

Impact of 3-deazapurine nucleobases on RNA properties

Raphael Bereiter^{1,†}, Maximilian Himmelstoß^{1,†}, Eva Renard², Elisabeth Mairhofer¹, Michaela Egger¹, Kathrin Breuker¹, Christoph Kreutz¹, Eric Ennifar² and Ronald Micura^{1,*}

¹Institute of Organic Chemistry, Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria and ²Architecture et Réactivité de l'ARN - CNRS UPR 9002, Université de Strasbourg, Strasbourg, France

Received March 05, 2021; Editorial Decision March 26, 2021; Accepted March 29, 2021

ABSTRACT

Deazapurine nucleosides such as 3-deazaadenosine (c³A) are crucial for atomic mutagenesis studies of functional RNAs. They were the key for our current mechanistic understanding of ribosomal peptide bond formation and of phosphodiester cleavage in recently discovered small ribozymes, such as twister and pistol RNAs. Here, we present a comprehensive study on the impact of c³A and the thus far underinvestigated 3-deazaguanosine (c³G) on RNA properties. We found that these nucleosides can decrease thermodynamic stability of base pairing to a significant extent. The effects are much more pronounced for 3-deazapurine nucleosides compared to their constitutional isomers of 7-deazapurine nucleosides (c⁷G, c⁷A). We furthermore investigated base pair opening dynamics by solution NMR spectroscopy and revealed significantly enhanced imino proton exchange rates. Additionally, we solved the X-ray structure of a c³A-modified RNA and visualized the hydration pattern of the minor groove. Importantly, the characteristic water molecule that is hydrogen-bonded to the purine N3 atom and always observed in a natural double helix is lacking in the 3-deazapurine-modified counterpart. Both, the findings by NMR and X-ray crystallographic methods hence provide a rationale for the reduced pairing strength. Taken together, our comparative study is a first major step towards a comprehensive understanding of this important class of nucleoside modifications.

INTRODUCTION

Deazapurine nucleosides have been intensively applied in atomic mutagenesis studies of biologically relevant RNAs (1–5). The replacement of a single nitrogen atom by a carbon atom can be critical because the acid-base properties of the nucleobase are dramatically changed and hydrogen acceptor/donor properties are erased at the specific position (6–8). This can be determining for RNA base pairing (8,9), for RNA recognition of other nucleic acids (e.g. DNA, 2'-OCH₃ RNA) (6), proteins (3), small molecules (8), and ions (9), and this can also be crucial with respect to RNA-catalyzed reactions (4,10–13). Concerning the latter, atomic mutagenesis lead to our current in-depth understanding of the chemical mechanism of phosphodiester cleavage of the recently discovered twister and pistol ribozymes (14–16). Similarly, atomic mutagenesis of the peptidyl transferase center in the ribosome center allowed to critically evaluate the mechanistic proposals for peptide bond formation that arose from interpretations of static crystal structures (17–20). These functional assays contributed significantly to a profound comprehension of ribosomal protein synthesis.

Suitable deazanucleosides for informative RNA atomic mutagenesis experiments are 7-deazaadenosine (c⁷A) (4,11–13,15,16), 3-deazaadenosine (c³A) (14,20), 1-deazaadenosine (c¹A) (11,12,14,20), 7-deazaguanosine (c⁷G) (13,15) and 3-deazacytidine (c³C) (13,21). Moreover, 3-deazaguanosine (c³G) and 1-deazaguanosine (c¹G) would be highly useful for RNA atomic mutagenesis, however, such studies are very rare (22) because synthetic access to appropriate phosphoramidite building blocks and to the corresponding RNAs is challenging (22,23). In general, 3-deazapurine modified RNAs appear underinvestigated when compared to the 7-deazapurine modified counterparts (24–28). Besides, more is known about deaza-modified DNA compared to deaza-modified RNA (29–38).

