Experimental and computational investigations of RNA duplexes containing N7-regioisomers of adenosine and LNA-adenosine

Ilyas Yildirim[1*†], Witold Andralojc[2†], Amirhossein Taghavi[1,3], Daniel Baranowski[2], Zofia Gdaniec[2], Ryszard Kierzek[2*], Elzbieta Kierzek[2*]

[1] Department of Chemistry and Biochemistry, Florida Atlantic University, Jupiter, FL 33458 USA

[2] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

[3] Department of Chemistry, The Scripps Research Institute, 130 Scripps Way, Jupiter, FL 33458 USA

† These authors contributed equally to this work

* Authors to whom correspondence should be addressed:

iyildirim@fau.edu, rkierzek@ibch.poznan.pl, elzbieta.kierzek@ibch.poznan.pl

ABSTRACT

RNA has a broad range of roles in cellular processes, including regulation of gene expression, translation, and formation of molecular machinery. Its implication in disease progression makes RNA a precious target for treating currently untreatable disorders. Modified RNA residues offer unique properties that can be utilized to control binding affinity, selectivity, and biostability of RNA. In this study, we conducted comprehensive experimental and computational investigations to elucidate the structural and thermodynamic properties of N7-ribofuranosyladenine (7A) and its locked nucleic acid analog ($7A^L$). Our results demonstrate that 7A and $7A^L$ enhance thermodynamic stabilities of 1×1 mismatches when paired with purines, while exhibiting the opposite effect when paired with pyrimidines. Utilizing NMR and computational techniques, we discovered that 1×1 7A:A and $7A^L$:A prefer *anti-anti* conformations, while 1×1 7A:G and $7A^L$:G prefer *syn-anti* orientations, both forming two hydrogen bond states, resulting in enhanced duplex stabilities. Additionally, UV melting data indicates that 7A and $7A^L$ enhance duplex stabilities specifically when used at 5′-dangling ends. The unique properties of 7A and $7A^L$ can advance structure-based RNA studies when resolving properties of dynamic RNA regions while stabilizing RNA mismatches involving purines, thereby potentially yielding novel therapeutic strategies targeting RNA-associated diseases.

INTRODUCTION

Central dogma of biology states that DNA transcribes to RNA and RNA translates to protein, which is an oversimplification of the functional roles of RNA in cell biology.(1) For example, ribosomal RNA (rRNA),(2-4) microRNAs (miRNAs),(5,6) RNA aptamers,(7-12) and riboswitches(13,14) can function as catalysts or bind to specific targets to regulate gene expression in cell. Furthermore, genome of several viruses such as hepatitis papilloma virus (HPV), human immunodeficiency virus (HIV), smallpox, influenza and SARS-based viruses are made of RNA. Moreover, several genetic diseases are caused by expanded RNA repeats causing myotonic dystrophy type 1 and 2 (DM1/DM2),(15-22) Fragile X-associated tremor ataxia syndrome (FXTAS),(23-25) Huntington's Disease (HD),(26-29) and amyotrophic lateral sclerosis/frontotemporal dementia (ALS/FTD).(30-34) As a result, new approaches and technologies to understand unique structural and/or thermodynamic properties of RNA could provide unique solutions in RNA-centered therapeutics.

Binding affinity, selectivity, and biostability of RNA are some of the properties, which can be regulated by modified RNA residues in oligonucleotide therapeutics. For example, locked nucleic acid (LNA) residues have been shown to enhance binding affinities and specificities in RNA and DNA,(35-38) and, thus, have been exploited for different biomolecular purposes such as protecting short interfering RNAs (siRNA) against ribonucleases,(39,40) gene silencing(41,42) and RNA splicing.(43,44) LNA residues are unnatural nucleotides, where O2′ and C4′ atoms are connected by a methylene bridge locking the sugar pucker specifically to C3′-endo conformation. This unique property of LNAs affects neighboring residues in RNA and DNA to favor C3′-endo conformations, thereby inducing global structural changes in the duplex.(45-48) The rigidification of the sugar pucker, and, therefore, the structure is one of the reasons why LNA residues enhance duplex stabilities. The investigation and extraction of such unique properties of modified RNA residues are, thus, critical as they hold the potential to offer novel applications in oligonucleotide-based therapeutics for human diseases.

Even though glycosidic bonds in purines involve N9 position, the formation of N7-ribofuranosyladenine (7A), which is a kinetically favorable ribosylation product during chemical synthesis of adenosine, has been known for years (**Figure 1AB**). Many efforts have been dedicated to N9-regioisomers of adenosine while 7A has been totally neglected and considered as side-product. A similar result was observed in the chemical synthesis of LNA-

adenosine (A^L), where a minor product, N7-LNA-adenosine ($7A^L$), was formed as well (**Figure 1CD**). While there is a limited amount of information available on 7A and $7A^L$, several reports investigated the biological activities of N7-guanosine (7G) and its derivatives. Hunziker et al.(49) and Radhakrishnan et al.(50) observed that substitution of the cytosine with 7G in a triple helix can increase its stability by changing the base-pairing pattern. Koshlap et al. showed that 7G mimics the arrangement of hydrogen bond donors seen in protonated cytosine and form hydrogen bonding scheme with Watson-Crick GC pairs within a triplex to increase specificity and binding affinity compared to guanine.(51) Rao et al. studied several oligonucleotides, where 2′-deoxyguanosine was replaced by its 7-regioisomer, and showed that this exchange did not enhance formation of antiparallel triple helix.(52) In turn, Clair et al. showed that acyclic 7G is capable of forming DNA triplexes, where they concluded that amino group of the modified unit could be involved in recognition of G in GC base pairs. (53) More recently, Leonard and co-workers compared the base pairing properties of 7-deoxyriboguanine and its 9-deaza analog in duplexes and triplexes.(54,55) Moreover, Seela and Zulauf investigated the effects of N7-2′-deoxyadenosine and N7-2′-deoxy-3-deazaguanosine in the duplex stabilities of DNA, where the nucleoside composition of oligomers containing the nucleosides was determined by hydrolysis with snake-venom phosphodiesterase followed by alkaline phosphatase.(56) As described above, studies involving N7-regioisomers of adenosine are lacking that can reveal unique properties of these systems. Thus, detailed studies of 7A and $7A^L$ are required that can provide unique solutions in biophysical and therapeutic studies.

RNA molecules are inherently flexible, which creates unique structural motifs such as internal, hairpin, multibranch, and bulge loops to yield unique functional roles to RNA important in cell biology.(57,58) Depending on the preferred sugar pucker, base orientation with respect to sugar, and stacking, RNA residues can adopt a variety of configurations in loop motifs that can affect stability. For example, while C3′-endo sugar pucker is mainly preferred by RNA residues, C2′-endo is another conformation they can adopt. Furthermore, RNA residues can adopt *syn* and *anti*-orientations with regards to the glycosidic bond angle, known also as χ torsional angle, that can dramatically change the conformational stability. The 1×1 RNA internal loop motifs are one of the simplest systems one can utilize to understand the properties of single mismatches, which have been a focus in multiple studies.(59-63) Furthermore, conformational variability adopted by internal loops can be exploited for

therapeutic purposes.(64-68) Incorporation of modified nucleotides such as A^L, 7A, and $7A^L$ can affect the dynamics of RNA. Thus, discovery of the unique properties of these modified RNA residues can have potential uses in the design principles of oligonucleotides targeting RNA in cell.

In this contribution, we report a comprehensive study on the thermodynamic stabilities of RNA duplexes incorporating 7A and $7A^L$ at select positions using both experimental and computational methods. UV melting experiments were performed on RNA duplexes holding 1×1 X:Y mismatches (X=A, A^L, 7A, $7A^L$; Y=A, C, G, U) to measure duplex hybridization free energies. Extensive structural studies using molecular dynamics (MD) calculations supported by Nuclear Magnetic Resonance (NMR) experiments were performed to explain both the structural and thermodynamic properties of RNA duplexes carrying 7A, A^L, and $7A^L$ residues. Thermodynamics data verified by MM/3D-RISM calculations displays that 7A and $7A^L$ destabilize RNA duplexes when paired with U and C and stabilize RNA duplexes when paired with A and G. NMR data and MD calculations provide evidence that both 1×1 7A:G and $7A^L$:G mismatches adopt *syn-anti* orientations. These orientations give rise to states characterized by two hydrogen bonds and intramolecular salt-bridge-like interactions. Remarkably, these structural features are akin to those observed in 1×1 G:G internal loop motifs. Furthermore, we determined that 1×1 7A:A and 1×1 $7A^L$:A both adopt *anti-anti* orientations and form two hydrogen bond states creating more stable duplexes than 1×1 A:A, which has a single hydrogen bond. Finally, compared to A and A^L, we found out that 7A and $7A^L$ enhance duplex stabilities by 0.6 kcal/mol when used at 5′-dangling ends of a duplex and destabilize duplexes dramatically when used as 3′-dangling ends. We propose that properties of 7A and $7A^L$ can be utilized in stabilizing RNA mismatches involving purines in structure-based RNA studies, thereby potentially yielding novel therapeutic strategies targeting RNA-associated diseases.

MATERIALS AND METHODS

Synthesis of protected N7-regioisomer of adenosine and its phosphoramidite. N7-regioisomer of adenosine (7A) was synthetized according to published procedure with some modifications.(69) Adenine was suspended in hexamethyldisilazane (HMDS, 4 mL/1 mM adenine) and 10 mg of ammonium sulfate per 1 mM of adenine was added and refluxed over 16 h. Reaction mixture was then evaporated, dissolved in anhydrous acetonitrile (4 mL/1 mM scale of synthesis), combined with ß-D-ribofuranose 1,2,3, 5-tetraacetate (1.5 equivalent) and trimethylsilyl

trifluoromethanesulfonate (TMSOTf, 1.5 equivalent), and finally stirred at room temperature (RT) for 80 min. After performing thin layer chromatography (TLC) analysis, a saturated aqueous solution of sodium bicarbonate was added to the reaction mixture, which was then extracted three times with dichloromethane. The combined organic layers were dried over anhydrous sodium sulfate, filtered, and evaporated. Silica gel column chromatography purification was performed, and acetylated N7-adenosine was isolated with an approximate yield of 60%. The acetylated N7-adenosine was initially dissolved in methanol (2 mL per 1 mM) and subsequently mixed with 25% aqueous ammonia (4 mL per 1 mM). After TLC analysis confirmed the complete removal of acetyl groups within 2 hours at RT, the reaction mixture was first evaporated and then co-evaporated with pyridine three times. Afterwards, the mixture was co-evaporated with anhydrous methanol three times. The resulting product was then suspended in anhydrous methanol (3 mL per 1 mM scale of synthesis). N,N-dimethylformamide dimethyl acetal (2.5 equivalents per 1 mM substrate) was added to the reaction mixture, and the mixture was stirred for 48 hours. Once the reaction completion was confirmed by TLC, water was added to the mixture, which was then evaporated after 15 minutes. After co-evaporating the reaction mixture with anhydrous pyridine three times, dimethoxytrityl chloride (1.05 equivalents per 1 mM used for the synthesis of N7-adenosine) was added to the mixture and stirred at RT for 2 hours. Again, after confirming reaction completion by TLC, a saturated aqueous solution of sodium bicarbonate was added to the mixture, which was then extracted three times with dichloromethane. The combined organic layers were dried over anhydrous sodium sulfate, filtered, and evaporated. Solution was first evaporated and co-evaporated with toluene. The reaction mixture was purified by silica gel column chromatography using a dichloromethane eluent with a methanol gradient. The overall yield of the last three steps was approximately 50%. The final step was protection of 2′-hydroxyl with tert-butyldimethylsilyl protecting group. Product was evaporated three times with anhydrous pyridine and dissolved in anhydrous pyridine (5 mL/1 mM). To the solution of 5′-O-dimethoxytrityl-N6-dimethylformamidine-N7-adenosine, tert-butyldimethylsilyl chloride (1.2 equivalents) and imidazole (2.5 equivalents) were added, and the mixture was stirred for 16 hours at RT. After verifying the completion of the reaction with TLC, an aqueous solution of sodium dihydrogen phosphate was added to the mixture, which was then extracted three times with dichloromethane. The combined organic layers were extracted with aqueous solution of sodium bicarbonate, dried

over anhydrous sodium sulfate, filtered, and evaporated. The solution was co-evaporated with toluene three times before being purified by silica gel column chromatography using a dichloromethane eluent with a gradient of ethyl acetate. The yield of the synthesis of the 2′-silyl 7N-adenosine derivative was approximately 45%. The structure of N7-adenosine was verified through 1D and 2D ^{1}H and ^{13}C NMR experiments (see SI, section "NMR spectra of 7A and $7A^{L}$"). Protected 3′-O-phosphoramidite of N7-A was prepared according to published procedures using 2-cyanoethyl N,N,N′,N′-tetraisopropylphosphorodiamidite and equivalent amount of tetrazole.(70,71) The yield of the phosphoramidite after silica gel column chromatography was approximately 80%.

Synthesis of protected N7-regioisomer of LNA-adenosine and its phosphoramidite. The N7-regioisomer of LNA-adenosine ($7A^{L}$) was a minor product with a yield of approximately 20% during the synthesis of N9-LNA-adenosine (A^{L}). The condensation reaction utilized 4-C-methanesulfonyloxymethyl-1,2-di-O-acetyl-3,5-di-O-benzyl-D-ribofuranose and HMDS-silylated adenine in the presence of trimethylsilyl triflate in anhydrous acetonitrile. The remaining seven steps of the reaction leading to 5′-O-dimethoxytrylit-N6-isobutyryl-LNA-adenosine were similar to the procedure described by Koshkin et al.(43,72) During silica gel column chromatography, the protected derivative of N7-LNA-A was isolated with an approximate yield of 20%. The structure of $7A^{L}$ was confirmed through 1D and 2D 1H and 13C NMR experiments (see SI, section "NMR spectra of 7A and $7A^{L}$"). Protected 3′-O-phosphoramidite of $7A^{L}$ was prepared according to published procedure using 2-cyanoethyl N,N,N′,N′-tetraisopropylphosphorodiamidite and equivalent amount of tetrazole.(70,71) After silica gel column chromatography, the yield of the phosphoramidite was approximately 80%.

