex containing one inverted C-G pair with a K_D of 22 nM in a buffer of 100 mM sodium acetate, pH 5.5. The sequence selectivity is relatively good as the P-incorporated 9-mer PNA binds relatively weakly to other RNA duplexes with the C-G pair replaced by G-C, A-U, or U-A, with the K_D ranging from 0.1 to 0.3 μM . However, the binding affinity of the P-incorporated 9-mer PNA to the target RNA duplex is slightly weakened compared to PNAs with the P replaced by unmodified base C ($K_D = 14$ nM) or T ($K_D = 16$ nM).

To improve the binding affinity, the linker between the PNA backbone and P nucleobase was extended to give a PNA monomer P_{ex} (**Figure 6H**).⁷⁹ ITC measurements reveal that, compared to the P-modified PNA, a P_{ex} -modified 9-mer PNA (CTCTC P_{ex} TCC) can recognize the target RNA duplex containing one inverted C-G pair with significantly enhanced binding affinity with the K_D decreased from 22 nM to 3 nM in a buffer of 100 mM sodium acetate, pH 5.5. The sequence selectivity of P_{ex} -incorporated 9-mer PNA, however, decreases as it also binds to an RNA duplex with the C-G pair replaced by an A-U pair ($K_D = 6$ nM). Interestingly, sequence selectivity is significantly improved upon increasing the pH from 5.5 to 6.25. At pH 6.25, the P_{ex} -incorporated 9-mer PNA binds to the target RNA duplex with one C-G pair ($K_D = 0.2$ μM), with no appreciable binding ($K_D > 100$ μM) to other RNA duplexes with the C-G pair replaced by G-C, A-U, or U-A.

The 3-oxo-2,3-dihydropyridazine (E) nucleobase was designed to recognize a DNA Watson–Crick T-A pair to form a PNA·DNA-PNA base triple E·T-A (see **Figure 6I** for the base triple structure).¹¹¹ As evidenced by ITC studies, a 9-mer PNA (CTCTCETCC) incorporating a PNA E residue (**Figure 6I**) shows selective binding ($K_D = 36$ nM) to a complementary RNA duplex containing a U-A pair at 100 mM sodium acetate, pH 6.25.⁷⁹ The binding is at least about 10 times weaker to an RNA duplex with the U-A pair replaced by an A-U, G-C or C-G pair. Upon decreasing the pH to 5.5, a PNA·RNA₂ triplex containing an E·U-A base triple shows no significant change in binding affinity ($K_D = 26$ nM). Note that replacing the C residues in the 9-mer PNA with M or L (**Figure 6D,E**) is expected to facilitate the binding to the RNA duplex target at more physiological pH.

Similar to the P_{ex}-incorporated PNA (**Figure 6H**), the sequence selectivity of the E-incorporated PNA (**Figure 6I**) decreases upon decreasing the pH from 6.25 to 5.5.⁷⁹ The reduced selectivity of matched base triple over mismatched base triples may be due to the fact that a PNA E residue has a relatively flexible linker – originally designed to avoid the steric clash with the 5-methyl group of a DNA T base.¹¹¹ In addition, at pH 5.5, an E residue may have structural rearrangement, and/or exist in alternative tautomeric forms and protonation states to form mismatched base triples E·G-C and E·A-U.⁷⁹ Thus, one needs to consider conformational plasticity and chemical dynamics^{108,112,113} in designing RNA triplexes to specifically target RNA duplexes. Detailed structural characterizations of the base pairs/triples formed between RNA and PNA are needed to facilitate further modifications for enhancing the binding affinity and selectivity.

In summary, the reported modified PNA nucleobases for the recognition of inverted RNA Watson–Crick C-G and U-A pairs (**Figure 6F-I**) contain only one six-member heterocycle, and thus only one base (i.e., C or U) of the RNA Watson–Crick pair is targeted, resulting in suboptimal binding affinity and selectivity. We expect that enhanced and sequence specific triplex formation at physiological condition will be facilitated by incorporating modified larger bases⁹³ into triplex-forming PNAs through the recognition of not only the pyrimidine base (U or C), but also the purine base (A or G) in inverted Watson–Crick C-G and U-A pairs in an RNA duplex.

In addition, small molecules may be covalently linked with TFOs and triplex-forming PNAs to enhance the binding. In a major-groove triplex structure, the TFO strand divides the original RNA duplex major groove into two grooves (**Figure 2E**),³⁰ and thus provides unique recognition sites for small molecules.¹¹⁴ It has been shown that the triplex stability but not the duplex stability may be selectively enhanced by neomycin, which binds to the triplex groove formed between the Hoogsteen strand and the pyrimidine strand.¹¹⁴

4. Examples of potential applications

Compared to small molecule binders, TFOs and triplex-forming PNAs are advantageous for sequence-specifically binding to RNA duplex structures and can do so in a programmable manner (**Figures 6 and 7**). Cellular delivery of bioactive TFOs and PNAs has been facilitated by attaching small molecules and short peptides.^{83,99-101,115,116} As discussed above, it is possible to develop chemically modified PNAs and possibly TFOs for selectively and tightly binding to RNA duplex but not single-stranded regions.¹⁰² Compared to a traditional antisense strand,^{99-101,115-117} which forms a duplex structure with the target sequence, the advantages of a TFO or a triplex-forming PNA are as follows: (i) The triplex formation

strategy facilitates the targeting of not only the sequence but also the secondary structure of RNA, which results in high selectivity. (ii) The triplex formation does not involve the disruption of the pre-formed RNA secondary structure of the target sequence, which presumably facilitates fast binding kinetics.