*To whom correspondence should be addressed. Tel: +43 512 507 57710; Fax: +43 512 507 57799; Email: ronald.micura@uibk.ac.at

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Therefore, we set out to accomplish a thorough chemical and biophysical analysis of 3-deazapurine containing RNA, along with the improvement of the synthesis of c^3G containing RNA beforehand. We here describe the impact of c^3G and c^3A on RNA properties. Based on UV-spectroscopic melting experiments, a detailed thermodynamic analysis of duplex and hairpin stabilities is provided and the effects on base pairing are discussed in the light of the sequence context. Furthermore, solution NMR spectroscopy sheds light on base pair opening dynamics. In addition, we have solved the X-ray structure of a c^3A containing RNA at atomic resolution to disclose crucial structural features, such as ribose puckers, hydrogen-bonding networks, and hydration patterns of the deazanucleoside, and to correlate them to base pairing properties.

MATERIAL AND METHODS

Synthesis and characterization of organic compounds

Reagents were purchased in the highest available quality from commercial suppliers (Merck/Sigma-Aldrich, ABCR, VWR, ChemGenes, CarboSynth) and used without further purification. All reactions were carried out under argon atmosphere, unless otherwise noted. Analytical thin-layer chromatography (TLC) was performed on Macherey-Nagel Polygram[®] SIL G/UV₂₅₄ plates. Silica gel 60 (mesh size 0.04–0.063 mm) for column chromatography was purchased from Macherey-Nagel. The procedures for chemical synthesis of phosphoramidites **8** and **9**, and the 6-chloro-3-deazapurine **14** and their characterization data are available in the Supporting Information (phosphoramidite **8**: eight steps, eight chromatographic purifications, 13% overall yield; total amount synthesized: 0.90 g; phosphoramidite **9**: ten steps, 10 chromatographic purifications, 9% overall yield; total amount synthesized: 0.80 g; 6-amino-3-deazapurine **15**: 4 steps, four chromatographic purifications, 54% overall yield; total amount synthesized: 5 g). ¹H, ¹³C, and ³¹P NMR spectra were recorded on a Bruker UltraShield[™] 400 Plus spectrometer. Chemical shifts (δ) are reported relative to tetramethylsilane (TMS), referenced to the residual solvent signal (DMSO-*d*₆: 2.50 ppm for ¹H and 39.52 ppm for ¹³C NMR spectra; CDCl₃: 7.26 ppm for ¹H and 77.16 ppm for ¹³C NMR spectra). The following abbreviations were used to denote multiplicities: s = singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet, b = broad. Signal assignments are based on ¹H–¹H-COSY, ¹H–¹³C-HSQC and ¹H–¹³C-HMBC experiments. High resolution mass spectra were recorded in positive ion mode on a Thermo Scientific Q Exactive Orbitrap, ionized via electrospray at 3.7 kV spray voltage.

RNA solid-phase synthesis

Standard phosphoramidite chemistry was applied for RNA strand elongation and incorporation of 3-deazaguanosine (*N*²-Tfa, 2'-*O*-Tbs and *N*²-Tfa, 2'-*O*-CEM; >98% coupling yield), 3-deazadenosine (*N*⁶-Bz, 2'-*O*-TIPS; >92% coupling yield), 7-deazaguanosine (*N*²-*i*Bu, 2'-*O*-Tbs; >98% coupling yield) and 7-deazaadenosine (*N*⁶-Bz, 2'-*O*-Tbs; >98% coupling yield). 2'-*O*-TOM and acetyl protected