Oligonucleotide synthesis. Oligonucleotides were synthesized on a BioAutomation MerMade12 DNA/RNA synthesizer using β-cyanoethyl phosphoramidite chemistry and commercially available phosphoramidites (ChemGenes, GenePharma), where standard protocols were followed as described previously.(70,73) For deprotection, oligoribonucleotides were treated with a mixture of 30% aqueous ammonia and ethanol (3:1 v/v) for 16 hours at 55°C. Silyl protecting groups were removed with the use of triethylamine trihydrofluoride. The deprotected oligonucleotides were purified using silica gel TLC in a mixture of 1-propanol, aqueous ammonia, and water (55:35:10 v/v/v), as described previously.(74,75)

UV-melting experiments. Thermodynamic measurements were performed for RNA duplexes at nine different concentrations ranging from 0.1 mM to 1 μM. The measurements were carried out in a buffer containing 1 M sodium chloride, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA at pH 7. The concentrations of single-strand oligonucleotides were calculated using the absorbance at 80 °C, and the extinction coefficients of the single strands were approximated using a nearest-neighbor model.(76,77) Absorbance vs. temperature melting curves were measured at 260 nm with a heating rate of 1°C/min from 0 to 90 °C on JASCO V-650 UV/Vis spectrophotometer with a thermoprogrammer. The melting curves were analyzed, and the thermodynamic parameters were calculated from a two-state model using MeltWin 3.5.(78) For most duplexes, the ΔH^o values derived from the T_M^{-1} vs. $\ln(C_T/4)$ plots were within 15% of those derived from averaging the fits to individual melting curves, as expected from the two-state model

NMR spectroscopy of RNA duplexes. For the acquisition of the NMR data, all RNA duplexes were dissolved in a 10 mM sodium phosphate buffer (pH 6.8) containing 150 mM sodium chloride and 0.1 mM EDTA. In order to ensure that RNA is present uniquely in the duplex form, the samples were further washed with the same buffer on an Amicon centrifugal filter with 3 kDa molecular weight cutoff, where any excess single-stranded RNA would pass through the pores of the Amicon membrane. NMR data was collected on a Bruker AVANCE III 700 MHz spectrometer equipped with a QCI CryoProbe. The resonance assignment of non-exchangeable aromatic and anomeric protons was achieved using standard procedures through the analysis of NOESY spectra recorded in 100% D2O at 25 °C and 35 °C.(79) In the case of duplexes incorporating $7A^L$:A and A^L:A mismatches, the assignment of the H2 protons of the two adenine bases in the mismatch was confirmed using an HMBC-type bond experiment. The H2 and H8 protons of most adenine residues were correlated to each other through cross-peaks to the same C4 carbon frequency. Unfortunately, for $7A^L$ the H8 resonance was slightly broadened and did not yield HMBC correlations. The H2 proton of this residue was thus assigned based on its atypical C4 chemical shift of 161.2 ppm consistent with monomeric $7A^L$ (**Table S1**). All the H2 protons of "standard" adenosine residues in the same duplex correlated to C4 carbon atoms resonating between 148.8 and 150.0 ppm, which is a typical range for A-C4 carbons. No attempt was made to assign the rest of the ribose protons. The exchangeable protons

were assigned using NOESY spectra measured in 90% H2O / 10% D2O at 5 °C and 25 °C. All spectra were analyzed in NMRFAM-Sparky.(80)

Parameterization of 7A and 7A^{L}. Nucleoside versions of 7A and 7A^{L} were first created with the LEAP module of AMBER 18.(81) Each molecule was first optimized and then electrostatic potentials (ESP) at a set of grid points were calculated with HF/6-31G(d) QM level of theory using Gaussian 09.(82) Restrained electrostatic potential (RESP) charges were derived following the RESP protocol.(83) During the charge calculations of 7A and 7A^{L}, RESP charges of all except the C1′, H1′, and the base atoms were set to be the same as in adenosine and LNA-adenosine, respectively (**Tables S2** and **S3**). Furthermore, χ and α/γ torsional parameters of adenosine and LNA-adenosine were used in 7A and 7A^{L}, respectively. (45,84,85)

Preparation of the model systems for MD studies. 10 RNA systems were created to investigate the properties of 7A and 7A^{L} (**Table S4**). Simulations were carried out with the AMBER 18(81) simulation package using the PARM99 force field(86) with revised χ(87) and α/γ(85) torsional parameters. The NAB module of AMBER 18(81) was utilized to build the initial structures for the unmodified systems of A-U, A-C, A-A, and A-G in A-form RNA orientations, which then were used to homology model 7A-U, 7A-C, 7A-A, 7A-G, 7A^{L}-A, and 7A^{L}-G (see **Table S4** for details). To determine the preferred conformational state for each system, four initial structures were created for each system, where the initial structures of the middle base pairs were designed to be in *syn-syn*, *anti-syn*, *syn-anti*, and *anti-anti* orientations, which allowed us to scan a wider conformational space while determining the global minimum structures. For this purpose, torsional restraints on χ were imposed on these residues to create the initial states in *syn* and *anti* orientations, which were then used in explicit solvent MD simulations without any restraints. Each system was first neutralized with Na^{+} ions,(88) which then was solvated with 9000 TIP3P(89) water molecules in a truncated octahedral box using the LEAP module of AMBER 18.(81) Each system was then added with extra 13 Na^{+} and Cl^{-} ions to mimic physiological conditions, where after equilibration each system had 0.17 M Na^{+} concentrations.

Molecular dynamics simulations. The structures were minimized in two steps, where positional restraints with restraint weights of 1 kcal mol^{-1} $Å^{-2}$ were applied only in the first step while no restraints were imposed in the second step. Each minimization step included 5000 steps of steepest-descent minimization subsequently followed

by 5000 steps of conjugate-gradient minimization. Minimization was followed by an equilibration protocol first in constant volume with positional restraints imposed on all the atoms of RNA with a restraint weight of 1 kcal mol^{-1} $Å^{-2}$. During the first step of equilibration, temperature was increased from 0 to 300 K in 2 ns using the Langevin thermostat with a 1 ps^{-1} collision frequency at constant volume. Another 2 ns of MD run was followed in the second step of equilibration using constant pressure MD with temperature at 300 K, and pressure at 1 bar with pressure relaxation time of 1 ps. No restraints were used in explicit solvent MD simulations. After minimization and equilibration, a 4 μs MD simulation with a 2 ps time step was performed on each system at constant pressure MD with isotropic position scaling. In 7A-A and 7A-G, which were started in *syn-syn* orientations, MD simulations were extended to cover over 6 and 10 μs, respectively (**Table S5**). SHAKE(90) was turned on for constraining bonds involving hydrogen atoms. An atom-based long-range cutoff of 10.0 Å was used in all the minimization and MD runs. PME was used to handle the electrostatics(91). Simulations were performed using pmemd.MPI, pmemd.cuda, and pmemd.cuda.MPI of AMBER 18.

Computational Analyses. Dihedral, hydrogen-bond, and root-mean-square deviation (rmsd) analyses were completed using the cpptraj module of AMBER 18.(81) Hydrogen bond analyses were done only on the middle base pairs to see if formation or loss of hydrogen bonds are correlated with the experimental data (**Script S1**). An in-house code was utilized to perform cluster analyses as described before (20,30,32,33,92-98). For each system, all the four MD trajectories (*anti-anti*, *syn-anti*, *anti-syn*, *syn-syn*) were first combined. Snapshots with root-mean-square deviation (RMSD) <= 0.5 Å of the middle base pair were clustered into the same group. Average structures of each cluster were calculated at the end using all the structures in that cluster. In order to calculate the hybridization energies, we utilized the MM/3D-RISM and NMODE approaches of AMBER 18, where both of the terminal two base pairs were removed prior to calculations (**Script S2**). The reference interaction site model (RISM) of molecular solvation provides the 3D map of density distribution of the solvent around the solute.(99,100) The probability density $\rho_\gamma g_\gamma(r)$ of finding the interaction sites (γ) of solvent molecules positioned around the solute, in the 3D space (r) is used to represent the solvent structure. The Kovalenko-Hirata closure (3D-RISM-KH) (101) molecular solvation theory was utilized in the calculations.

RESULTS.

Determination of the structure of 7A and 7-A^{L}. The structures of 7A and 7A^{L} nucleosides were verified by a series of various NMR experiments. The diagnostic signals that unequivocally confirmed the substitution at N7 position of adenine with D-ribose were heteronuclear long-range correlations C5-H1′ and C8-H1′ observed on ^{1}H-^{13}C gHMBC spectra (see SI, section "NMR spectra of 7A and 7A^{L}", and **Figures S1-S14**). Further evidence supporting N7 substitution is provided by a comparative analysis of the ^{1}H and ^{13}C chemical shifts of compounds A and 7A (see SI, section "NMR spectra of 7A and 7A^{L}", and **Tables S1** and **S6**). Additionally, orientation around glycosidic bond of 7A and 7A^{L} preferred *anti* or *high-anti* orientation and conformation of β-D-ribofuranose moiety was determined as C1′-*exo* and C3′-*endo*, respectively (**Table S7**). These N7-derivatives were used to synthesize the 3′-O-phosphoramitites and respective N7-modified RNA oligonucleotides for thermodynamic UV-melting studies.

Thermodynamic properties of RNA duplexes with 7A and 7A^{L}. Natural RNA and DNA nucleotides contain glycosidic bonds between C1′ of ribose and N9/N1 of purine/pyrimidine. Exception is pseudouridine (Ψ) and its derivatives, which display glycosidic bonds between C1′ and C5 atoms. Herein, we investigate the properties of 7A and 7A^{L} modified RNA oligonucleotides. Inclusion of N7-regioisomer analogues of adenosine in oligonucleotides dramatically changes base paring and thermodynamic properties. In order to perform the thermodynamic studies, we utilized RNA duplexes, such as 5′UCAG**X**CAGU/3′AGUC**Y**GUCA, where X was A, 7A, A^{L}, 7A^{L} and Y was A, C, G and U (**Table 1**). Moreover, the thermodynamic stabilities of duplexes 5′UCAGGCAGU**X**/3′AGUCCGUCA and 5′**X**UCAGGCAGUX/3′AGUCCGUCA containing A, 7A, A^{L}, 7A^{L} at 3′- and 5′-terminal unpaired positions also known as 3′- and 5′-dangling ends were investigated. As a result, model duplexes were forming either canonical base pairs in the middle of the duplexes such as A:U and A^{L}:U, or 1×1 non-canonical base pairs, or carrying 7A, 7A^{L} at the 3′- and 5′-terminal unpaired positions.

Contribution of 7A to the thermodynamic stabilities of RNA duplexes. The major difference between A and 7A is related to the glycoside bond linking ribose to adenine. Inversion of adenine moiety with respect to sugar changes the locations of functional groups involved in base pairing with the complementary strand. To evaluate the effects of 7A, thermodynamic stabilities of different RNA duplexes were measured using UV-melting method (**Table 1**). Duplex hybridization energies, ΔG°_{37}, were measured to be -7.71, -6.93, -8.53 and -7.70 kcal/mol in

7A-U, 7A-C, 7A-A and 7A-G, respectively, whereas ΔG°_{37} of A-U, A-C, A-A, and A-G were measured to be -11.63, -9.00, -7.84 and -7.10 kcal/mol, respectively, suggesting A and 7A to have different structural and thermodynamic properties when paired with pyrimidines and purines. Substitution of A with 7A diminishes thermodynamic stabilities by 3.92 and 2.07 kcal/mol, when paired with U and C, respectively. On the other hand, enhancements in thermodynamic stabilities by 0.69 and 0.60 kcal/mol were observed for 7A-A and 7A-G pairs as compared to A-A and A-G, respectively.

Contribution of $7A^L$ to thermodynamic stabilities of RNA duplexes. Compared to natural RNA residues, locked nucleic acids (LNAs) enhance thermodynamic stabilities of RNA helices by 1.5-1.7 kcal/mol per each LNA residue. (73,102) It was previously reported that improved stacking and hydrogen bonding patterns as well as structural preorganization, which is due to exclusive preference of C3′-endo sugar pucker in LNA residues, were responsible for the enhancements.(45,102) Using UV-melting experiments, duplex hybridization energies for $7A^L$-U, $7A^L$-C, $7A^L$-A and $7A^L$-G were measured to be -9.66, -8.61, -9.60 and -9.89 kcal/mol, respectively, while the corresponding energies for A^L-U, A^L-C, A^L-A and A^L-G were measured to be -12.81, -9.27, -8.13 and -9.32 kcal/mol, respectively (**Table 1**). Substitution of A^L with $7A^L$ indicates that 1×1 $7A^L$:U and 1×1 $7A^L$:C mismatches destabilize the duplexes by 3.15 and 0.66 kcal/mol, respectively (**Table 1**). In contrast 1×1 $7A^L$:A and 1×1 $7A^L$:G mismatches stabilize the duplexes by 1.47 and 0.57 kcal/mol, respectively, compared to 1×1 A^L:A and 1×1 A^L:G, respectively (**Table 1**). These results are similar to the outcomes observed on substitution of A with 7A as described above. When paired with pyrimidines, both 7A and 7AL destabilize the duplexes, whereas when paired with purines, they stabilize them (**Table 1**).

Contributions of terminal unpaired 7A and $7A^L$ to thermodynamic stability of RNA duplexes. Hydrogen bonds and base stackings with adjacent nucleotides are interactions affecting thermodynamic stabilities of RNA duplexes. Depending on the position, sequence, and orientation of closing base pairs, placement of an unpaired nucleotide at 5′- or 3′-terminal position, which are known as 5′- or 3′-dangling ends, can stabilize RNA duplexes to some extent that is due to stacking interactions with the adjacent base pairs. (102-105) In order to determine the properties of A, 7A, A^L and $7A^L$ as dangling ends, we utilized a model RNA duplex, 5′UCAGGCAGU/3′AGUCCGUCA (M1) and studied several RNA duplexes (**Table 1**). We determined that

enhancement in duplex stabilities by -0.48, -1.24, -0.77, and -1.22 kcal/mol were observed in 5′-A, 5′-7A, 5′-A^L, and 5′-7A^L, respectively, displaying an order for the measured ΔG°_{37} as M1 > 5′-A > 5′-A^L > 5′-7A^L > 5′-7A (**Table 1**). On the other hand, loss of duplex stabilities by 0.42 and 0.35 kcal/mole were observed in 3′-7A and 3′-7A^L, compared to M1 (**Table 1**). As expected, utilization of A and A^L as 3′-dangling ends as seen in 3′-A and 3′-A^L enhances duplex stabilities by around -0.48 kcal/mol (**Table 1**). Analysis of these results show that when used as 5′-dangling ends 7A and 7A^L enhance duplex stabilities by 0.6 kcal/mol on average compared to A and A^L. Conversely, when utilized as 3′-dangling ends, they significantly reduce the hybridization energies, where the measured ΔG°_{37} values follow the order of 3′-7AL > 3′-7A > M1 > 3′-A > 3′-AL. Overall, the observed outcomes suggest that stacking properties of 7A and 7A^L are different compared to A and A^L when utilized as dangling ends.