TFOs and triplex-forming PNAs are thus promising ligands for inhibition of RNA tertiary (see **Section 1**) and RNA-protein¹¹⁸ interactions involving RNA duplex structures. In addition, TFOs and triplex-forming PNAs may be utilized for the specific stabilization of RNA duplex structures, and thus modulate functions involving RNA (un)folding. Carefully designed TFOs and triplex-forming PNAs may also be useful tools for mapping complex RNA secondary structures **by identifying double helices**. Development of TFOs and triplex forming PNAs for biological applications is still at its early stage; a few examples are discussed below.

4.1 Inhibiting the RNA-dependent protein kinase (PKR) binding to RNA duplexes

RNA-dependent protein kinase (PKR) is an interferon-inducible, RNA-dependent kinase that undergoes autophosphorylation reactions, followed by the phosphorylation of eukaryotic translation initiation factor 2 α (eIF-2 α), resulting in translation inhibition.¹¹⁹ Thus, interactions between cellular PKR and viral RNA duplexes are crucial for the host's antiviral response. Some DNA viruses (e.g., adenovirus and Epstein–Barr virus) produce highly structured inhibitory RNAs that bind to and inactivate PKR, and thus evade the host's antiviral defense.^{120,121} Thus, TFOs or triplex forming PNAs may be designed to bind to the duplex regions of the viral inhibitory RNAs to probe and modulate the interaction between the viral inhibitory RNA and PKR. As a proof of principle study, EMSA and in vitro autophosphorylation assays reveal that unmodified long (>18 nucleotides) RNA TFOs may inhibit the PKR binding to RNA duplexes, through a major-groove triplex formation at near-

physiological condition (**Figure 7A**).¹²² The IC_{50} values are 35 and 210 nM, respectively, for a 28-nucleotide TFO and a 20-nucleotide TFO. Thus, chemically modified TFOs and triplex-forming PNAs have the potential to serve as tools for studying PKR signalling, and as therapeutics to inhibit the interaction between viral decoy RNAs and PKR.

4.2 Inhibiting HIV-1 viral Rev protein binding to Rev response element RNA (RRE)

Assembly of HIV-1 Rev protein on a viral Rev response element RNA (RRE) facilitates the transport of viral genomic and partially spliced viral mRNAs for virus packaging and viral protein translation, respectively.¹²³⁻¹²⁶ Thus, anti-HIV therapeutics may be developed by targeting the highly conserved RRE structure to inhibit Rev binding to RRE. An unmodified chimeric DNA-RNA oligonucleotide was designed to form both duplex (8 base pairs) and triplex (6-12 base triples) structures (**Figure 7B**) with a variant of RRE RNA.¹²⁷ EMSA binding studies show that the binding affinities range from 100 nM to 30 μ M under physiological buffer condition. The simultaneous duplex and triplex formation results in enhanced inhibition of viral Rev protein binding compared to the duplex-forming oligonucleotide alone,¹²⁷ and modulation of other biological activities such as -1 ribosomal frameshift.⁸⁹ Thus, a TFO or a triplex-forming PNA may be covalently attached to a duplex-forming oligonucleotide or PNA, to enhance the binding affinity and selectivity targeting the natural RRE (**Figure 7B**) and other functional RNAs.

4.3 Targeting HIV-1 ribosomal frameshift-inducing RNA

Many viruses including HIV-1 have evolved programmed ribosomal frameshift to express defined ratios of structural and enzymatic proteins. In HIV-1 group M (main), the frameshift stimulating signals of all subtypes are a hairpin.¹²⁸ The frameshift efficiencies of all subtypes are low and within a narrow range of 4-9% *in vitro* and 1-3% in cultured cells.¹²⁸ Thus, TFOs

and triplex-forming PNAs may serve as novel anti-HIV therapeutics by stabilizing frameshift stimulatory RNA structure and/or inhibiting its interaction with ribosome and/or other cellular factors. It is promising that a short PNA (with the sequence of LLTTLL, see **Figure 6A,E**) binds to an HIV-1 frameshift stimulatory RNA by forming a 6-base-triple PNA·RNA₂ triplex (**Figure 7C**) with a binding affinity of 1 μ M at a near-physiological buffer condition (200 mM NaCl, pH 7.5).¹⁰² It is worth noting that an unmodified PNA (CCTTCC) or a J modified PNA (JJTTJJ) does not bind to the HIV-1 frameshift stimulatory RNA.¹⁰²

4.4 Targeting bacterial ribosomal A-site RNA

The A-minor triplex (see **Figure 4A,B** for example) formation facilitates the selection of the cognate tRNAs in the decoding process of mRNA translation.^{40,49-51} The ribosomal A-site RNA contains a 1 $\times$ 2 internal loop (**Figure 7D,E**) in helix 44 of small subunit rRNA. The two A residues on the same side of the loop are in a dynamic equilibrium of two conformations (stacked within helix 44 or bulged out). In the bulged out conformation, the two A residues form A-minor interactions with the mini-helix formed between an mRNA codon and a cognate tRNA anticodon loop.^{40,49-51} However, binding of A-site internal loop-targeting aminoglycoside antibiotics causes the two A residues to bulge out, which stimulates the A-minor interaction with the codon-anticodon mini-helix with either cognate or near-cognate tRNA. The enhanced A-minor interactions upon antibiotics binding cause an increased frequency of the incorporation of near-cognate tRNAs, resulting in aberrant bacterial protein translation.^{49,50}