nucleoside phosphoramidite building blocks and 2'-*O*-Tbs 1000 Å CPG solid support (>15nt) were purchased from ChemGenes, Primer support[™] 5G (<15nt) was purchased from GE Healthcare. All oligonucleotides were synthesized on an ABI 391 or ABI 392 Nucleic Acid Synthesizer following standard methods: detritylation (90 sec) with dichloroacetic acid/1,2-dichloroethane (4/96); coupling (5.0 min) with phosphoramidites/acetonitrile (100 mM, 200 μ l) and benzylthiotetrazole / acetonitrile (300 mM, 500 μ l); capping (2 \times 25 s) with Cap A/Cap B (1/1) for c^3A , c^7G & c^7A modified RNA, Cap A: 4-(dimethylamino)pyridine/acetonitrile (500 mM), Cap B: acetic anhydride/*sym*-collidine/acetonitrile (2/3/5) and Cap A mild/Cap B mild (1/1) for c^3G modified RNA, Cap A mild: phenoxyacetic anhydride/acetonitrile (100 mM), Cap B mild: *N*-methylimidazole/*sym*-collidine/tetrahydrofuran (0.160/0.265/10); oxidation (60 sec) with iodine (20 mM) in tetrahydrofuran/pyridine/H₂O (35/10/5). Solutions of phosphoramidites, tetrazole and Cap were dried over activated molecular sieves (3 Å) overnight.

Deprotection, purification and quantification of natural and modified RNA

For basic deprotection of natural, c^3G , c^3A , c^7G and c^7A modified RNA, the solid support was mixed with aqueous methylamine (40%, 0.65 ml) and aqueous ammonia (28%, 0.65 ml) for 15 min at 65°C or 3 h at 37°C. The supernatant was removed and the solid support was washed twice with 0.5 ml tetrahydrofuran/H₂O (1/1). Combined supernatant and washings were evaporated to dryness and the residue was dissolved in a solution of tetrabutylammonium fluoride in tetrahydrofuran (1.0 M, 1.5 ml) and incubated for 14 h at 37°C for removal of 2'-*O*-silyl protecting groups. The reaction was quenched by addition of triethylammonium acetate/H₂O (1.0 M, 1.5 ml, pH 7.4). Tetrahydrofuran was removed under reduced pressure and the sample was desalted with size-exclusion column chromatography (GE Healthcare, HiPrep[™] 26/10 Desalting; Sephadex G25) eluting with H₂O; collected fractions were evaporated and the RNA dissolved in H₂O (1 ml). The crude RNA was purified by anion exchange chromatography on a semipreparative Dionex DNAPac[®] PA-100 column (9 mm \times 250 mm) at 80°C with a flow rate of 1 ml/min (eluent A: 6 M urea, 25 mM Tris-HCl, pH 8.0; eluent B: 500 mM NaClO₄, 6 M urea, 25 mM Tris-HCl, pH 8.0). Fractions containing RNA were diluted with 0.1 M triethylammonium bicarbonate solution, loaded on a C18 Sep-Pak Plus[®] cartridge (Waters/Millipore), washed with H₂O and eluted with acetonitrile/H₂O (1/1). Crude and purified RNA were analyzed by anion exchange chromatography on a GE Healthcare Äkta Explorer HPLC System containing a Dionex DNAPac[®] PA-100 column (4 mm \times 250 mm) at 80°C with a flow rate of 1 ml/min. For RNA shorter or equal to 15 nucleotides, a gradient of 0–40% B in 30 minutes and for RNA longer than 15 nucleotides a gradient of 0–60% B was used; Eluent A: 6 M urea, 25 mM Tris-HCl, pH 8.0; Eluent B: 500 mM NaClO₄, 6 M urea, 25 mM Tris-HCl, pH 8.0. HPLC traces were recorded at UV absorption by

260 nm. RNA quantification was performed on an Implen P300 Nanophotometer.