Conformational preferences of 7A and 7A^L according to NMR. Our thermodynamic results display that N7-glycosylated adenine stabilizes duplexes when paired with G and A as compared to the standard N9-glycosylated counterparts in both RNA and LNA series (**Table 1**). We thus decided to investigate the structural basis for such increased stabilities using NMR spectroscopy. Four duplexes containing the 7A-G, 7A^L-G, 7A-A and 7A^L-A pairings (**Table 1**) were obtained in NMR-compatible quantities and studied using 1D (**Figure 2**) and 2D NMR experiments in order to provide hints regarding the mutual arrangement of the two interacting bases. The NMR observables used for this purpose included: 1) the intensities of intra-residue H8/H6-H1′ NOE cross peaks reporting the states of glycosidic bonds of the bases (*syn/anti*) (**Table 2**), 2) the identities of exchangeable protons in the central mismatch to identify groups likely involved in hydrogen bonding interactions (*vide infra*) and 3) a list of NOEs and their intensities, which are structurally significant and involve base protons of the central mismatch (**Tables S8-S11**).

For the 7A-G and 7A^L-G containing duplexes the NMR results show that the central 1×1 mismatches adapt *syn-anti* conformations (**Table 2**). In both systems, the NOE intensities of 7A5H1′-H8 and 7A^L5H1′-H8 are four times greater than the averages established over the other residues throughout the duplexes (**Figures 2** and **3**). For the opposing G14 residue, on the other hand, the intensities of G14H1′-H8 in 7A^L-G and 7A-G are within the said average and slightly above one standard deviation, respectively (**Figures 2** and **3**). In addition, the 7A5H8 and

7A^{L}5H8 protons do not produce the expected cross-peaks to the preceding G4H1′ but correlate to the following C6H1′ (**Figures 2** and **3**), which is a characteristic pattern observed in helices having a single nucleotide in *syn* state.(106) On the other hand, the exchangeable imino proton of G14, G14H1, is observed in 7A^{L}-G but not in 7A-G (**Figure 2**). At low temperatures, the imino resonance in 7A^{L}-G appears to be relatively sharp, but as the temperature increases, it progressively broadens and becomes barely detectable at 35 °C. Such a level of protection against solvent exchange suggests that this resonance is involved in a relatively stable hydrogen bonding interaction. The -NH_2 group of the central G14 was unfortunately not observable in the spectra of both systems, which is expected even for -NH_2 groups of standard Watson-Crick GC pairs.(107) As a result, we could not evaluate the hydrogen bonding states of the amino groups of G14. Interestingly, the imino proton of the central G14 exhibits a relatively strong NOE cross-peak to 7A^{L}5-H2 proton, but not to 7A^{L}5-H8 (**Figure S15**). NOESY spectrum recorded in D_2O also reveals that the 7A5H2 and 7A^{L}5H2 protons give only a single and weak NOE to C6-H5 (**Figures 2** and **3**).

Given the fact that the 7A^{L} is in *syn* conformation, which is supported by the NOE patterns mentioned above, the hydrogen bonding partner of the imino proton of G14 in 7A^{L}-G is most likely the N3 nitrogen of 7A^{L} (**Figure 4).** Such a geometry would also bring the -NH_2 group of G14 close to N9 nitrogen of 7A^{L} suggesting a base pairing stabilized by two hydrogen bonds (**Figures 1**, **2,** and **4**). For the 7A-G system, where the imino proton of G in 1×1 7A:G mismatch is not observable, it is more difficult to confidently establish the base pairing geometry. However, given the fact that both 7A^{L}-G and 7A-G adopt the same glycosidic states and exhibit very similar aromatic and anomeric chemical shift patterns (**Table S12**) it is reasonable to suggest that both systems actually display similar pairing geometries, with perhaps the pairing in 7A-G being more labile, precluding the observation of the hydrogen bonded imino proton.

In 7A-A and 7A^{L}-A (**Table 2**), H1′-H8 NOE cross peak intensities within the central pairs do not differ significantly from other pairs implying that the 7A-A and 7A^{L}-A mismatches adopt the *anti-anti* glycosidic conformations (**Figures 2** and **5**). Unfortunately, no direct information regarding hydrogen bonding patterns can be established for these systems as the -NH_2 protons of all adenine bases in question are missing in the NMR spectra. This is usually the case even for adenosine residues forming stable Watson-Crick A-U pairs.(107) The

H2 protons within the central mismatch exhibit very similar NOE patterns with the neighboring base pairs in both 7A-A and 7AL-A systems (**Figures 2** and **5,** and **Tables S8-S11**). Also, the chemical shifts measured for 7A-A and $7A^{L}$-A are very similar except for H8 proton, which is expected due to the differences in ribose sugar of RNA and LNA residues (**Table S13**). Similar to the results observed in 7A-G and $7A^{L}$-G, the findings suggest that the geometries of the central 1×1 mismatches in 7A-A and $7A^{L}$-A are identical. However, for the 7A-A and $7A^{L}$-A duplexes it is much more difficult to establish the common geometry from the NMR data alone due to the absence of hydrogen bond information. One pairing geometry satisfying the NMR observations is shown in **Figure 4B**. Finally, the systems of $7A^{L}$-C and $7A^{L}$-U were also investigated. However, significant line broadening was observed in the central mismatches, making thorough analysis impossible due to the conformational exchange process occurring in these systems (data not shown).

Predicted binding free energies are in line with experimental results. As described above, N7-glycosylation of adenine as seen in 7A and $7A^{L}$ places the functional groups to different regions on adenine compared to natural N9-glycosylated adenine that sets different properties to 7A and $7A^{L}$ (**Table 1**). In order to describe an in-depth understanding of the properties of 7A and $7A^{L}$, we utilized molecular dynamics (MD) simulations, and performed binding free energy calculations to determine if any structural differences observed in the duplexes containing 7A and $7A^{L}$ can explain the experimental thermodynamics and NMR data (**Table 1**). In order to scan structurally important space, four initial structures were created for each system, where the middle base pairs of each duplex were homology modeled to be in *anti-anti*, *syn-anti*, *anti-syn*, and *syn-syn* orientations (**Table S4**). After cluster and then binding analyses, we found that predicted binding free energies, $\Delta\Delta G_{pred}$, are qualitatively in line with experimental results (**Tables 2** and **S14-S25**). While destabilizing RNA duplexes when paired with pyrimidines, 7A and $7A^{L}$ stabilize systems when paired with purines (**Table 2**). For example, 7A-A, $7A^{L}$-A, 7A-G, and $7A^{L}$-G are predicted to have lower hybridization energies compared to A-A and A-G, which is in line with experimental thermodynamics data (**Tables 1** and **2**). As an example, $\Delta\Delta G_{pred,X}$, defined as $\Delta\Delta G_{pred,X} = \Delta G_{pred,X,} - \Delta G_{pred,Y}$ (X=7A-U, 7A-C, 7A-A, 7A-G, $7A^{L}$-U, $7A^{L}$-C, $7A^{L}$-A, $7A^{L}$-G and Y=A-U, A-C, A-A, A-G), was predicted to be -1.98, -2.04, -0.71, and -1.59 kcal/mol for 7A-A, $7A^{L}$-A, 7A-G, and $7A^{L}$-G, respectively, while experimental thermodynamics data measured the corresponding $\Delta\Delta G^{\circ}_{37}$ values as -0.69, -1.76, -0.60, and -2.79 kcal/mol (**Table**

2). Furthermore, computational predictions also display destabilization in 7A-U, 7A^{L}-U, 7A-C, and 7A^{L}-C, which is also observed in thermodynamics data (**Tables 1** and **2**). We calculated $\Delta\Delta G_{pred}$ values of 7A-U, 7A^{L}-U, 7A-C, and 7A^{L}-C as 3.57, 2.58, 1.17, and 0.42 kcal/mol, respectively, while measured $\Delta\Delta G^{\circ}_{37}$ values are 3.92, 1.97, 2.07, and 0.39 kcal/mol (**Table 2**). After verifying a good correlation between the predictions and experiments, we investigated the structural details extracted from MD calculations.

When paired with pyrimidines, 7A and 7A^{L} destabilize RNA duplexes due to loss of hydrogen bond and stacking. Figure 6A-F displays the global minimum structures predicted for the middle base pairs in A-U, 7A-U, 7A^{L}-U, A-C, 7A-C, and 7A^{L}-C by MD calculations (**Table 2**). As expected, the middle 1×1 A:U base pair in A-U is forming 1.98 hydrogen bonds in *anti-anti* orientation (**Tables 2 and S14**). When A is substituted with 7A and 7A^{L} as it is seen in 7A-U and 7A^{L}-U, respectively, the total number of hydrogen bonds between the middle base pairs drops to 1.16 and 1.00 (**Tables 2**, **S15-S16**, and **Figure 6A-C**). Analyses also display that 1×1 7A:U mismatch is flexible and can form *anti-anti* and *syn-anti* states with relatively similar ΔG_{pred} values (**Table S15**). Furthermore, 7A^{L} in 7A^{L}-U is in *syn* orientation, which is due to the interaction of its amino group, -NH2, with its own O4′ atom (**Figure 6C**). Due to N7-glycosylation of adenine in 7A and 7A^{L}, the amino groups in these residues can interact with the phosphate backbone or with its own O4′ atom by switching to the *syn* orientation as seen in 1×1 7A^{L}:U mismatch (**Figure 6C**). The *anti*-orientation, however, is the favored state of nucleosides.(87) To make an *anti* → *syn* transformation, enough energetics is required such as formation of hydrogen bonds as observed in 1×1 G:G internal loops, which are in *syn-anti* orientations and have two hydrogen bonds.(105,108-110) In 1×1 7A^{L}:U, however, one hydrogen bond is totally lost compared to 1×1 A:U (**Figure 6AC**).

Predicted structure of A-C displays the middle 1×1 A:C base pair forming a single hydrogen bond where C is in *syn* orientation (**Figure 6D**). It is noteworthy to highlight here that 1×1 A:C internal loop is a flexible motif, where cluster analysis displays both *anti-syn* and *anti-anti* orientations forming single hydrogen bonds with relatively similar ΔG_{pred} values (**Table S17**). Nevertheless, 7A-C presents a 1×1 7A:C mismatch, where 7A is in *syn* orientation forming on average 0.97 hydrogen bonds (**Table 2** and **Figure 6E**). Like A-C, 1×1 7A:C mismatch is dynamic, which can also form *anti-anti* orientations with relatively similar ΔG_{pred} values and single hydrogen

bonds (**Table S18**). Similar to the case observed in $7A^{L}$-U, the *anti* → *syn* transformation does not create an extra hydrogen bond with the complementary strand to stabilize the duplex, which explains why $\Delta\Delta G^{\circ}_{37}$ is observed to be 2.07 kcal/mol and ΔG_{pred} is predicted to be 1.17 kcal/mol (**Tables 1** and **2**). The predicted global minimum structure of $7A^{L}$-C displays a similar scenario. Although cluster analyses display this system to have *anti-anti*, *syn-syn*, *syn-anti*, and *anti-syn* orientations with relatively similar ΔG_{pred} values with single hydrogen bonds (**Table S19**), predicted global minimum displays 1×1 $7A^{L}$:C in *anti-anti* orientation preserving the A-form RNA orientation (**Figure 6F**). This might be the reason why the $\Delta\Delta G^{\circ}_{37}$ and $\Delta\Delta G_{pred}$ values to be marginal, with values of 0.39 and 0.42 kcal/mol, respectively (**Tables 1** and **2**).

When paired with adenosine, 7A and $7A^{L}$ stabilize the RNA duplexes by forming base pairs with two hydrogen bonds in *anti-anti* orientations. Previous studies showed that 1×1 A:A base pairs form a single hydrogen bond state in *anti-anti* orientations. (16,29,111-113) Our cluster analyses and binding studies also determined that the global minimum state of the 1×1 A:A mismatch in A-A display similar properties (**Figure 6G** and **Tables 2** and **S20**). When one of the A is substituted with either 7A or $7A^{L}$, however, the 1×1 loop forms a two-hydrogen bond state in *anti-anti* orientation, where each amino group is forming a hydrogen bond with the opposing basic nitrogen located at N1 position (**Figure 6HI** and **Tables S21-S22**). Predictions are also in line with the NMR results predicting 1×1 $7A^{L}$:A mismatch in *anti-anti* orientation (**Table 2**). The geometry observed in the simulations is also in agreement with the measured NOE patterns (**Tables S10-S11**). The formation of an extra hydrogen bond with no distortions at the backbone as seen in **Figure 6HI** is the reason why 7A-A and $7A^{L}$-A are forming more stable duplexes compared to A-A (**Tables 1** and **2**).

When paired with guanosine, 7A and $7A^{L}$ stabilize the RNA duplexes by forming two hydrogen bond states in *syn-anti* orientations similar to the structures observed in 1×1 G:G internal loop motifs. Cluster analyses determined that the 1×1 A:G internal loop in A-G prefers *syn-anti* orientation forming two hydrogen bonds between the amino group of A and carbonyl group of G, and basic nitrogen of A located at N7 and imino group of G (**Figures 1** and **6J,** and **Table S23**). The 1×1 7A:G and 1×1 $7A^{L}$:G mismatches also form two hydrogen bond states in *syn-anti* orientations, where the two basic nitrogen atoms of 7A and $7A^{L}$ located at N3 and N9 form hydrogen bonds with the amino and imino groups of G, respectively (**Figure 6KL**, and **Tables 2**, **S8-S9, and**

S24-S25). The *syn-anti* orientations of 1×1 7A:G and 1×1 7A^{L}:G are in agreement with the NMR observations (**Tables S8-S9**). This state is very similar to the non-canonical base pairs observed in 1×1 G:G internal loop motifs except the location of amino groups.(105,108-110) In 1×1 G:G internal loops, one of G is in *syn* orientation, which displays an intramolecular salt bridge between amino group and phosphate backbone, as well as two hydrogen bonds with the opposite G. As described above, when 7A and 7A^{L} are in *syn* orientations, the amino groups are not forming intramolecular interactions with the phosphate backbone but the O4′ atoms that stabilize the duplexes further (**Figure 6KL**). These interactions can explain why 7A-G and 7A^{L}-G are forming more stable duplexes compared to A-G (**Tables 1** and **2**).

Here, it is worth taking a look at the predicted structure of 1×1 A:G in A-G and the limitation of our binding calculations. Previous NMR studies on 1×1 A:G internal loops determined that these loops form one hydrogen bond states in *anti-syn* orientations opposite to our prediction. (114,115) Our cluster analyses also display the *anti-syn* orientations as stable states but with stabilities less than 2.5 kcal/mol compared to global minimum structure predicted, which is in *syn-anti* orientation (**Table S23**). Binding energy calculations performed on two RNA strands to predict the hybridization energies will anticipate a global minimum, which optimizes the intermolecular interactions such as stacking and hydrogen bonds. In the binding energy calculations, formation of intramolecular salt bridges observed in *anti-syn* form of 1×1 A:G are not included, which can be the reason why this state is not predicted to be the global minimum for A-G by MM/3D-RISM calculations (**Table S23**). Furthermore, distortions at the backbone due to base orientations such as *syn-anti* and *anti-syn* states in 1×1 A:G cannot be captured by MM/3D-RISM, which is another limitation of binding energy calculations.