It would be interesting to see how triplex formation in the duplex region of helix 44 in small subunit rRNA adjacent to the ribosomal A-site 1 $\times$ 2 internal loop may affect ribosome

structure, dynamics, decoding, and translation. New antibiotics based on TFOs and triplex-forming PNAs may be developed targeting the bacterial ribosomal A-site RNA duplex. ITC studies show **binding affinities of about 0.5 μM** for the binding of 8-mer or 9-mer triplex-forming PNAs to the model bacterial A-site RNA duplexes (**Figure 7D,E**) at relatively low pH (6.3 or 5.5).^{79,81} No M or L modifications (**Figure 6D,E**),^{82,102} however, were incorporated in the triplex-forming PNAs, which are known to specifically stabilize PNA·RNA₂ triplexes at physiological pH. The results suggest that a G·U pair and a G·A pair in an RNA duplex can also form base triples with a PNA C residue to form C⁺·G·U and C⁺·G·A triples, respectively. In addition, the two A residues in the 1 × 2 internal loop can be targeted by forming a PNA·RNA base pair (T·A) and a PNA·RNA·RNA base triple (T·A·A), respectively (**Figure 7D,E**). Thus, several mismatches and non-base-paired residues within an RNA duplex are tolerable for the formation of a stable PNA·RNA₂ triplex.

4.5 Targeting miRNA hairpin precursors

Cellular miRNAs are matured from hairpin precursors (pre-miRNAs) by Dicer proteins and are involved in many disease-associated gene regulation processes.¹²⁹ miRNA activity can be inhibited by an antisense strategy through duplex formation using the so-called anti-miR oligomers.^{99,101,116} Inhibition of Dicer binding and thus pre-miRNA maturation by triplex formation may also be a viable strategy. A 10-mer PNA (**Figure 7F**) incorporating three M and one E residues (**Figure 6D,I**) was designed to target a duplex region of a model hairpin derived from pre-miRNA-215.⁸² The ITC measurement reveals a K_D of 83 nM at physiological buffer condition (90 mM KCl, 10 mM NaCl, 2 mM MgCl₂, 50 mM potassium phosphate, pH 7.4 at 37 °C).⁸² The results again suggest that a stable PNA·RNA₂ triplex can form in the presence of several mismatches in the RNA duplex by forming non-standard PNA·RNA·RNA base triples such as T·A·C, T·A·A, and M⁺·G·U (**Figure 7F**).

5. Summary

In summary, naturally-occurring RNA triplexes are emerging as important tertiary structures of many functional RNAs, which inspired the recent development of chemically modified TFOs and triplex-forming PNAs for targeting RNA duplexes with high affinity and sequence specificity. The recent work on triplex-forming PNAs shows that a total of 6-10 consecutive PNA·RNA₂ base triples are sufficient for the formation of a relatively stable and sequence specific triplex at physiological condition. In this review, we discuss only a few examples as the targets for TFOs and triplex-forming PNAs. There is a great potential in developing chemically modified TFOs and triplex-forming PNAs as biological tools and therapeutics targeting many other functional structured RNAs.

Acknowledgment

We thank Dr Walter N. Moss and Prof Xavier Roca for critically reading the manuscript. We thank Nanyang Technological University start-up grant, Singapore Ministry of Education (MOE) Tier 1 (RGT3/13) and MOE Tier 2 (MOE2013-T2-2-024) for the financial support.

References

1. Walberer BJ, Cheng AC, Frankel AD. Structural diversity and isomorphism of hydrogen-bonded base interactions in nucleic acids. *J Mol Biol* 2003, 327:767-780.
2. Xin Y, Laing C, Leontis NB, Schlick T. Annotation of tertiary interactions in RNA structures reveals variations and correlations. *RNA* 2008, 14:2465-2477.
3. Ulyanov NB, James TL. RNA structural motifs that entail hydrogen bonds involving sugar-phosphate backbone atoms of RNA. *New J Chem* 2010, 34:910-917.
4. Butcher SE, Pyle AM. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc Chem Res* 2011, 44:1302-1311.

5. Firdaus-Raih M, Harrison AM, Willett P, Artymiuk PJ. Novel base triples in RNA structures revealed by graph theoretical searching methods. *BMC bioinformatics* 2011, 12 Suppl 13:S2.
6. Abu Almakarem AS, Petrov AI, Stombaugh J, Zirbel CL, Leontis NB. Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res* 2012, 40:1407-1423.
7. Appasamy SD, Ramlan EI, Firdaus-Raih M. Comparative sequence and structure analysis reveals the conservation and diversity of nucleotide positions and their associated tertiary interactions in the riboswitches. *PLoS One* 2013, 8:e73984.
8. Buske FA, Mattick JS, Bailey TL. Potential in vivo roles of nucleic acid triple-helices. *RNA Biol* 2011, 8:427-439.
9. Conrad NK. The emerging role of triple helices in RNA biology. *Wiley Interdiscip Rev RNA* 2014, 5:15-29.
10. Theimer CA, Blois CA, Feigon J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* 2005, 17:671-682.
11. Ulyanov NB, Shefer K, James TL, Tzfati Y. Pseudoknot structures with conserved base triples in telomerase RNAs of ciliates. *Nucleic Acids Res* 2007, 35:6150-6160.
12. Shefer K, Brown Y, Gorkovoy V, Nussbaum T, Ulyanov NB, Tzfati Y. A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Mol Cell Biol* 2007, 27:2130-2143.
13. Kim NK, Zhang Q, Zhou J, Theimer CA, Peterson RD, Feigon J. Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J Mol Biol* 2008, 384:1249-1261.
14. Qiao F, Cech TR. Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol* 2008, 15:634-640.
15. Cash DD, Cohen-Zontag O, Kim NK, Shefer K, Brown Y, Ulyanov NB, Tzfati Y, Feigon J. Pyrimidine motif triple helix in the *Kluyveromyces lactis* telomerase RNA pseudoknot is essential for function in vivo. *Proc Natl Acad Sci USA* 2013, 110:10970-10975.
16. Klein DJ, Ferre-D'Amare AR. Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science* 2006, 313:1752-1756.
17. Cochrane JC, Lipchock SV, Strobel SA. Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chem Biol* 2007, 14:97-105.
18. Gilbert SD, Rambo RP, Van Tyne D, Batey RT. Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat Struct Mol Biol* 2008, 15:177-182.
19. Kang M, Peterson R, Feigon J. Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol Cell* 2009, 33:784-790.
20. Klein DJ, Edwards TE, Ferre-D'Amare AR. Cocystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat Struct Mol Biol* 2009, 16:343-344.
21. Spitale RC, Torelli AT, Krucinska J, Bandarian V, Wedekind JE. The structural basis for recognition of the PreQ0 metabolite by an unusually small riboswitch aptamer domain. *J Biol Chem* 2009, 284:11012-11016.
22. Batey RT. Structure and mechanism of purine-binding riboswitches. *Q Rev Biophys* 2012, 45:345-381.
23. Liberman JA, Salim M, Krucinska J, Wedekind JE. Structure of a class II preQ1 riboswitch reveals ligand recognition by a new fold. *Nat Chem Biol* 2013, 9:353-355.
24. Su L, Chen L, Egli M, Berger JM, Rich A. Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat Struct Biol* 1999, 6:285-292.
25. Michiels PJ, Versleijen AA, Verlaan PW, Pleij CW, Hilbers CW, Heus HA. Solution structure of the pseudoknot of SRV-1 RNA, involved in ribosomal frameshifting. *J Mol Biol* 2001, 310:1109-1123.