Mass spectrometry of oligoribonucleotides

RNA samples (3 μ l) were diluted with 40 mM $\text{Na}_2\text{H}_2(\text{EDTA})/\text{H}_2\text{O}$ (5/4) for a total volume of 30 μ l, injected onto a C18 XBridge column (2.5 μ m, 2.1 mm \times 50 mm) at a flow rate of 0.1 ml/min and eluted with 0–100% B gradient at 30°C (eluent A: 8.6 mM triethylamine, 100 mM 1,1,1,3,3,3-hexafluoroisopropanol in H_2O ; eluent B: methanol). RNA HPLC runs were analyzed on a Finnigan LCQ Advantage Max electrospray ionization mass spectrometer with 4.0 kV spray voltage in negative mode. Characterization of the $c^3\text{G}$ modified 47mer was performed on a 7 T FT-ICR-MS (Bruker Daltonics – Apex Ultra 70). Sample preparation: 2 μ M RNA in 1:1 $\text{H}_2\text{O}/\text{CH}_3\text{OH}$ with 20 mM imidazole and 20 mM piperidine (pH \sim 10). The sample was measured in negative ion mode at a flow rate of 90 μ l/h and a skimmer potential of –50 V for ESI and –60 V for CAD (spectra averaged CAD: 400, spectra averaged ESI: 50, capillary voltage 3.1 kV, spray shield 2.7 kV, dry temperature 200°C. A collision energy of 12 V was applied in the CAD experiments.

Melting curve measurements of oligoribonucleotides

RNA samples were lyophilized as triethylammonium salts, dissolved in 800 or 330 μ l buffer (10 mM Na_2HPO_4 (pH 7.0) containing 150 mM NaCl) and transferred into UV permeable high precision cells made of quartz SUPRASIL[®] with a light path of 10 mm or 1 mm. UV melting profiles were recorded at 250 and 260 nm on a Varian Cary-100 spectrophotometer equipped with a multiple cell holder and a peltier temperature control device. Each RNA was measured at five different concentrations ranging from 1 to 100 μ M. T_m values were determined by calculating the first derivative. Thermodynamic parameters were determined according to (49,50). The error limits for ΔG , ΔH and ΔS reflect the standard deviation of three (duplex) or five (hairpin) independent measurements for a confidence interval of 95%.

NMR experiments

RNA samples were lyophilized as triethylammonium salts, dissolved in 500 μ l NMR buffer (15 mM $\text{Na}[\text{AsO}_2(\text{CH}_3)_2]\cdot 3\text{H}_2\text{O}$, 25 mM NaCl, 3 mM NaN_3 , in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1, pH 6.5) and transferred into 5 mm NMR tubes. Sample concentrations varied between 0.05 and 0.3 mM and experiments were run at 298 K unless otherwise stated. All NMR experiments were conducted on a Bruker 600 MHz Avance II+ NMR or a 700 MHz Avance Neo NM both equipped with a Prodigy TCI probe.

Determination of imino proton water exchange rates by CLEANEX-PM experiment. The CLEANEX-PM NMR experiments were carried out in the following buffer: 15 mM $\text{Na}[\text{AsO}_2(\text{CH}_3)_2]\cdot 3\text{H}_2\text{O}$, 25 mM NaCl, 3 mM NaN_3 , in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1, pH 6.5. The same buffer stock solution was used for all samples to rule out changes in imino

proton—bulk water exchange rates due to differences in the buffer composition. The CLEANEX-PM pulse sequence is available from the Bruker standard experiment collection (*zgcxesgp*) with an excitation sculpting water suppression element. A standard excitation sculpting water suppression experiment (*zgesgp*) was used as the reference experiment. For the determination of the water T_1 relaxation times under the experimental conditions a saturation recovery experiment was used and the longitudinal water relaxation time ranged between 3.03 s ($c^3\text{G}$ RNA) and 3.13 s (unmodified RNA). The following experimental parameters were used: spectral width 24 ppm, o1p: 4.7 ppm, number of scans 2048, dummy scans 32, interscan delay 1.5 s. Shaped pulse parameters were set via the *getprosol* command and the hard 90° ^1H pulse. The mixing times were set to 5, 25, 50, 50, 100, 100, 150, 200, 300, 400, 400 and 500 ms for both duplexes. The NMR spectra were processed in TOPSPIN 4.0.9 using a line broadening factor of 10 Hz (lb value set to 10). The absolute peak intensities of the CLEANEX-PM experiments and the reference experiments were determined and used to obtain the relative intensities. The data was then exported to MATLAB and the build-up curves were fitted to the following equation:

$$I/I_0 = \frac{k}{(R_{1,A} + k - R_{1,\text{water}})} \times \{e^{-R_{1,\text{water}}\tau_{\text{mix}}} - e^{-(R_{1,A}+k)\tau_{\text{mix}}}\}$$

with I/I_0 relative peak intensity, k imino proton—bulk water exchange rate (s^{-1}), $R_{1,\text{water}}$ longitudinal water exchange rate, τ_{mix} CLEANEX-PM mixing times (s) and $R_{1,A}$ is a combination of the longitudinal and transverse relaxation rate of the imino proton resonance (floating parameter during fitting). Errors in exchange rates were estimated from replicate experiments and obtained from 1000 Monte-Carlo runs.

Determination of the pK_a value of 3-deazaguanosine by NMR spectroscopy. For the pH dependent ^1H and ^{13}C NMR experiments, 22 mg of 3-deazaguanosine were dissolved in 600 μ l buffer (15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.9) and transferred into a 5 mm NMR tube. The chemical shifts of C2 and C6 were recorded on a Bruker 600 MHz Avance II+ equipped with a Prodigy TCI probe. The pH-value was determined using a Sigma-Aldrich micro pH combination glass electrode. At each pH-value, ^1H NMR spectrum and ^{13}C NMR spectra at natural abundance with power gated decoupling were measured.

The pK_a values were obtained by fitting the data to the following equation:

$$\delta_{\text{obs}} = \frac{\delta_{\text{deprotonated}} \times 10^{(pH-pK_a)} + \delta_{\text{protonated}}}{1 + 10^{(pH-pK_a)}}$$

With δ_{obs} observed ^{13}C chemical shift at the respective pH value, $\delta_{\text{deprotonated}}$ ^{13}C chemical shift of deprotonated species and $\delta_{\text{protonated}}$ ^{13}C chemical shift of the protonated species.

Table 1. Thermodynamic data of deazapurine modified RNA^a

No.	RNA sequences 5' to 3'	T_m (°C) ^b	ΔT_m (°C)	ΔG°_{298} (kcal mol ⁻¹) ^c	ΔH°_{298} (kcal mol ⁻¹) ^c	ΔS° (cal mol ⁻¹ K ⁻¹) ^c
I	GGCAGAGGC / GCCUCUGCC	66.7	-	-16.5 ± 0.4	-79.7 ± 4.6	-212 ± 14
Ia	GGCAc ³ GAGGC / GCCUCUGCC	62.8	-3.9	-15.2 ± 0.5	-74.8 ± 4.7	-199 ± 14
Ib	GGCAc ⁷ GAGGC / GCCUCUGCC	63.8	-2.9	-16.0 ± 0.3	-80.0 ± 1.8	-215 ± 5
Ic	GGCAc ⁷ GAGGC / GCCUCUGCC	67.2	+0.5	-16.7 ± 0.6	-80.1 ± 5.4	-213 ± 16
Id	GGCAc ⁷ GAGGC / GCCUCUGCC	66.5	-0.2	-16.1 ± 0.6	-76.6 ± 4.2	-203 ± 12
II	GAAGGGCAACCUUCG	73.3	-	-7.0 ± 0.2	-52.8 ± 1.3	-153 ± 4
IIa	GAAc ³ GGGCAACCUUCG	64.2	-9.1	-6.2 ± 0.2	-53.5 ± 1.6	-159 ± 5
IIb	GAc ³ AGGGCAACCUUCG	65.4	-7.9	-6.4 ± 0.3	-53.8 ± 2.0	-159 ± 6
IIc	GAAc ⁷ GGGCAACCUUCG	68.5	-4.8	-6.5 ± 0.3	-51.5 ± 1.9	-151 ± 5
IId	GAc ⁷ AGGGCAACCUUCG	69.8	-3.5	-6.6 ± 0.4	-52.3 ± 3.3	-153 ± 10
III	GGUCGACC	58.3	-	-13.2 ± 0.9	-64.6 ± 8.6	-172 ± 26
IIIa	GGUCc ³ GACC	44.8	-13.5	-11.4 ± 0.5	-74.7 ± 6.4	-213 ± 20
IIIb	GGUCGc ³ ACC	48.5	-9.8	-11.7 ± 0.9	-68.1 ± 8.8	-189 ± 26
IIIc	GGUCc ⁷ GACC	55.9	-2.4	-13.4 ± 1.0	-70.6 ± 8.5	-192 ± 25
IIId	GGUCGc ⁷ ACC	55.6	-2.7	-13.0 ± 0.6	-67.0 ± 4.6	-181 ± 14
III'	GGCUAGCC	60.7	-	-14.5 ± 1.1	-72.3 ± 9.5	-194 ± 28
III'a	GGCUAc ³ GCC	47.7	-13.0	-11.8 ± 0.8	-70.9 ± 8.3	-199 ± 25
III'b	GGCUc ³ AGCC	47.7	-13.0	-10.7 ± 0.6	-56.5 ± 6.2	-154 ± 19
III'c	GGCUAc ⁷ GCC	58.6	-2.1	-13.7 ± 0.7	-68.8 ± 7.3	-185 ± 22
III'd	GGCUc ⁷ AGCC	59.7	-1.0	-14.2 ± 0.6	-71.8 ± 6.9	-193 ± 21