DISCUSSION

Modification of glycosidic linkage from N9 to N7 as observed in 7A and 7A^{L} result in almost 180° inversion of adenine, which exposes different functional groups and basic nitrogen atoms located at N1, N3, and N9 accessible to form hydrogen bonds with the complementary strands (**Tables S2** and **S3**). Furthermore, steric perturbation of A-form RNA is another consequence of N7-glycosylation when incorporated in RNA duplexes, which can either stabilize or destabilize the systems. As a result, it is expected that substitution of A and A^{L} with 7A and 7A^{L} will yield different properties as we discuss in this contribution. Moreover, while RNA residues can

prefer C2′-endo and C3′-endo sugar pucker conformations, LNA residues locks the sugar pucker into C3′-endo due to a methylene bridge formed between C2′ and C4′ atoms.(116) In ribonucleotide forms, RNA residues are observed to prefer equivalent mixtures of C2′-endo and C3′-endo conformers whereas in ribonucleotide forms they exclusively prefer C3′-endo conformation when forming Watson-Crick base pairs. As a result, differences in the sugar pucker and/or the orientation of the N-glycoside bond will cause structural and thermodynamic differences in A→7A and A^L→$7A^L$ mutations.

A dramatic destabilization was observed when A:U pair was substituted by 7A:U with a $\Delta\Delta G^\circ_{37}$ of 3.92 kcal/mol. Computational work displays 1×1 7A:U pair forming a single hydrogen bond while A:U pairs are known to have two hydrogen bonds. A single hydrogen bond will contribute a stability between 1.0 and 1.5 kcal/mol.(117) As a result, loss of such a single hydrogen bond cannot explain how 1×1 7A:U destabilizes the duplex stability by 3.92 kcal/mol, which suggests that inversion from N9 to N7 perturbs stacking interaction with the closing G:C base pairs. Substitution of the central 1×1 A:C mismatch with 1×1 7A:C results also in destabilization of the duplex stability by $\Delta\Delta G^\circ_{37}$ of 2.07 kcal/mol, but at a much lesser level compared to A:U→7A:U substitution. Computational results display both A:C and 7A:C pairs to form one hydrogen bond implying that stacking interactions with the closing G:C pairs are perturbed in 7A:C, too. Surprisingly, however, A:A→7A:A and A:G→7A:G substitutions result in stabilization of RNA duplexes by 0.69 and 0.60 kcal/mol, respectively. Computational and NMR results display that 7A:A forms two hydrogen bonds in *anti-anti* orientation while A:A is known to prefer a single hydrogen bond state. Furthermore, NMR data and computational results display 1×1 7A:G to prefer *syn-anti* orientation and form 1.54 hydrogen bonds and an intramolecular salt-bridge-like interaction similar to the motif observed in 1×1 G:G internal loops that enhances duplex stabilities in 7A-G.

As described above, LNA residues exclusively lock the sugar pucker to C3′-endo. As a result, any differences observed between A^L and $7A^L$ is related to N9- and N7-glycosylation. Previously, it was reported that substitution of ribonucleotides with LNA analogues enhances duplex stabilities by 1-1.5 kcal/mol per LNA residue due to helical preorganization of single-stranded oligonucleotides.(73,102,118) We have observed that A→A^L substitutions in A-U, A-C, A-A, and A-G enhance the thermodynamic stabilities by 1.18, 0.27, 0.29 and 2.22

kcal/mol, respectively. Nevertheless, we have observed that $A^L \rightarrow 7A^L$ substitutions destabilize $7A^L$-U and $7A^L$-C by 3.15 and 0.66 kcal/mol, respectively, compared to A^L-U and A^L-C, and stabilize $7A^L$-A and $7A^L$-G by 1.47 and 0.57 kcal/mol, respectively, compared to A^L-A and A^L-G. Moreover, NMR data and computational results display the structures of 1×1 $7A^L$:A and 1×1 $7A^L$:G to be in *anti-anti* and *syn-anti* conformational states, similar to the results observed in 7A-A and 7A-G. As described above, in general, substitution of A and A^L with 7A and $7A^L$ diminishes thermodynamic stabilities of RNA duplexes when paired with pyrimidines while enhances thermodynamic stabilities of RNA duplexes when paired with purines.

Hydrogen bonding, stacking, hydrophobic and electrostatic interactions play crucial roles in stabilizing RNA duplexes, where stacking with adjacent nucleobases is one of the dominant contributors in duplex stabilities. The best and most simple method to evaluate stacking properties of a modified nucleotide is through placement of single unpaired versions of it at 5′- or 3′-ends of a duplex, also known as dangling ends. Stacking interactions of residues in an RNA duplex depend on the position of the nucleotide, the sequence and orientation of the base pairs in the duplex. (73,102,117,119) By comparing the hybridization free energies of duplexes with and without the use of dangling ends, one can establish the stacking properties of a modified nucleotide. We, therefore, utilized the dangling ends of an RNA duplex with 7A and $7A^L$ to determine their stacking properties. A core duplex of r(5′-UCAGGCAGU/3′AGUCCGUCA) was utilized for this purpose that indicates that single unpaired 7A destabilizes the duplex by 0.42 kcal/mol and stabilizes it by 1.24 kcal/mol when used at 3′- and 5′-ends, respectively. A similar trend was observed in $7A^L$, too, where placement of single unpaired $7A^L$ at 3′- and 5′-ends of the core duplex causes the system to be destabilized by 0.35 kcal/mol and stabilized by 1.22 kcal/mol, respectively. These results are different than what is observed for the typical N9-regioisomers, which enhance thermodynamic stabilities of RNA duplexes when used as unpaired nucleotides at both ends of the duplex. Furthermore, use of single unpaired natural RNA residues at 3′-dangling ends contribute to stability equally or more than 5′-dangling ends.(120,121)

CONCLUSION

Although mixtures of N9- and N7-regioisomers are obtained during the chemical synthesis of adenosine, it remains unclear why nature would favor N9-adenosine over N7-adenosine. During the prebiotic RNA world era,

nature likely favored RNA molecules that formed thermodynamically stable structures, such as N9-regioisomers of adenosine, uridine, cytidine, and guanosine, allowing the formation of Watson-Crick AU and GC base pairs. There are numerous modified RNA residues available in the literature, where specific sugar and base atoms are modified to manipulate and control the properties of RNA molecules.(122-125) These modifications can enhance the stability and biostability of RNA and improve specificity and selectivity.(124,126) Furthermore, they can be utilized in structural studies to stabilize RNA molecules while investigating specific RNA motifs. For example, 2′-O-methylated RNA and LNA residues stabilize RNA structures by keeping the sugar conformations in C3′-endo conformation.(45) Modified RNA residues, such as m^6A, m^5C, and m^7G, which are methylated at specific sites, can stabilize the RNA structures due to the hydrophobic properties of the methyl groups. Inosine, a modified adenosine residue, displays structural properties of adenosine and guanosine, where the base region includes a carbonyl group at C6 and an imino group at N1, yielding unique base-pairing properties compared to adenosine and guanosine. Pseudouridine (Ψ) is another modified residue, displaying an imino group at the C5 position of uridine, generating unique base pairing capabilities. While 4-thiouridine (s^4U) residues destabilize RNA duplexes, 2-thiouridine (s^2U) enhances duplex stabilities dramatically because of the modification substantially stabilizing the C3′-endo sugar conformation.(127)

The extended repertoire of RNA residues offers complex uses, such as the utilization of N1-methylguanosine (m^1G) in structural studies, which, when paired with guanosine residues, will specifically stabilize 1×1 m^1G:G mismatches in *syn-anti* orientations.(34) In this context, the structural and thermodynamic properties of 7A and $7A^L$ can provide unique applications in structural studies. For example, 1×1 A:A internal loops observed in RNA C<u>A</u>G repeat expansions causing Huntington's Disease are known to interact with muscleblind-like 1 protein (MBNL1). This interaction can make use of 7A and $7A^L$ to investigate the specifics of RNA C<u>A</u>G/MBNL1 interaction. Furthermore, 1×1 G:G internal loops are known to form *syn-anti* orientations, which are observed in RNA C<u>G</u>G and G_4C_2 repeat expansions causing fragile X-associated tremor/ataxia syndrome (FXTAS) and amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), respectively.(34) By mutating one of the guanosines in 1×1 G:G with 7A, one can force the 1×1 G:G mismatches, which prefer *anti-syn* orientations, to transform into 1×1 7A:G mismatches, which prefer *syn-anti* orientations. These kinds of mutation studies can

provide invaluable help in structural studies, such as deciphering specific chemical shifts in NMR experiments, which are hard to deconvolute otherwise.

DATA AVAILABILITY

All data is contained within the manuscript and/or supplementary files. Computational data that support the findings of this study are available from the corresponding author upon reasonable request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Computations were performed using the High-Performance Computing (HPC) cluster, KoKo, at the Florida Atlantic University.

FUNDING

National Institute of Health [R15 GM146199 to I.Y.]; Florida Atlantic University startup grant (to I.Y.); David and Lynn Center for Degenerative Disease Research program (to I.Y.); National Science Center (Poland) grants [UMO-2021/41/B/NZ1/03819 to E.K., UMO-2019/33/B/ST4/01422 to R.K., UMO-2020/37/B/ST4/03182 to W.A., and UMO-2018/31/D/ST4/01467 to W.A.].

Table 1. Thermodynamic parameters of RNA duplexes containing adenosine (A), N7-adenosine (7A), LNA-adenosine (A^L) and N7-LNA-adenosine ($7A^L$).[a]

Duplexes (5′-3′)	Short Notation	Average of curve fits					T_M^{-1} vs log C_T plots				
		-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_M[b] (°C)	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_M[b] (°C)	ΔΔG°$_{37}$ (kcal/mol)	ΔT_M[b] (°C)
5′ UCAGACAGU 3′ AGUCUGUCA	A-U	75.9 ± 2.4	206.1 ± 7.2	12.01 ± 0.24	61.1	68.9 ± 10.0	184.5 ± 30.5	11.63 ± 0.63	61.8	0.00	0.0
5′ UCAG7ACAGU 3′AGUC U GUCA	7A-U	71.1 ± 5.7	204.4 ± 18.6	7.73 ± 0.14	42.3	61.4 ± 3.6	173.2 ± 11.5	7.71 ± 0.06	43.1	3.92	-18.7
5′ UCAGA^LCAGU 3′AGUCU GUCA	A^L-U	83.3 ± 2.3	224.5 ± 6.6	13.66 ± 0.32	66.0	72.4 ± 5.5	192.2 ± 16.4	12.81 ± 0.44	66.4	-1.18/0.00[c]	4.6/0.0[d]
5′ UCAG7A^LCAGU 3′ AGUC U GUCA	7A^L-U	74.6 ± 2.8	209.8 ± 8.5	9.47 ± 0.26	49.7	80.8 ± 7.6	229.5 ± 23.8	9.66 ± 0.28	49.5	1.97/3.15[c]	-12.3/-16.9[d]
5′ UCAGACAGU 3′ AGUCCGUCA	A-C	78.4 ± 3.9	224.3 ± 12.4	8.87 ± 0.14	46.5	89.7 ± 12.7	260.4 ± 40.7	9.00 ± 0.21	45.8	0.00	0.0
5′ UCAG7ACAGU 3′AGUC C GUCA	7A-C	74.0 ± 5.6	216.1 ± 17.9	6.99 ± 0.16	38.9	82.7 ± 14.1	244.4± 45.6	6.93 ± 0.38	38.5	2.07	-7.3
5′ UCAGA^LCAGU 3′ AGUCC GUCA	A^L-C	71.3 ± 0.9	199.2 ± 2.8	9.48 ± 0.11	50.4	62.1 ± 1.0	170.2 ± 3.1	9.27 ± 0.02	51.3	-0.27/0.00[c]	5.5/0.0[d]
5′ UCAG7A^LCAGU 3′ AGUC C GUCA	7A^L-C	71.8 ± 2.8	203.5 ± 8.9	8.73 ± 0.12	46.8	67.6 ± 2.1	190.1 ± 6.8	8.61 ± 0.05	46.8	0.39/0.66[c]	1.0/-4.5[d]
5′ UCAGACAGU 3′ AGUCAGUCA	A-A	63.4 ± 8.6	178.9 ± 20.2	7.86 ± 0.49	43.6	57.4 ± 7.5	159.8 ± 8.3	7.84 ± 0.78	44.2	0.00	0.0
5′ UCAG7ACAGU 3′AGUC AGUCA	7A-A	72.3 ± 4.1	205.9 ± 12.7	8.49 ± 0.19	45.6	79.1 ± 7.1	227.9 ± 22.6	8.53 ± 0.14	45.0	-0.69	0.8
5′ UCAGA^LCAGU 3′ AGUCA GUCA	A^L-A	64.2 ± 3.8	180.5 ± 11.7	8.23 ± 0.18	45.4	58.5 ± 4.1	162.5 ± 13.0	8.13 ± 0.10	45.7	-0.29/0.00[c]	1.5/0.0[d]
5′ UCAG7A^LCAGU 3′ AGUC A GUCA	7A^L-A	73.2 ± 2.0	204.6 ± 5.9	9.80 ± 0.17	51.5	67.6 ± 2.1	187.0 ± 6.7	9.60 ± 0.07	51.8	-1.76/-1.47[c]	7.6/6.1[d]
5′ UCAGACAGU 3′ AGUCGGUCA	A-G	64.9 ± 5.9	186.4 ± 9.7	7.15 ± 0.08	40.0	69.0 ± 4.6	199.5 ± 9.9	7.10 ± 0.50	39.6	0.00	0.0
5′ UCAG7ACAGU 3′AGUC G GUCA	7A-G	68.8 ± 18.6	196.6 ± 58.7	7.86 ± 0.45	43.1	62.7 ± 7.8	177.2 ± 25.0	7.70 ± 0.26	42.9	-0.60	-3.3