26. Nixon PL, Rangan A, Kim YG, Rich A, Hoffman DW, Hennig M, Giedroc DP. Solution structure of a luteoviral P1-P2 frameshifting mRNA pseudoknot. *J Mol Biol* 2002, 322:621-633.
27. Pallan PS, Marshall WS, Harp J, Jewett FC, 3rd, Wawrzak Z, Brown BA, 2nd, Rich A, Egli M. Crystal structure of a luteoviral RNA pseudoknot and model for a minimal ribosomal frameshifting motif. *Biochemistry* 2005, 44:11315-11322.
28. Giedroc DP, Cornish PV. Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res* 2009, 139:193-208.
29. Olsthoorn RC, Reumerman R, Hilbers CW, Pleij CW, Heus HA. Functional analysis of the SRV-1 RNA frameshifting pseudoknot. *Nucleic Acids Res* 2010, 38:7665-7672.
30. Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA. Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* 2010, 330:1244-1247.
31. Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA. Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. *Cell Rep* 2012, 2:26-32.
32. Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA. Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MEN β noncoding RNAs. *Proc Natl Acad Sci USA* 2012, 109:19202-19207.
33. Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* 2012, 26:2392-2407.
34. Adams PL, Stahley MR, Kosek AB, Wang J, Strobel SA. Crystal structure of a self-splicing group I intron with both exons. *Nature* 2004, 430:45-50.
35. Golden BL, Kim H, Chase E. Crystal structure of a phage Twort group I ribozyme-product complex. *Nat Struct Mol Biol* 2005, 12:82-89.
36. Guo F, Gooding AR, Cech TR. Structure of the Tetrahymena ribozyme: base triple sandwich and metal ion at the active site. *Mol Cell* 2004, 16:351-362.
37. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science* 2008, 320:77-82.
38. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The kink-turn: a new RNA secondary structure motif. *EMBO J* 2001, 20:4214-4221.
39. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc Natl Acad Sci USA* 2001, 98:4899-4903.
40. Noller HF. RNA structure: reading the ribosome. *Science* 2005, 309:1508-1514.
41. Felsenfeld G, Davies DR, Rich A. Formation of a 3-stranded polynucleotide molecule. *J Am Chem Soc* 1957, 79:2023-2024.
42. Hoogsteen K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr* 1959, 12:822-823.
43. Carmona P, Molina M. Binding of oligonucleotides to a viral hairpin forming RNA triplexes with parallel G*G*C triplets. *Nucleic Acids Res* 2002, 30:1333-1337.
44. Chen G, Chang KY, Chou MY, Bustamante C, Tinoco I, Jr. Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc Natl Acad Sci USA* 2009, 106:12706-12711.
45. Hengesbach M, Kim NK, Feigon J, Stone MD. Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew Chem Int Ed Engl* 2012, 51:5876-5879.
46. Haller A, Rieder U, Aigner M, Blanchard SC, Micura R. Conformational capture of the SAM-II riboswitch. *Nat Chem Biol* 2011, 7:393-400.
47. Chen B, Zuo X, Wang YX, Dayie TK. Multiple conformations of SAM-II riboswitch detected with SAXS and NMR spectroscopy. *Nucleic Acids Res* 2012, 40:3117-3130.
48. Brierley I, Pennell S, Gilbert RJ. Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nat Rev Microbiol* 2007, 5:598-610.
49. Ogle JM, Brodersen DE, Clemons WM, Tarry MJ, Carter AP, Ramakrishnan V. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 2001, 292:897-902.

50. Lescoute A, Westhof E. The A-minor motifs in the decoding recognition process. *Biochimie* 2006, 88:993-999.
51. Demeshkina N, Jenner L, Westhof E, Yusupov M, Yusupova G. A new understanding of the decoding principle on the ribosome. *Nature* 2012, 484:256-259.
52. Doherty EA, Batey RT, Masquida B, Doudna JA. A universal mode of helix packing in RNA. *Nat Struct Biol* 2001, 8:339-343.
53. Tamura M, Holbrook SR. Sequence and structural conservation in RNA ribose zippers. *J Mol Biol* 2002, 320:455-474.
54. Disney MD, Haidaris CG, Turner DH. Recognition elements for 5' exon substrate binding to the *Candida albicans* group I intron. *Biochemistry* 2001, 40:6507-6519.
55. Disney MD, Turner DH. Molecular recognition by the *Candida albicans* group I intron: tertiary interactions with an imino G.A pair facilitate binding of the 5' exon and lower the K_M for guanosine. *Biochemistry* 2002, 41:8113-8119.
56. Schroeder SJ, Fountain MA, Kennedy SD, Lukavsky PJ, Puglisi JD, Krugh TR, Turner DH. Thermodynamic stability and structural features of the J4/5 loop in a *Pneumocystis carinii* group I intron. *Biochemistry* 2003, 42:14184-14196.
57. Znosko BM, Kennedy SD, Wille PC, Krugh TR, Turner DH. Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and *Candida dubliniensis* group I introns. *Biochemistry* 2004, 43:15822-15837.
58. Chen G, Znosko BM, Kennedy SD, Krugh TR, Turner DH. Solution structure of an RNA internal loop with three consecutive sheared GA pairs. *Biochemistry* 2005, 44:2845-2856.
59. Chen G, Turner DH. Consecutive GA pairs stabilize medium-size RNA internal loops. *Biochemistry* 2006, 45:4025-4043.
60. Leontis NB, Westhof E. Analysis of RNA motifs. *Curr Opin Struct Biol* 2003, 13:300-308.
61. Davis JH, Tonelli M, Scott LG, Jaeger L, Williamson JR, Butcher SE. RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J Mol Biol* 2005, 351:371-382.
62. Schroeder KT, Daldrop P, McPhee SA, Lilley DM. Structure and folding of a rare, natural kink turn in RNA with an A*A pair at the 2b*2n position. *RNA* 2012, 18:1257-1266.
63. Lilley DM. The structure and folding of kink turns in RNA. *Wiley Interdiscip Rev RNA* 2012, 3:797-805.
64. Fiore JL, Nesbitt DJ. An RNA folding motif: GNRA tetraloop-receptor interactions. *Q Rev Biophys* 2013, 46:223-264.
65. Chastain M, Tinoco I, Jr. Nucleoside triples from the group I intron. *Biochemistry* 1993, 32:14220-14228.
66. Bailor MH, Sun X, Al-Hashimi HM. Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science* 2010, 327:202-206.
67. Kolk MH, van der Graaf M, Wijmenga SS, Pleij CW, Heus HA, Hilbers CW. NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. *Science* 1998, 280:434-438.
68. Nonin-Lecomte S, Felden B, Dardel F. NMR structure of the *Aquifex aeolicus* tmRNA pseudoknot PK1: new insights into the recoding event of the ribosomal trans-translation. *Nucleic Acids Res* 2006, 34:1847-1853.
69. Turner DH. Fundamental interactions in RNA: Questions answered and remaining. *Biopolymers* 2013, 99:1097-1104.
70. Grohman JK, Gorelick RJ, Lickwar CR, Lieb JD, Bower BD, Znosko BM, Weeks KM. A guanosine-centric mechanism for RNA chaperone function. *Science* 2013, 340:190-195.
71. Roberts RW, Crothers DM. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* 1992, 258:1463-1466.

72. Escude C, Francois JC, Sun JS, Ott G, Sprinzl M, Garestier T, Helene C. Stability of triple helices containing RNA and DNA strands: experimental and molecular modeling studies. *Nucleic Acids Res* 1993, 21:5547-5553.
73. Plum GE, Pilch DS, Singleton SF, Breslauer KJ. Nucleic acid hybridization: triplex stability and energetics. *Annu Rev Biophys Biomol Struct* 1995, 24:319-350.
74. Holland JA, Hoffman DW. Structural features and stability of an RNA triple helix in solution. *Nucleic Acids Res* 1996, 24:2841-2848.
75. Wilds CJ, Damha MJ. Duplex recognition by oligonucleotides containing 2'-deoxy-2'-fluoro-D-arabinose and 2'-deoxy-2'-fluoro-D-ribose. Intermolecular 2'-OH-phosphate contacts versus sugar puckering in the stabilization of triple-helical complexes. *Bioconjug Chem* 1999, 10:299-305.
76. Zhou Y, Kierzek E, Loo ZP, Antonio M, Yau YH, Chuah YW, Geifman-Shochat S, Kierzek R, Chen G. Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. *Nucleic Acids Res* 2013, 41:6664-6673.
77. Salim NN, Feig AL. Isothermal titration calorimetry of RNA. *Methods* 2009, 47:198-205.
78. Li M, Zenggeya 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.
79. Gupta P, Zenggeya T, Rozners E. Triple helical recognition of pyrimidine inversions in polypurine tracts of RNA by nucleobase-modified PNA. *Chem Commun* 2011, 47:11125-11127.
80. Zenggeya T, Li M, Rozners E. PNA containing isocytidine nucleobase: synthesis and recognition of double helical RNA. *Bioorg Med Chem Lett* 2011, 21:2121-2124.
81. Gupta P, Muse O, Rozners E. Recognition of double-stranded RNA by guanidine-modified peptide nucleic acids. *Biochemistry* 2012, 51:63-73.
82. Zenggeya T, Gupta P, Rozners E. Triple-helical recognition of RNA using 2-aminopyridine-modified PNA at physiologically relevant conditions. *Angew Chem Int Ed Engl* 2012, 51:12593-12596.
83. Muse O, Zenggeya T, Mwaura J, Hnedzko D, McGee DW, Grewer CT, Rozners E. Sequence selective recognition of double-stranded RNA at physiologically relevant conditions using PNA-peptide conjugates. *ACS Chem Biol* 2013, 8:1683-1686.
84. Alberti P, Arimondo PB, Mergny JL, Garestier T, Helene C, Sun JS. A directional nucleation-zipping mechanism for triple helix formation. *Nucleic Acids Res* 2002, 30:5407-5415.
85. Carrascosa LG, Gomez-Montes S, Avino A, Nadal A, Pla M, Eritja R, Lechuga LM. Sensitive and label-free biosensing of RNA with predicted secondary structures by a triplex affinity capture method. *Nucleic Acids Res* 2012, 40:e56.
86. Yang M, Ghosh SS, Millar DP. Direct measurement of thermodynamic and kinetic parameters of DNA triple helix formation by fluorescence spectroscopy. *Biochemistry* 1994, 33:15329-15337.
87. Han H, Dervan PB. Sequence-specific recognition of double helical RNA and RNA.DNA by triple helix formation. *Proc Natl Acad Sci USA* 1993, 90:3806-3810.
88. Mortimer SA, Weeks KM. Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. *Nat Protoc* 2009, 4:1413-1421.
89. Chou MY, Chang KY. An intermolecular RNA triplex provides insight into structural determinants for the pseudoknot stimulator of -1 ribosomal frameshifting. *Nucleic Acids Res* 2010, 38:1676-1685.
90. Tinoco I, Chen G, Qu X. RNA reactions one molecule at a time. *Cold Spring Harbor perspectives in biology* 2010, 2:a003624.
91. Lee IB, Hong SC, Lee NK, Johnner A. Kinetics of the triplex-duplex transition in DNA. *Biophys J* 2012, 103:2492-2501.