5′ UCAG**A^{L}**CAGU 3′ AGUC **G** GUCA	A^{L}-G	69.1 ± 2.8	192.1 ± 9.0	9.49 ± 0.02	50.9	61.1 ± 1.0	167.0 ± 3.2	9.32 ± 0.02	51.8	-2.22/0.00[c]	12.2/0.0[d]
5′ UCAG**7A^{L}**CAGU 3′AGUC **G** GUCA	7A^{L}-G	74.8 ± 2.4	208.9 ± 7.7	9.98 ± 0.08	52.0	71.8 ± 2.6	199.8 ± 8.3	9.89 ± 0.08	52.2	-2.79/-0.57[c]	12.6/0.4[d]
5′ UCAGGCAGU 3′ AGUCCGUCA	M1	83.6 ± 4.8	227.6 ± 14.5	13.01 ± 0.41	63.1	72.8 ± 7.2	195.3 ± 21.5	12.28 ± 0.55	63.6	0.0	0.0
5′ UCAGGCAGU**A** 3′ AGUCCGUCA	3-A	85.5 ± 4.2	233.6 ± 12.8	13.11 ± 0.22	62.8	75.7 ± 5.1	203.7 ± 15.3	12.48 ± 0.34	63.5	-0.20/0.0[e]	-0.1/0[f]
5′**A**UCAGGCAGU 3′ AGUCCGUCA	5-A	83.6 ±6.9	227.1 ± 20.5	12.93 ± 0.56	62.8	81.2 ± 6.3	220.5 ± 19.0	12.76 ± 0.42	62.8	-0.48/-0.28[e]	-0.8/-0.7[f]
5′ UCAGGCAGU**7A** 3′ AGUCCGUCA	3-7A	99.9 ± 2.0	275.8 ± 6.3	14.33± 0.43	63.3	63.8 ± 7.2	167.6 ± 21.7	11.86± 0.56	65.2	0.42/0.0[e]	1.6/0 [f]
5′**7A**UCAGGCAGU 3′ AGUCCGUCA	5-7A	87.7 ± 8.8	239.8 ± 26.7	13.32 ± 0.53	63.0	91.7 ± 6.3	252.2 ± 18.9	13.52± 0.43	62.6	-1.24/-1.66[e]	-1.0/-2.6[f]
5′ UCAGGCAGU**A^{L}** 3′ AGUCCGUCA	3-A^{L}	102.3 ± 1.8	280.0 ± 6.3	15.47 ± 0.46	66.7	68.9± 7.7	180.9 ± 22.8	12.77± 0.70	67.9	-0.49/0.0[e]	4.3/0[f]
5′**A^{L}**UCAGGCAGU 3′ AGUCCGUCA	5-A^{L}	82.8 ± 1.6	225.5 ± 4.8	12.84 ± 0.16	62.6	85.9 ± 3.3	234.9± 9.9	13.05± 0.21	62.5	-0.77/-0.28[e]	-1.1/-5.4[f]
5′UCAGGCAGU**7A^{L}** 3′AGUCCGUCA	3-7A^{L}	91.0 ± 3.6	248.6 ± 11.5	13.86 ± 0.21	64.2	65.6 ± 2.8	173.0 ± 8.3	11.93± 0.22	64.8	0.35/0.0[e]	1.2/0[f]
5′**7A^{L}**UCAGGCAGU 3′ AGUCCGUCA	5-7A^{L}	82.1 ± 3.9	223.6± 11.3	12.73 ± 0.39	62.3	93.2 ± 2.4	256.9 ± 7.4	13.50± 0.16	62.1	-1.22/-1.57[e]	-1.5/-2.7[f]

[a] Solution: 1 M sodium chloride, 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7. [b] Calculated for 10^{-4} M oligomer concentration; where: 7A is N7 regioisomer of A and 7A^{L} is N7 regioisomer of LNA-A and A^{L} is LNA-A. [c] Comparison of free energies (ΔG°_{37}) to A-X and A^{L}-X, respectively (X=A, C, G, U). [d] Comparison of melting temperatures to A-X and A^{L}-X, respectively (X=A, C, G, U). [e] Comparison of free energies (ΔG°_{37}) to M1 and 3-X, respectively (X=A, 7A, A^{L}, 7A^{L}). [f] Comparison of melting temperatures to M1 and 3-X, respectively (X=A, 7A, A^{L}, 7A^{L}).

Table 2. Comparison of experimental results with computational predictions (for details see **Tables 1** and **S14-S25**).

Short Notation	Duplexes (5′-3′)	Experimental		Computational			
		$\Delta\Delta G^{\circ}_{37}$[a]	Loop Orientation[b]	ΔG_{pred}[c]	$\Delta\Delta G_{pred}$[d]	Loop Orientation	Hydrogen Bonds[e]
A-U	5′ UCAGACAGU 3′ AGUCUGUCA	0.00	*anti-anti*	1.86	0.00	*anti-anti*	1.98
7A-U[f]	5′ UCAG7ACAGU 3′AGUC U GUCA	3.92	-	5.43	3.57	*anti-anti*	1.16
$7A^{L}$-U	5′ UCAG7A^{L}CAGU 3′ AGUC U GUCA	1.97	-	2.85	2.58	*syn-anti*	1.00
A-C	5′ UCAGACAGU 3′ AGUCCGUCA	0.00	-	4.34	0.00	*anti-syn*	1.00
7A-C[g]	5′ UCAG7ACAGU 3′AGUC C GUCA	2.07	-	5.51	1.17	*syn-anti*	0.97
$7A^{L}$-C[h]	5′ UCAG7A^{L}CAGU 3′ AGUC C GUCA	0.39	-	4.76	0.42	*anti-anti*	0.99
A-A	5′ UCAGACAGU 3′ AGUCAGUCA	0.00	*anti-anti*[i]	3.48	0.00	*anti-anti*	0.96
7A-A	5′ UCAG7ACAGU 3′AGUC AGUCA	-0.69	*anti-anti*	1.50	-1.98	*anti-anti*	1.99
$7A^{L}$-A	5′ UCAG7A^{L}CAGU 3′ AGUC A GUCA	-1.76	*anti-anti*	1.44	-2.04	*anti-anti*	1.99
A-G	5′ UCAGACAGU 3′ AGUCGGUCA	0.00	-	2.81	0.00	*syn-anti*	1.99
7A-G	5′ UCAG7ACAGU 3′AGUC G GUCA	-0.60	*syn-anti*	2.10	-0.71	*syn-anti*	1.54
$7A^{L}$-G	5′ UCAG7A^{L}CAGU 3′AGUC G GUCA	-2.79	*syn-anti*	1.22	-1.59	*syn-anti*	1.75

[a] Experimental $\Delta\Delta G^{\circ}_{37}$ results in kcal/mol measured from T_M^{-1} vs log C_T plots (**Table 1**). [b] Loop orientations observed in NMR. [c] $\Delta G_{pred} = \Delta G_{MM/3D\text{-}RISM} - \Delta G_{NMODE}$ (in kcal/mol) (**Tables S4-S25**). [d] $\Delta\Delta G_{pred}$ is the predicted free energy difference with respect to A-X, where X=U, C, A, G. For example, $\Delta\Delta G_{pred,7A\text{-}U} = \Delta G_{pred,–A\text{-}U} - \Delta G_{pred,A\text{-}U}$. [e] Total number of hydrogen bonds observed in the middle base pair. [f] Cluster analyses display that this system has both *anti-anti* and *syn-anti* orientations with almost similar predicted ΔG_{pred} and hydrogen bonds

(**Table S15**). [g] Cluster analyses display that this system has both *syn-anti* and *anti-anti* orientations with relatively similar predicted ΔG_{pred} and hydrogen bonds (**Table S18**). [h] Cluster analyses display that this system has *anti-anti*, *syn-syn*, *syn-anti*, and *anti-syn* orientations with relatively similar predicted ΔG_{pred} and hydrogen bonds (**Table S19**). [i] From literature. (16,29,111).

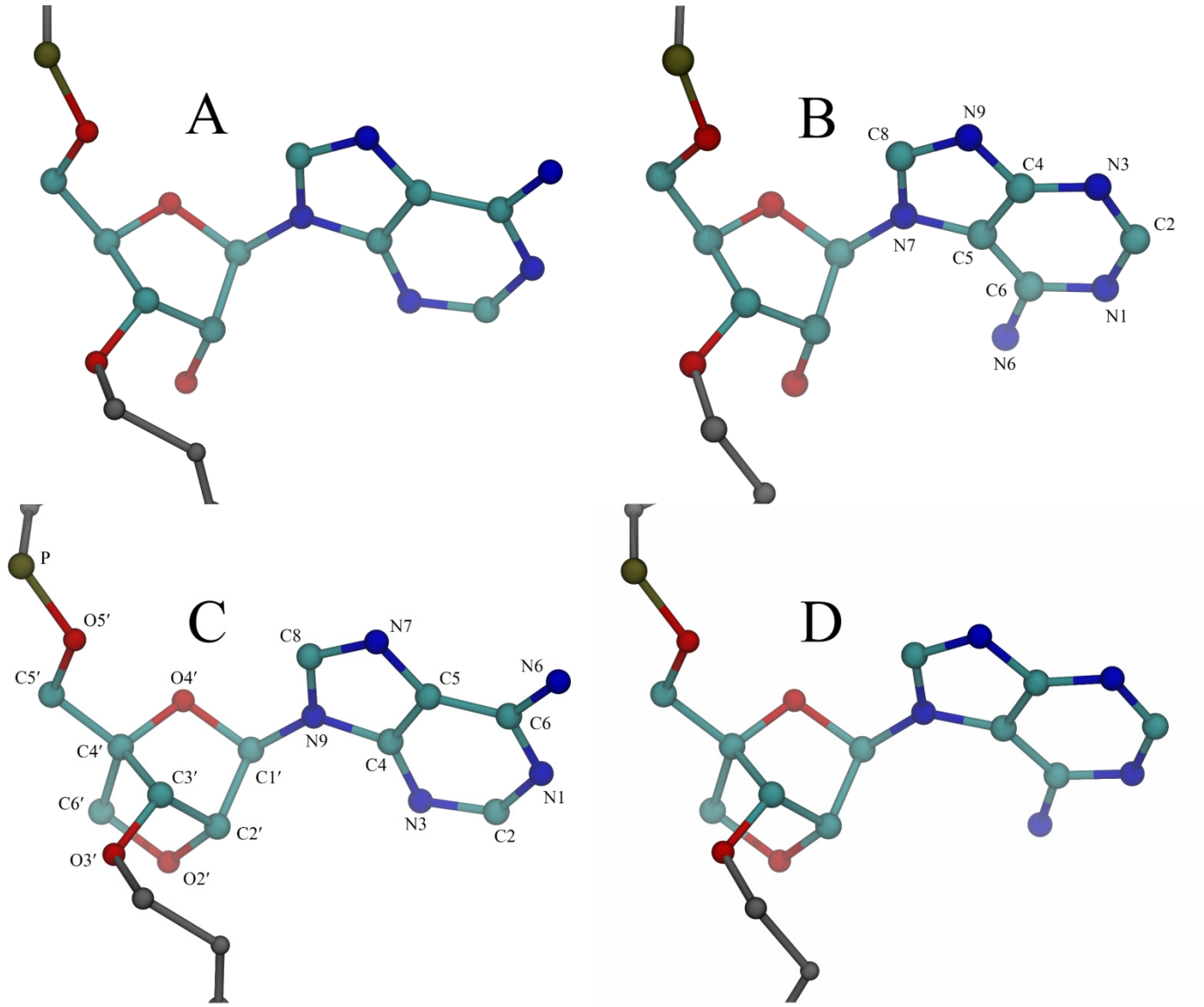

Figure 1. (A) N9-ribofuranosyladenine, A, (B) N7-ribofuranosyladenine, 7A, (C) N9-LNA-adenosine, A^{L}, and (D) N7-LNA-adenosine, $7A^{L}$. For simplicity, no hydrogen atoms are shown. Common atom names are displayed in B and C. Note that the O2′−C4′ methylene bridges seen in C and D lock the sugar pucker to C3′-endo.

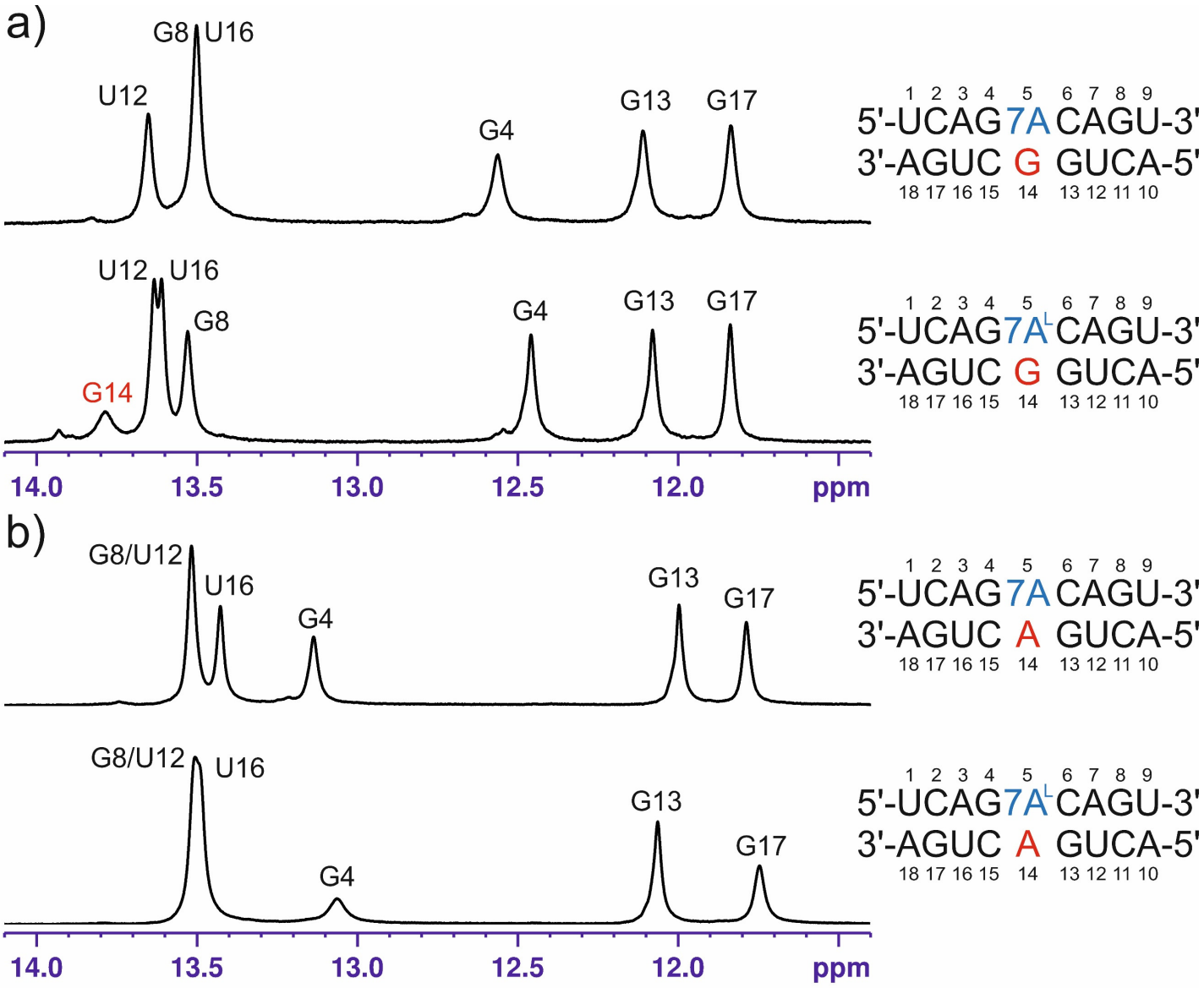

Figure 2. The imino proton regions of the 1D ^{1}H NMR spectra of duplexes containing (a) 1×1 7A:G and 1×1 7A^{L}:G mismatches recorded at 20 °C, and (b) 1×1 7A:A and 1×1 7A^{L}:A mismatches recorded at 25 °C.

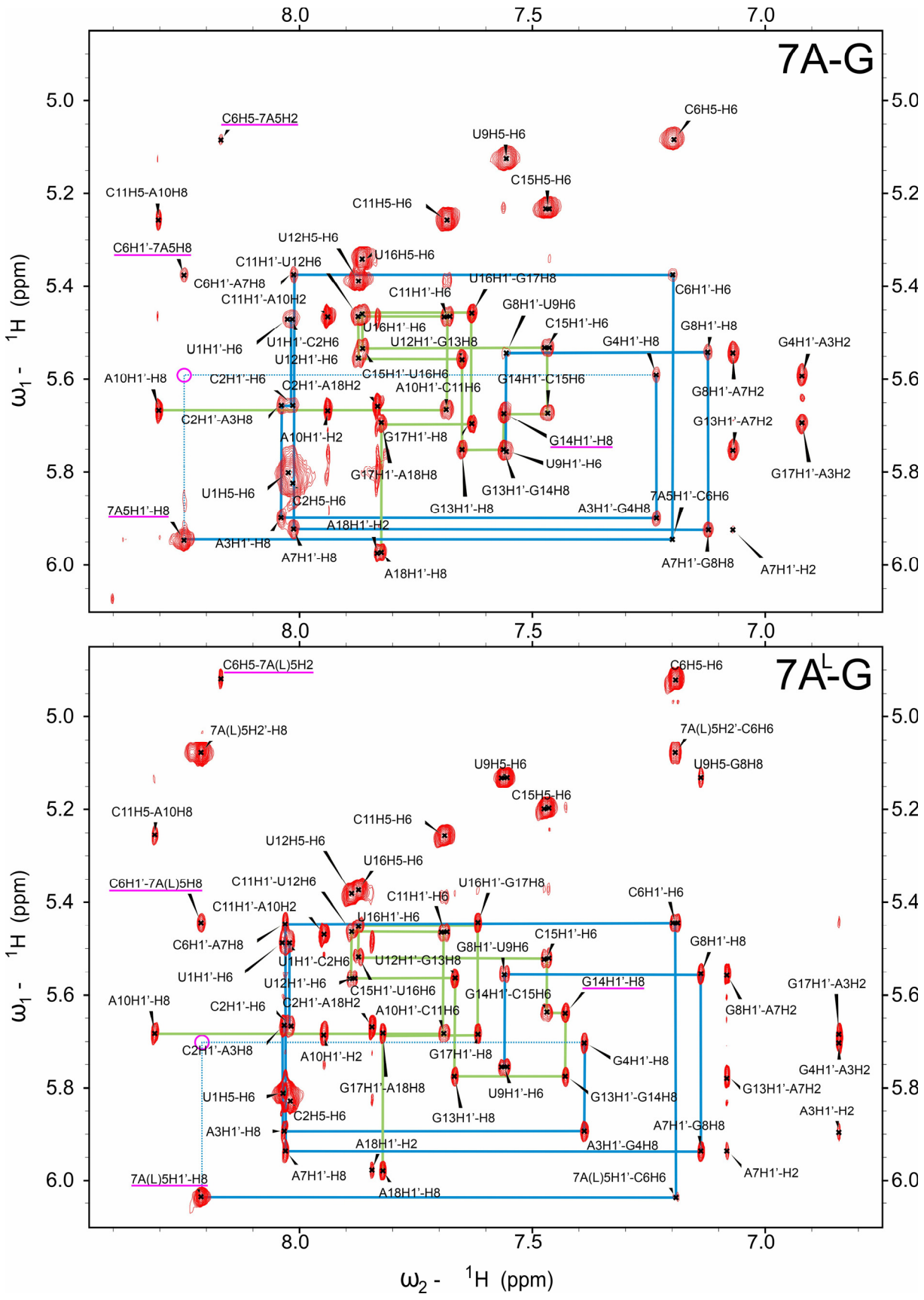

Figure 3. The aromatic-to-anomeric regions of the 2D ^{1}H-^{1}H NOESY spectra recorded for duplexes containing 1×1 7A:G and 1×1 7A^{L}:G mismatches. Significant NOEs discussed in the text are marked in magenta. Sequential connectivity paths for the two strands are displayed in green and blue lines. For residue numbering see **Figure 2**.

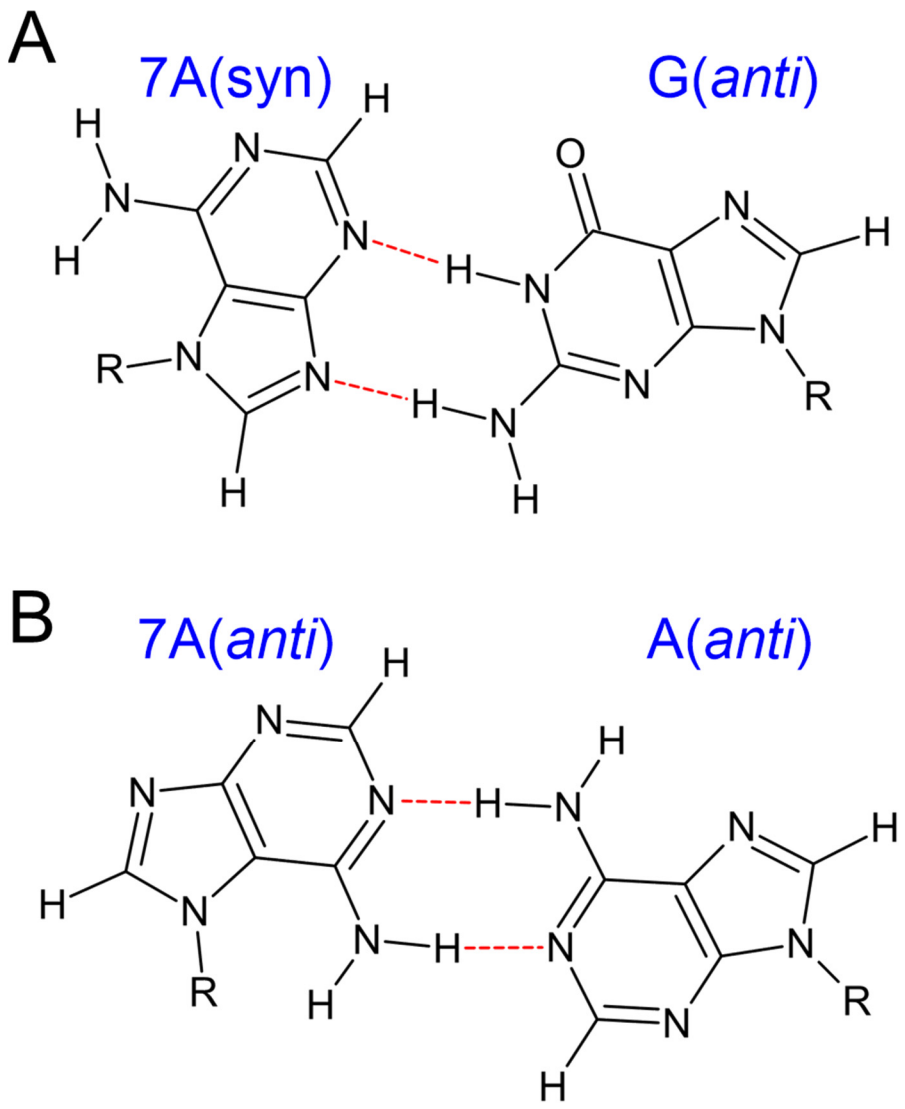

Figure 4. Proposed base pairing geometries for 1×1 7A:G and 7A^{L}:G (A), and 1×1 7A:A and 7A^{L}:A (B) mismatches using NMR data.

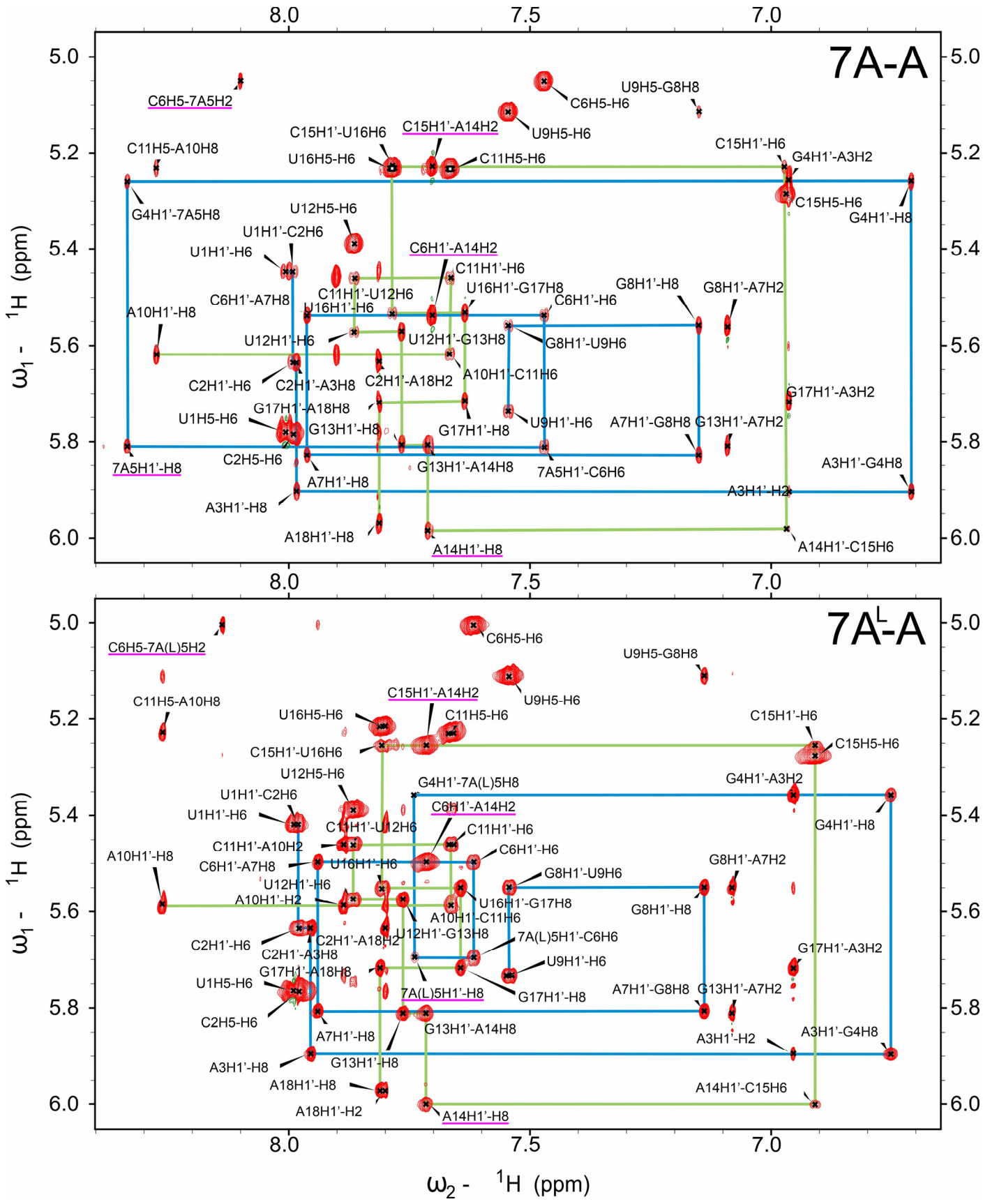

Figure 5. The aromatic-to-anomeric regions of the 2D ^{1}H-^{1}H NOESY spectra recorded for duplexes containing 1×1 7A:A and 1×1 7A^{L}:A mismatches. Significant NOEs discussed in the text are marked in magenta. Sequential connectivity paths for the two strands are displayed in green and blue lines. For residue numbering see **Figure 2**.

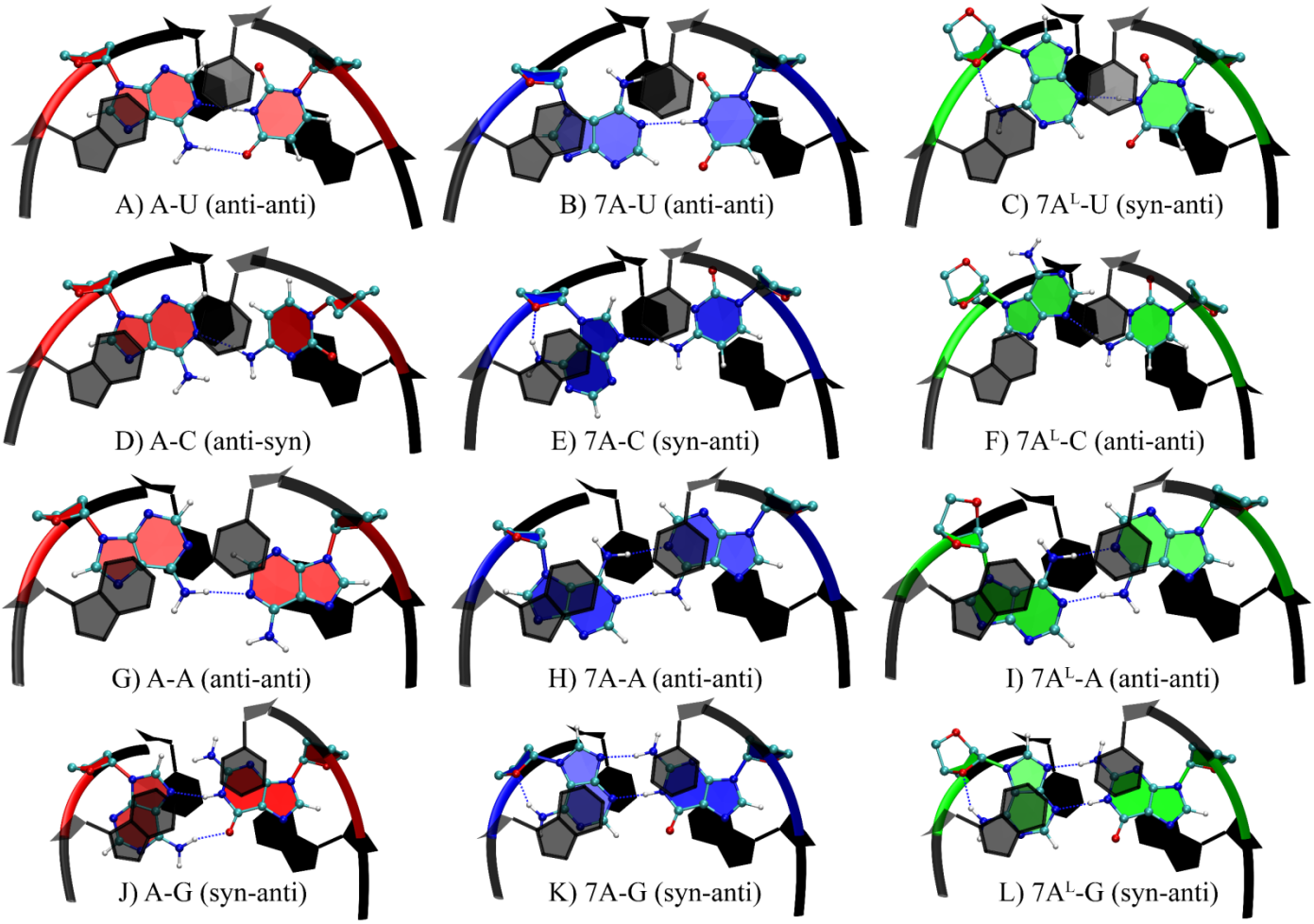

Figure 6. Predicted conformations of middle base pairs in (A) A-U, (B) 7A-U, (C) 7A^{L}-U, (D) A-CI, (E) 7A-C, (F) 7A^{L}-C, (G) A-A, (H) 7A-A, (I) 7A^{L}-A, (J) A-G, (K) 7A-G, and (G) 7A^{L}-G. Red, blue, and green are used to highlight base pairs formed by A, 7A, and 7A^{L}, respectively. Base pairs highlighted in black represent the closing GC base pairs with transparent ones illustrating the top GC base pairs. Dashed blue lines represent the hydrogen bonds. Note that due to inversion of adenine in 7A and 7A^{L}, functional groups are re-located, which either can destabilize the system as observed in B and C by losing a hydrogen bond or stabilize the system as observed in H, I, K, and L by forming extra hydrogen bonds (**Table 2**).