92. Ritchie DB, Foster DA, Woodside MT. Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci USA* 2012, 109:16167-16172.
93. Fox KR, Brown T. Formation of stable DNA triplexes. *Biochem Soc Trans* 2011, 39:629-634.
94. Wang S, Kool ET. Relative stabilities of triple helices composed of combinations of DNA, RNA and 2'-O-methyl-RNA backbones: chimeric circular oligonucleotides as probes. *Nucleic Acids Res* 1995, 23:1157-1164.
95. Shimizu M, Konishi A, Shimada Y, Inoue H, Ohtsuka E. Oligo(2'-O-methyl)ribonucleotides - Effective probes for duplex DNA. *FEBS Lett* 1992, 302:155-158.
96. Acharya P, Cheruku P, Chatterjee S, Acharya S, Chattopadhyaya J. Measurement of nucleobase pKa values in model mononucleotides shows RNA-RNA duplexes to be more stable than DNA-DNA duplexes. *J Am Chem Soc* 2004, 126:2862-2869.
97. Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Norden B, Nielsen PE. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* 1993, 365:566-568.
98. Egholm M, Christensen L, Dueholm KL, Buchardt O, Coull J, Nielsen PE. Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA. *Nucleic Acids Res* 1995, 23:217-222.
99. Fabani MM, Abreu-Goodger C, Williams D, Lyons PA, Torres AG, Smith KG, Enright AJ, Gait MJ, Vigorito E. Efficient inhibition of miR-155 function in vivo by peptide nucleic acids. *Nucleic Acids Res* 2010, 38:4466-4475.
100. Shiraishi T, Nielsen PE. Nanomolar cellular antisense activity of peptide nucleic acid (PNA) cholic acid ("umbrella") and cholesterol conjugates delivered by cationic lipids. *Bioconjug Chem* 2012, 23:196-202.
101. Avitabile C, Saviano M, D'Andrea L, Bianchi N, Fabbri E, Brognara E, Gambari R, Romanelli A. Targeting pre-miRNA by peptide nucleic acids: a new strategy to interfere in the miRNA maturation. *Artif DNA PNA XNA* 2012, 3:88-96.
102. 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.
103. Ono A, Tso POP, Kan LS. Triplex formation of oligonucleotides containing 2'-O-methylpseudoisocytidine in substitution for 2'-deoxycytidine. *J Am Chem Soc* 1991, 113:4032-4033.
104. Kan LS, Lin WC, Yadav RD, Shih JH, Chao I. NMR studies of the tautomerism in pseudoisocytidine. *Nucleos Nucleot* 1999, 18:1091-1093.
105. Cassidy SA, Slickers P, Trent JO, Capaldi DC, Roselt PD, Reese CB, Neidle S, Fox KR. Recognition of GC base pairs by triplex forming oligonucleotides containing nucleosides derived from 2-aminopyridine. *Nucleic Acids Res* 1997, 25:4891-4898.
106. Hildbrand S, Blaser A, Parel SP, Leumann CJ. 5-substituted 2-aminopyridine C-nucleosides as protonated cytidine equivalents: Increasing efficiency and selectivity in DNA triple-helix formation. *J Am Chem Soc* 1997, 119:5499-5511.
107. Switzer C, Moroney SE, Benner SA. Enzymatic incorporation of a new base pair into DNA and RNA. *J Am Chem Soc* 1989, 111:8322-8323.
108. Chen G, Kierzek R, Yildirim I, Krugh TR, Turner DH, Kennedy SD. Stacking effects on local structure in RNA: changes in the structure of tandem GA pairs when flanking GC pairs are replaced by isoG-isoC pairs. *J Phys Chem B* 2007, 111:6718-6727.
109. Buchini S, Leumann CJ. Stable and selective recognition of three base pairs in the parallel triple-helical DNA binding motif. *Angew Chem Int Ed Engl* 2004, 43:3925-3928.
110. Gildea B, McLaughlin LW. The synthesis of 2-pyrimidinone nucleosides and their incorporation into oligodeoxynucleotides. *Nucleic Acids Res* 1989, 17:2261-2281.

111. Eldrup AB, Dahl O, Nielsen PE. A novel peptide nucleic acid monomer for recognition of thymine in triple-helix structures. *J Am Chem Soc* 1997, 119:11116-11117.
112. Gilbert SD, Reyes FE, Edwards AL, Batey RT. Adaptive ligand binding by the purine riboswitch in the recognition of guanine and adenine analogs. *Structure* 2009, 17:857-868.
113. Hermann T, Patel DJ. Adaptive recognition by nucleic acid aptamers. *Science* 2000, 287:820-825.
114. Arya DP. New approaches toward recognition of nucleic acid triple helices. *Acc Chem Res* 2011, 44:134-146.
115. Das I, Desire J, Manvar D, Baussanne I, Pandey VN, Decout JL. 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* 2012, 55:6021-6032.
116. Stenvang J, Petri A, Lindow M, Obad S, Kauppinen S. Inhibition of microRNA function by anti-miR oligonucleotides. *Silence* 2012, 3:1.
117. Peacey E, Rodriguez L, Liu Y, Wolfe MS. Targeting a pre-mRNA structure with bipartite antisense molecules modulates tau alternative splicing. *Nucleic Acids Res* 2012, 40:9836-9849.
118. Tian B, Bevilacqua PC, Diegelman-Parente A, Mathews MB. The double-stranded-RNA-binding motif: interference and much more. *Nat Rev Mol Cell Biol* 2004, 5:1013-1023.
119. Nallagatla SR, Toroney R, Bevilacqua PC. Regulation of innate immunity through RNA structure and the protein kinase PKR. *Curr Opin Struct Biol* 2011, 21:119-127.
120. McKenna SA, Kim I, Liu CW, Puglisi JD. Uncoupling of RNA binding and PKR kinase activation by viral inhibitor RNAs. *J Mol Biol* 2006, 358:1270-1285.
121. McKenna SA, Lindhout DA, Shimoike T, Aitken CE, Puglisi JD. Viral dsRNA inhibitors prevent self-association and autophosphorylation of PKR. *J Mol Biol* 2007, 372:103-113.
122. Vuyisich M, Beal PA. Regulation of the RNA-dependent protein kinase by triple helix formation. *Nucleic Acids Res* 2000, 28:2369-2374.
123. Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, Frankel AD, Williamson JR. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* 1996, 273:1547-1551.
124. Jain C, Belasco JG. Structural model for the cooperative assembly of HIV-1 Rev multimers on the RRE as deduced from analysis of assembly-defective mutants. *Mol Cell* 2001, 7:603-614.
125. Moehle K, Athanassiou Z, Patora K, Davidson A, Varani G, Robinson JA. Design of beta-hairpin peptidomimetics that inhibit binding of alpha-helical HIV-1 Rev protein to the rev response element RNA. *Angew Chem Int Ed Engl* 2007, 46:9101-9104.
126. Pond SJ, Ridgeway WK, Robertson R, Wang J, Millar DP. HIV-1 Rev protein assembles on viral RNA one molecule at a time. *Proc Natl Acad Sci USA* 2009, 106:1404-1408.
127. Moses AC, Huang SW, Schepartz A. Inhibition of Rev.RRE complexation by triplex tethered oligonucleotide probes. *Bioorg Med Chem* 1997, 5:1123-1129.
128. Baril M, Dulude D, Gendron K, Lemay G, Brakier-Gingras L. Efficiency of a programmed -1 ribosomal frameshift in the different subtypes of the human immunodeficiency virus type 1 group M. *RNA* 2003, 9:1246-1253.
129. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 2014, 42:D68-73.

Figure 1. Standard major-groove base triples of U·A-U (A), s²U·A-U (B), C⁺·G-C (C), and U·G-C (D). The hydrogen, carbon, nitrogen, oxygen, sulfur atoms are shown in white, cyan, blue, red, and yellow, respectively. The Watson–Crick and Hoogsteen base pairs are indicated by green arrows in the U·A-U base triple (panel A). According to Leontis/Westhof base pairing classification,⁶ in a Watson–Crick pair, the Watson–Crick edges of two bases are involved in hydrogen bonding. In a Hoogsteen pair, the Hoogsteen edge of A or G (containing atoms 5–8) are exposed in the major groove and involved in hydrogen bonding with the Watson–Crick edges of U or C, respectively. The letter R represents the ribose-phosphate backbone, which separates the Hoogsteen edge and the sugar edge. The sugar edge also contains the 2′-OH group in the riboses. Most of the hydrogen bonds are indicated by gray dashed lines. In the base triple C⁺·G-C (panel C), the hydrogen bond formed between protonated N3 in C⁺ and N7 in G is indicated by a red dashed line. The U·A-U, C⁺·G-C, and U·G-C base triple structures are taken from the [RNA Base Triple Database](http://rna.bgsu.edu/triples) (<http://rna.bgsu.edu/triples>).⁶ In the base triple s²U·A-U (panel B, modelled from the U·A-U base triple structure), the van der Waals interaction between the sulfur atom of s²U (2-thio U) and H8 atom of adenine is indicated by a green dashed line.