References:

1. Atkins, J.F., Gesteland, R.F. and Cech, T.R. (2010) *RNA Worlds: From Life's Origins to Diversity in Gene Regulation* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849-857.
3. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, **31**, 147-157.
4. Nissen, P., Hansen, J., Ban, N., Moore, P.B. and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289**, 920-930.
5. Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The C-ELEGANS heterochronic gene LIN-4 encodes small RNAs with antisense complementarity to LIN-14. *Cell*, **75**, 843-854.
6. Ruvkun, G. (2001) Molecular Biology: Glimpses of a Tiny RNA World. *Science*, **294**, 797-799.
7. Lee, J.F., Stovall, G.M. and Ellington, A.D. (2006) Aptamer therapeutics advance. *Curr. Opin. Chem. Biol.*, **10**, 282-289.
8. Ng, E.W.M., Shima, D.T., Calias, P., Cunningham, E.T., Guyer, D.R. and Adamis, A.P. (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat. Rev. Drug Discovery*, **5**, 123-132.
9. Willner, I. and Zayats, M. (2007) Electronic aptamer-based sensors. *Angew. Chem. Int. Ed.*, **46**, 6408-6418.
10. Chu, T.C., Twu, K.Y., Ellington, A.D. and Levy, M. (2006) Aptamer mediated siRNA delivery. *Nucleic Acids Res.*, **34**, e73-e73.
11. Duconge, F. and Toulme, J.J. (1999) In vitro selection identifies key determinants for loop-loop interactions: RNA aptamers selective for the TAR RNA element of HIV-1. *RNA*, **5**, 1605-1614.
12. Jeong, S., Han, S.R., Lee, Y.J. and Lee, S.W. (2010) Selection of RNA aptamers specific to active prostate-specific antigen. *Biotechnol. Lett.*, **32**, 379-385.
13. Epshtein, V., Mironov, A.S. and Nudler, E. (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 5052-5056.

14. Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L. and Breaker, R.R. (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science*, **306**, 275-279.

15. Yildirim, I., Chakraborty, D., Disney, M.D., Wales, D.J. and Schatz, G.C. (2015) Computational Investigation of RNA CUG Repeats Responsible for Myotonic Dystrophy 1. *J. Chem. Theory Comput.*, **11**, 4943-4958.

16. Chen, J.L., VanEtten, D.M., Fountain, M.A., Yildirim, I. and Disney, M.D. (2017) Structure and Dynamics of RNA Repeat Expansions That Cause Huntington's Disease and Myotonic Dystrophy Type 1. *Biochemistry*, **56**, 3463-3474.

17. Taghavi, A. and Yildirim, I. (2022) Computational Investigation of Bending Properties of RNA AUUCU, CCUG, CAG, and CUG Repeat Expansions Associated With Neuromuscular Disorders. *Front. Mol. Biosci.*, **9**, 830161.

18. Ashley, C.T., Jr. and Warren, S.T. (1995) Trinucleotide repeat expansion and human disease. *Annu. Rev. Genet.*, **29**, 703-728.

19. Emery, A.E.H. (2002) The muscular dystrophies. *Lancet*, **359**, 687-695.

20. Childs-Disney, J.L., Yildirim, I., Park, H., Lohman, J.R., Guan, L., Tran, T., Sarkar, P., Schatz, G.C. and Disney, M.D. (2014) Structure of the Myotonic Dystrophy Type 2 RNA and Designed Small Molecules That Reduce Toxicity. *ACS Chem. Biol.*, **9**, 538-550.

21. Raca, G., Siyanova, E.Y., McMurray, C.T. and Mirkin, S.M. (2000) Expansion of the $(CTG)_n$ repeat in the 5'-UTR of a reporter gene impedes translation. *Nucleic Acids Res.*, **28**, 3943-3949.

22. Kino, Y., Mori, D., Oma, Y., Takeshita, Y., Sasagawa, N. and Ishiura, S. (2004) Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.*, **13**, 495-507.

23. Sellier, C., Rau, F., Liu, Y.L., Tassone, F., Hukema, R.K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D.J. *et al.* (2010) Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J.*, **29**, 1248-1261.

24. Greco, C.M., Berman, R.F., Martin, R.M., Tassone, F., Schwartz, P.H., Chang, A., Trapp, B.D., Iwahashi, C., Brunberg, J., Grigsby, J. *et al.* (2006) Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain*, **129**, 243-255.

25. Jin, P., Duan, R.H., Qurashi, A., Qin, Y.L., Tian, D.H., Rosser, T.C., Liu, H.J., Feng, Y. and Warren, S.T. (2007) Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. *Neuron*, **55**, 556-564.

26. Li, L.B. and Bonini, N.M. (2010) Roles of trinucleotide-repeat RNA in neurological disease and degeneration. *Trends Neurosci.*, **33**, 292-298.

27. Li, L.B., Yu, Z.M., Teng, X.Y. and Bonini, N.M. (2008) RNA toxicity is a component of ataxin-3 degeneration in Drosophila. *Nature*, **453**, 1107-U1109.

28. Mykowska, A., Sobczak, K., Wojciechowska, M., Kozlowski, P. and Krzyzosiak, W.J. (2011) CAG repeats mimic CUG repeats in the misregulation of alternative splicing. *Nucleic Acids Res.*, **39**, 8938-8951.

29. Yildirim, I., Park, H., Disney, M.D. and Schatz, G.C. (2013) A dynamic structural model of expanded RNA CAG repeats: A refined X-ray structure and computational investigations using Molecular Dynamics and Umbrella Sampling simulations. *J. Am. Chem. Soc.*, **135**, 3528-3538.

30. Bush, J.A., Aikawa, H., Fuerst, R., Li, Y., Ursu, A., Meyer, S.M., Benhamou, R.I., Chen, J.L., Khan, T., Wagner-Griffin, S. *et al.* (2021) Ribonuclease recruitment using a small molecule reduced c9ALS/FTD r(G_4C_2) repeat expansion in vitro and in vivo ALS models. *Sci. Transl. Med.*, **13**, eabd5991.

31. Ursu, A., Baisden, J.T., Bush, J.A., Taghavi, A., Choudhary, S., Zhang, Y.-J., Gendron, T.F., Petrucelli, L., Yildirim, I. and Disney, M.D. (2021) A Small Molecule Exploits Hidden Structural Features within the RNA Repeat Expansion That Causes c9ALS/FTD and Rescues Pathological Hallmarks. *ACS Chem. Neurosci.*, **12**, 4076-4089.

32. Ursu, A., Wang, K.W., Bush, J.A., Choudhary, S., Chen, J.L., Baisden, J.T., Zhang, Y.J., Gendron, T.F., Petrucelli, L., Yildirim, I. *et al.* (2020) Structural Features of Small Molecules Targeting the RNA Repeat Expansion That Causes Genetically Defined ALS/FTD. *ACS Chem. Biol.*, **15**, 3112-3123.

33. Wang, Z.-F., Ursu, A., Childs-Disney, J.L., Guertler, R., Yang, W.-Y., Bernat, V., Rzuczek, S.G., Fuerst, R., Zhang, Y.-J., Gendron, T.F. *et al.* (2019) The Hairpin Form of r(G_4C_2)^{exp} in c9ALS/FTD Is Repeat-Associated Non-ATG Translated and a Target for Bioactive Small Molecules. *Cell Chem. Biol.*, **26**, 179-190.

34. Taghavi, A., Baisden, J.T., Childs-Disney, J.L., Yildirim, I. and Disney, Matthew D. (2023) Conformational dynamics of RNA G4C2 and G2C4 repeat expansions causing ALS/FTD using NMR and molecular dynamics studies. *Nucleic Acids Res.*, **51**, 5325-5340.

35. Kurreck, J., Wyszko, E., Gillen, C. and Erdmann, V.A. (2002) Design of antisense oligonucleotides stabilized by locked nucleic acids. *Nucleic Acids Res.*, **30**, 1911-1918.

36. Papargyri, N., Pontoppidan, M., Andersen, M.R., Koch, T. and Hagedorn, P.H. (2020) Chemical Diversity of Locked Nucleic Acid-Modified Antisense Oligonucleotides Allows Optimization of Pharmaceutical Properties. *Mol. Ther. Nucleic Acids*, **19**, 706-717.

37. Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hökfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K. *et al.* (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 5633-5638.

38. Wojtkowiak-Szlachcic, A., Taylor, K., Stepniak-Konieczna, E., Sznajder, L.J., Mykowska, A., Sroka, J., Thornton, C.A. and Sobczak, K. (2015) Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res.*, **43**, 3318-3331.

39. Elmén, J., Thonberg, H., Ljungberg, K., Frieden, M., Westergaard, M., Xu, Y., Wahren, B., Liang, Z., Ørum, H., Koch, T. *et al.* (2005) Locked nucleic acid (LNA) mediated improvements in siRNA stability and functionality. *Nucleic Acids Res.*, **33**, 439-447.

40. Mook, O.R., Baas, F., de Wissel, M.B. and Fluiter, K. (2007) Evaluation of locked nucleic acid-modified small interfering RNA in vitro and in vivo. *Mol. Cancer Ther.*, **6**, 833-843.

41. Jepsen, J.S., Sorensen, M.D. and Wengel, J. (2004) Locked nucleic acid: A potent nucleic acid analog in therapeutics and biotechnology. *Oligonucleotides*, **14**, 130-146.

42. Jepsen, J.S. and Wengel, J. (2004) LNA-antisense rivals siRNA for gene silencing. *Curr. Opin. Drug Discovery Dev.*, **7**, 188-194.

43. Koshkin, A.A., Singh, S.K., Nielsen, P., Rajwanshi, V.K., Kumar, R., Meldgaard, M., Olsen, C.E. and Wengel, J. (1998) LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron*, **54**, 3607-3630.

44. Childs, J.L., Disney, M.D. and Turner, D.H. (2002) Oligonucleotide directed misfolding of RNA inhibits *Candida albicans* group I intron splicing. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 11091-11096.

45. Yildirim, I., Kierzek, E., Kierzek, R. and Schatz, G.C. (2014) Interplay of LNA and 2'-0-Methyl RNA in the Structure and Thermodynamics of RNA Hybrid Systems: A Molecular Dynamics Study Using the Revised AMBER Force Field and Comparison with Experimental Results. *J. Phys. Chem. B*, **118**, 14177-14187.

46. Nielsen, K.E. and Spielmann, H.P. (2005) The structure of a mixed LNA/DNA : RNA duplex is driven by conformational coupling between LNA and deoxyribose residues as determined from C-13 relaxation measurements. *J. Am. Chem. Soc.*, **127**, 15273-15282.

47. Pande, V. and Nilsson, L. (2008) Insights into structure, dynamics and hydration of locked nucleic acid (LNA) strand-based duplexes from molecular dynamics simulations. *Nucleic Acids Res.*, **36**, 1508-1516.

48. Ivanova, A. and Rosch, N. (2007) The structure of LNA:DNA hybrids from molecular dynamics simulations: The effect of locked nucleotides. *J. Phys. Chem. A*, **111**, 9307-9319.

49. Hunziker, J., Priestley, E.S., Brunar, H. and Dervan, P.B. (1995) Design of an N7-Glycosylated Purine Nucleoside for Recognition of GC Base Pairs by Triple Helix Formation. *J. Am. Chem. Soc.*, **117**, 2661-2662.

50. Radhakrishnan, I., Patel, D.J., Priestley, E.S., Nash, H.M. and Dervan, P.B. (1993) NMR structural studies on a nonnatural deoxyribonucleoside which mediates recognition of GC base pairs in pyrimidine·purine·pyrimidine DNA triplexes. *Biochemistry*, **32**, 11228-11234.

51. Koshlap, K.M., Schultze, P., Brunar, H., Dervan, P.B. and Feigon, J. (1997) Solution Structure of an Intramolecular DNA Triplex Containing an N7-Glycosylated Guanine Which Mimics a Protonated Cytosine. *Biochemistry*, **36**, 2659-2668.

52. Rao, T.S., Durland, R.H. and Revankar, G.R. (1994) Synthesis of oligonucleotides containing 7-(2-deoxy-β-d-erythro-pentofuranosyl)guanine and 8-amino-2′-deoxyguanosine. *J. Heterocycl. Chem.*, **31**, 935-940.

53. St. Clair, A., Xiang, G. and McLaughlin, L.W. (1998) Synthesis and Triplex Forming Properties of an Acyclic N7-Glycosylated Guanine Nucleoside. *Nucleosides Nucleotides Nucleic Acids*, **17**, 925-937.

54. Leonard, P., Wiglenda, T. and Seela, F. (2001) Synthesis and base pairing properties of 9-deazapurine N7-nucleosides in oligonucleotide duplexes and triplexes. *Nucleosides Nucleotides Nucleic Acids*, **20**, 1279-1282.

55. Seela, F. and Leonard, P. (1997) Studies on the Base-Pairing Properties of N7-(2-Deoxy-β-D-erythro-pentofuranosyl)guanine (N7Gd). *Helv. Chim. Acta*, **80**, 1301-1318.

56. Seela, F. and Zulauf, M. (1999) Oligonucleotides Containing 7-Deazaadenines: The Influence of the 7-Substituent Chain Length and Charge on the Duplex Stability. *Helv. Chim. Acta*, **82**, 1878-1898.

57. Zacharias, M. and Hagerman, P.J. (1996) The influence of symmetric internal loops on the flexibility of RNA. *J. Mol. Biol.*, **257**, 276-289.