Figure 2. Triplex and pseudoknot structures. Watson–Crick pairs are indicated by solid lines. Non-Watson–Crick pairs are indicated by dots or dashed lines. The distance cutoff for the hydrogen bonds is 4 Å between two heteroatoms. (A,B) H-type pseudoknot structure found in human telomerase RNA (PDB: 2K96).^{10,13} The bulge U177 residue (between A176 and G178) was deleted to stabilize the structure for the initial NMR studies. (C,D) Crystal structure of an H-type pseudoknot of SAM-II riboswitch in complex with SAM (PDB: 2QWY).¹⁸ (E,F) Crystal structure of a trans pseudoknot formed between a truncated ENE core from KSHV PAN lncRNA and a 9-nucleotide oligo(A) sequence (PDB: 3P22).³⁰ The standard major-groove triplex with five U·A·U base triples is shown in panel E.

Figure 3. Representative major-groove and other base triples formed involving stem 2 in H-type pseudoknots found in human telomerase RNA (panel **A**, see **Figure 2A,B**, PDB: 2K96)^{10,13} and a SAM-II riboswitch in complex with SAM (panels **B-F**, see **Figure 2C,D**, PDB: 2QWY).¹⁸ The distance cutoff for the dashed lines representing hydrogen bonds is 4 Å between two heteroatoms. (**A**) A major-groove U·G·C base triple formed between stem 2 and loop 1 in human telomerase RNA pseudoknot. The standard major-groove U·A·U base triple structures (see **Figure 1A**) are not shown. (**B**) A non-standard major-groove base triple G·G·C. (**C**) A non-standard major-groove base triple A·C·G. (**D**) A base triple U·U·A formed involving the adenine moiety of SAM. (**E**) A standard major-groove U·A·U base triple is shown to highlight the backbone-backbone hydrogen bond formed between 2'-OH in U12 and a non-bridging phosphate oxygen in A46, in addition to the base-base hydrogen bonds shown in **Figure 1**. (**F**) A base triple formed between A13 in loop 1 and a non-Watson-Crick A48·U18 pair.

Figure 4. Representative minor-groove base triples. The carbon atoms are shown in cyan or green. The distance cutoff for the dashed lines representing hydrogen bonds is 4 Å between two heteroatoms. (A,B) Standard A-minor interactions formed between a truncated ENE core from KSHV PAN lncRNA and a 9-nucleotide oligo(A) sequence (see Figure 2E,F, PDB: 3P22).³⁰ In a standard A-minor base triple, adenosine residues in the third strand form hydrogen bonds with a stem utilizing sugar and/or Watson-Crick edges. Type II and type I A-minor interactions are present in panels A and B, respectively. (C) Packing of two duplexes facilitated by A-minor motif interactions between two sheared A·A pairs (A58·A86 and A87·A57) in J4/5 loop and a wobble pair (formed between G and 5-methyl U (m^5U)) in P1 stem in the 5' exon of a group I intron (PDB: 1U6B).³⁴ A58 and A87 are present in two opposite strands with cross-strand stacking, and are involved in type II and type I-like⁵⁰ A-

minor interactions, respectively. **(D,E)** “Twisted A-minor” base triple formed between stem 1 and loop 2 in human telomerase RNA pseudoknot (see **Figure 2A,B**, PDB: 2K96).^{10,13} **A171** (panel **D**) and **A172** (panel **E**) in loop 2 rotate along the minor groove of stem 1 and form minor-groove base triples using Hoogsteen edge (amino group) only, and Hoogsteen and Watson-Crick edges (amino group and N1), respectively. **(F-H)** “Twisted A-minor” base triples formed between stem 1 and loop 2 in a SAM-II riboswitch in complex with SAM (see **Figure 2C,D**, PDB: 2QWY).¹⁸ **A33** (panel **F**) and **A35** (panel **G**) in loop 2 form minor-groove base triples with stem 1 using Watson-Crick edges (amino group and/or N1). In panel **H**, **A35**, **A36**, and **A37** in loop 2 are observed to rotate along the minor groove of stem 1 and form minor-groove base triples using Hoogsteen edge (amino group) only, Watson-Crick (amino group and N1), and Watson-Crick and sugar edges (N1 and N3), respectively. **(I)** A minor-groove base triple involving **U38** residue in loop 2 and two consecutive base pairs **G6·U26** and **C7·G25** in stem 1 present in a SAM-II riboswitch (see **Figure 2C,D**, PDB: 2QWY).¹⁸

Figure 5. Chemical structures of DNA (**A**), RNA (**B**), 2'-OMe RNA (**C**), LNA (**D**), and PNA (**E**).

Figure 6. Major-groove base triples formed between PNA bases (shown in blue) and RNA Watson–Crick pairs (A-U, G-C, C-G, and U-A, shown in black). The letter R represents the sugar-phosphate backbone of RNA. **Most of** the hydrogen bonds between the bases are indicated by **gray** dashed lines. **The hydrogen bonds formed between protonated N3 in C⁺ and N7 in G (panel B), and protonated N1 in M⁺ and N7 in G (panel D) are indicated by red dashed lines.** Hydrogen bonds involving the backbones may also be present and need to be confirmed by future structural studies. J (panel C) is for pseudoisocytosine. M⁺ (panel D) is for **the protonated form of** 2-aminopyridine. L (panel E) is for thio-pseudoisocytosine. In the base triple L·G-C, the van der Waals interaction between the sulfur atom of L and H8 atom of G is indicated by a green dashed line. iC (panel F) is for 5-methylisocytosine. P and Pex

(panels **G** and **H**) are for 2-pyrimidinone base with carbonyl methylene and carbonyl ethylene linkages to the PNA backbone, respectively. E (panel **I**) is for 3-oxo-2,3-dihydropyridazine.