58. Dethoff, E.A., Chugh, J., Mustoe, A.M. and Al-Hashimi, H.M. (2012) Functional complexity and regulation through RNA dynamics. *Nature*, **482**, 322-330.

59. Kierzek, R., Burkard, M.E. and Turner, D.H. (1999) Thermodynamics of single mismatches in RNA duplexes. *Biochemistry*, **38**, 14214-14223.

60. Sugimoto, N., Nakano, M. and Nakano, S.-i. (2000) Thermodynamics−Structure Relationship of Single Mismatches in RNA/DNA Duplexes. *Biochemistry*, **39**, 11270-11281.

61. Davis, A.R. and Znosko, B.M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry*, **46**, 13425-13436.

62. SantaLucia, J., Kierzek, R. and Turner, D.H. (1990) Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. *Biochemistry*, **29**, 8813-8819.

63. Davis, A.R. and Znosko, B.M. (2008) Thermodynamic Characterization of Naturally Occurring RNA Single Mismatches with G-U Nearest Neighbors. *Biochemistry*, **47**, 10178-10187.

64. Malina, J., Scott, P. and Brabec, V. (2015) Recognition of DNA/RNA bulges by antimicrobial and antitumor metallohelices. *Dalton Trans.*, **44**, 14656-14665.

65. Tok, J.B.H., Bi, L. and Saenz, M. (2005) Specific recognition of napthyridine-based ligands toward guanine-containing bulges in RNA duplexes and RNA–DNA heteroduplexes. *Bioorg. Med. Chem. Lett.*, **15**, 827-831.

66. Disney, M.D., Labuda, L.P., Paul, D.J., Poplawski, S.G., Pushechnikov, A., Tran, T., Velagapudi, S.P., Wu, M. and Childs-Disney, J.L. (2008) Two-dimensional combinatorial screening identifies specific aminoglycoside - RNA internal loop partners. *J. Am. Chem. Soc.*, **130**, 11185-11194.

67. Guan, L.R. and Disney, M.D. (2012) Recent Advances in Developing Small Molecules Targeting RNA. *ACS Chem. Biol.*, **7**, 73-86.

68. Nakatani, K., Horie, S., Goto, Y., Kobori, A. and Hagihara, S. (2006) Evaluation of mismatch-binding ligands as inhibitors for Rev–RRE interaction. *Biorg. Med. Chem.*, **14**, 5384-5388.

69. Framski, G., Gdaniec, Z., Gdaniec, M. and Boryski, J. (2006) A reinvestigated mechanism of ribosylation of adenine under silylating conditions. *Tetrahedron*, **62**, 10123-10129.

70. Beaucage, S.L. and Caruthers, M.H. (1981) Deoxynucleoside phosphoramidites - A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.*, **22**, 1859-1862.

71. McBride, L.J. and Caruthers, M.H. (1983) An investigation of several deoxynucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides. *Tetrahedron Lett.*, **24**, 245-248.

72. Koshkin, A.A., Fensholdt, J., Pfundheller, H.M. and Lomholt, C. (2001) A Simplified and Efficient Route to 2‘-O, 4‘-C-Methylene-Linked Bicyclic Ribonucleosides (Locked Nucleic Acid). *J. Org. Chem.*, **66**, 8504-8512.

73. Kierzek, E., Ciesielska, A., Pasternak, K., Mathews, D.H., Turner, D.H. and Kierzek, R. (2005) The influence of locked nucleic acid residues on the thermodynamic properties of 2 '-O-methyl RNA/RNA heteroduplexes. *Nucleic Acids Res.*, **33**, 5082-5093.

74. Kierzek, E. and Kierzek, R. (2003) The synthesis of oligoribonucleotides containing N-6-alkyladenosines and 2-methylthio-N-6-alkyladenosines via post-synthetic modification of precursor oligomers. *Nucleic Acids Res.*, **31**, 4461-4471.

75. Xia, T., SantaLucia, J., Jr., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C. and Turner, D.H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry*, **37**, 14719-14735.

76. Richards, E.G. (1975) In Fasman, G. D. (ed.), *CRC Handbook of Biochemistry and Molecular Biology: Nucleic Acids*. 3rd ed. CRC Press, Cleveland, OH, Vol. 1, pp. 596-603.

77. Borer, P.N. (1975) In Fasman, G. D. (ed.), *CRC Handbook of Biochemistry and Molecular Biology: Nucleic Acids*. 3rd ed. CRC Press, Cleveland, OH, Vol. 1, pp. 589-595.

78. McDowell, J.A. and Turner, D.H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of (rGAGGUCUC)$_2$ by two-dimensional NMR and simulated annealing. *Biochemistry*, **35**, 14077-14089.

79. Wijmenga, S.S. and van Buuren, B.N.M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.*, **32**, 287-387.

80. Lee, W., Tonelli, M. and Markley, J.L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*, **31**, 1325-1327.

81. Case, D.A., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, T.E., Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Ghoreishi, D., Gilson, M.K. *et al.* (2018). University of California, San Francisco, CA.

82. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H. *et al.* (2016). Revision A.02 ed. Gaussian, Inc., Wallingford, CT.

83. Cieplak, P., Cornell, W.D., Bayly, C. and Kollman, P.A. (1995) Application of the multimolecule and multiconformational RESP Methodology to biopolymers - Charge derivation for DNA, RNA, and Proteins. *J. Comput. Chem.*, **16**, 1357-1377.

84. Yildirim, I., Stern, H.A., Tubbs, J.D., Kennedy, S.D. and Turner, D.H. (2011) Benchmarking AMBER force fields for RNA: Comparisons to NMR spectra for single-stranded r(GACC) are improved by revised χ torsions. *J. Phys. Chem. B*, **115**, 9261–9270.

85. Wales, D.J. and Yildirim, I. (2017) Improving Computational Predictions of Single-Stranded RNA Tetramers with Revised α/γ Torsional Parameters for the Amber Force Field. *J. Phys. Chem. B*, **121**, 2989–2999.

86. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. (1996) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules (vol 117, pg 5179, 1995). *J. Am. Chem. Soc.*, **118**, 2309-2309.

87. Yildirim, I., Stern, H.A., Kennedy, S.D., Tubbs, J.D. and Turner, D.H. (2010) Reparameterization of RNA χ torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine. *J. Chem. Theory Comput.*, **6**, 1520-1531.

88. Joung, I.S. and Cheatham, T.E. (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B*, **112**, 9020-9041.

89. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926-935.

90. Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) Numerical-Integration of cartesian equations of motion of a system with constraints: Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.*, **23**, 327-341.

91. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G. (1995) A smooth Particle Mesh Ewald method. *J. Chem. Phys.*, **103**, 8577-8593.

92. Childs-Disney, J.L., Stepniak-Konieczna, E., Tran, T., Yildirim, I., Park, H., Chen, C.Z., Hoskins, J., Southall, N., Marugan, J.J., Patnaik, S. *et al.* (2013) Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules. *Nat. Commun.*, **4**, 2044.

93. Benhamou, R.I., Vezina-Dawod, S., Choudhary, S., Wang, K.W., Meyer, S.M., Yildirim, I. and Disney, M.D. (2020) Macrocyclization of a Ligand Targeting a Toxic RNA Dramatically Improves Potency. *ChemBioChem*, **21**, 3229-3233.

94. Angelbello, A.J., Benhamou, R.I., Rzuczek, S.G., Choudhary, S., Tang, Z., Chen, J.L., Roy, M., Wang, K.W., Yildirim, I., Jun, A.S. *et al.* (2021) A Small Molecule that Binds an RNA Repeat Expansion Stimulates Its Decay via the Exosome Complex. *Cell Chem. Biol.*, **28**, 34-45.

95. Vezina-Dawod, S., Angelbello, A.J., Choudhary, S., Wang, K.W., Yildirim, I. and Disney, M.D. (2021) Massively Parallel Optimization of the Linker Domain in Small Molecule Dimers Targeting a Toxic r(CUG) Repeat Expansion. *ACS Med. Chem. Lett.*, **12**, 907-914.

96. Costales, M.G., Aikawa, H., Li, Y., Childs-Disney, J.L., Abegg, D., Hoch, D.G., Velagapudi, S.P., Nakai, Y., Khan, T., Wang, K.W. *et al.* (2020) Small-molecule targeted recruitment of a nuclease to cleave an oncogenic RNA in a mouse model of metastatic cancer. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 2406-2411.

97. Suresh, B.M., Li, W.C., Zhang, P.Y., Wang, K.W., Yildirim, I., Parker, C.G. and Disney, M.D. (2020) A general fragment-based approach to identify and optimize bioactive ligands targeting RNA. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 33197-33203.

98. Wang, K.W., Riveros, I., DeLoye, J. and Yildirim, I. (2023) Dynamic docking of small molecules targeting RNA CUG repeats causing myotonic dystrophy type 1. *Biophys. J.*, **122**, 180-196.

99. Luchko, T., Gusarov, S., Roe, D.R., Simmerling, C., Case, D.A., Tuszynski, J. and Kovalenko, A. (2010) Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J. Chem. Theory Comput.*, **6**, 607-624.

100. Kovalenko, A. (2003) In Hirata, F. (ed.), *Molecular Theory of Solvation*. Springer Netherlands, Dordrecht, pp. 169-275.

101. Genheden, S., Luchko, T., Gusarov, S., Kovalenko, A. and Ryde, U. (2010) An MM/3D-RISM Approach for Ligand Binding Affinities. *J. Phys. Chem. B*, **114**, 8505-8516.

102. Kierzek, E., Pasternak, A., Pasternak, K., Gdaniec, Z., Yildirim, I., Turner, D.H. and Kierzek, R. (2009) Contributions of stacking, preorganization, and hydrogen bonding to the thermodynamic stability of duplexes between RNA and 2'-O-Methyl RNA with Locked Nucleic Acids. *Biochemistry*, **48**, 4377-4387.

103. Freier, S.M., Sinclair, A., Neilson, T. and Turner, D.H. (1985) Improved free energies for GC base pairs. *J. Mol. Biol.*, **185**, 645-647.

104. Freier, S.M., Burger, B.J., Alkema, D., Neilson, T. and Turner, D.H. (1983) Effects of 3' dangling end stacking on the stability of GGCC and CCGG double helices. *Biochemistry*, **22**, 6198-6206.

105. Burkard, M.E. and Turner, D.H. (2000) NMR structures of r(GCAGGCGUGC)$_2$ and determinants of stability for single guanosine-guanosine base pairs. *Biochemistry*, **39**, 11748-11762.

106. Adrian, M., Heddi, B. and Phan, A.T. (2012) NMR spectroscopy of G-quadruplexes. *Methods*, **57**, 11-24.

107. Varani, G., Aboulela, F. and Allain, F.H.T. (1996) NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc.*, **29**, 51-127.

108. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2011) Crystal structures of CGG RNA repeats with implications for fragile X-associated tremor ataxia syndrome. *Nucleic Acids Res.*, **39**, 7308-7315.

109. Kumar, A., Fang, P.F., Park, H., Guo, M., Nettles, K.W. and Disney, M.D. (2011) A crystal structure of a model of the repeating r(CGG) transcript found in Fragile X syndrome. *ChemBioChem*, **12**, 2140-2142.

110. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2012) Crystallographic characterization of CCG repeats. *Nucleic Acids Res.*, **40**, 8155-8162.

111. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2010) Atomic resolution structure of CAG RNA repeats: structural insights and implications for the trinucleotide repeat expansion diseases. *Nucleic Acids Res.*, **38**, 8370-8376.

112. An, Y., Chen, Z.S., Chan, Ho Yin E. and Ngo, Jacky Chi K. (2022) Molecular insights into the interaction of CAG trinucleotide RNA repeats with nucleolin and its implication in polyglutamine diseases. *Nucleic Acids Res.*, **50**, 7655-7668.

113. Tawani, A. and Kumar, A. (2015) Structural Insights Reveal the Dynamics of the Repeating r(CAG) Transcript Found in Huntington's Disease (HD) and Spinocerebellar Ataxias (SCAs). *PLoS One*, **10**, e0131788.

114. Lukavsky, P.J., Kim, I., Otto, G.A. and Puglisi, J.D. (2003) Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol.*, **10**, 1033-1038.

115. Keane, S.C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., Barton, S., Hosic, A., Florwick, A., Santos, J. *et al.* (2015) Structure of the HIV-1 RNA packaging signal. *Science*, **348**, 917-921.

116. Wengel, J., Petersen, M., Frieden, M. and Koch, T. (2003) Chemistry of locked nucleic acids (LNA): Design, synthesis, and bio-physical properties. *Lett. Pept. Sci.*, **10**, 237-253.

117. Turner, D.H., Sugimoto, N., Kierzek, R. and Dreiker, S.D. (1987) Free-energy increments for hydrogen-bonds in nucleic-acid base pairs. *J. Am. Chem. Soc.*, **109**, 3783-3785.

118. Kierzek, E., Mathews, D.H., Ciesielska, A., Turner, D.H. and Kierzek, R. (2006) Nearest neighbor parameters for Watson–Crick complementary heteroduplexes formed between 2′-O-methyl RNA and RNA oligonucleotides. *Nucleic Acids Res.*, **34**, 3609-3614.

119. Turner, D.H. (2000) In Bloomfield, V. A., Crothers, D. M. and Tinoco, I., Jr. (eds.), *Nucleic Acids: Structures, Properties, and Functions*. University Science Books, Sausalito, California, pp. 259-334.

120. Bommarito, S., Peyret, N. and SantaLucia, J., Jr. (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.*, **28**, 1929-1934.

121. Ohmichi, T., Nakano, S.-i., Miyoshi, D. and Sugimoto, N. (2002) Long RNA Dangling End Has Large Energetic Contribution to Duplex Stability. *J. Am. Chem. Soc.*, **124**, 10367-10372.

122. Schaefer, M., Kapoor, U. and Jantsch, M.F. (2017) Understanding RNA modifications: the promises and technological bottlenecks of the 'epitranscriptome'. *Open Biology*, **7**.

123. Jonkhout, N., Tran, J., Smith, M.A., Schonrock, N., Mattick, J.S. and Novoa, E.M. (2017) The RNA modification landscape in human disease. *RNA*, **23**, 1754-1769.

124. Boo, S.H. and Kim, Y.K. (2020) The emerging role of RNA modifications in the regulation of mRNA stability. *Exp. Mol. Med.*, **52**, 400-408.

125. Shen, X.L. and Corey, D.R. (2018) Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs. *Nucleic Acids Res.*, **46**, 1584-1600.

126. Corey, D.R. (2007) Chemical modification: the key to clinical application of RNA interference? *J. Clin. Invest.*, **117**, 3615-3622.

127. Kumar, R.K. and Davis, D.R. (1997) Synthesis and Studies on the Effect of 2-Thiouridine and 4-Thiouridine on Sugar Conformation and RNA Duplex Stability. *Nucleic Acids Res.*, **25**, 1272-1280.