^aBuffer: 10 mM Na₂HPO₄, 150 mM NaCl, pH 7.0. ΔH and ΔS values were obtained by van't Hoff analysis or based on RNA concentration dependent measurements according to references 49 and 50.

^bThe estimated errors of UV-spectroscopically determined T_m values are $\pm 0.3^\circ\text{C}$.

^cErrors for ΔH and ΔS were determined from three (RNAs I, III, and III') and five (RNAs II) independent measurements; in general, errors arising from noninfinite cooperativity of two-state transitions and from the assumption of a temperature-independent enthalpy, are typically 10–15%. Additional error is introduced when free energies are extrapolated far from melting transitions; errors for ΔG are typically 3–5%.

Crystallization and structure solution of a c³A modified RNA

Oligoribonucleotides of 27 nucleotide length corresponding to the sequence of *Escherichia coli* 23 S rRNA sarcin loop (SRL) and containing a 3-deazaadenosine at position 2670 were used for crystallization. The RNA was dissolved at a concentration of $\sim 350 \mu\text{M}$ in a buffer made of Tris-HCl (10 mM), Na₂H₂EDTA (1 mM), pH 8.0. The RNA sample was then heated to 55°C and cooled down to 10°C using a temperature-controlled device equipped with a Peltier element. Only one unique cubic-shaped crystal of c³A2670-modified RNA could be obtained. It grew after one month at 20°C using vapor diffusion method by mixing two volumes of RNA sample with one volume of a crystallization buffer made of ammonium sulfate (2.5 M), magnesium acetate (10 mM), and 2-(*N*-morpholino)ethanesulfonic acid (MES) (50 mM), pH 5.6 (the other drops made in identical conditions led to spherulites). Prior data collection, the crystal was cryoprotected for about 5 min in a reservoir solution containing 15% of glycerol and 3.0 M of ammonium sulfate, flash-frozen in liquid ethane and then transferred into liquid nitrogen. The collection of X-ray diffraction data has been done on the X06DA beamline at the SLS synchrotron, Villigen, Switzerland. Processing of the data was done with the XDS Package (71) and the structure was solved by molecular replacement with MOLREP (72) using the related PDB ID 3DVZ unmodified SRL RNA model. The structure was refined with the PHENIX package (73). Models were built using Coot (74). Alternative conformations were visible for residues 2647–2650. Coordinates have been deposited with the PDB database (PDB ID 7L3R for c³A2670-modified SRL).

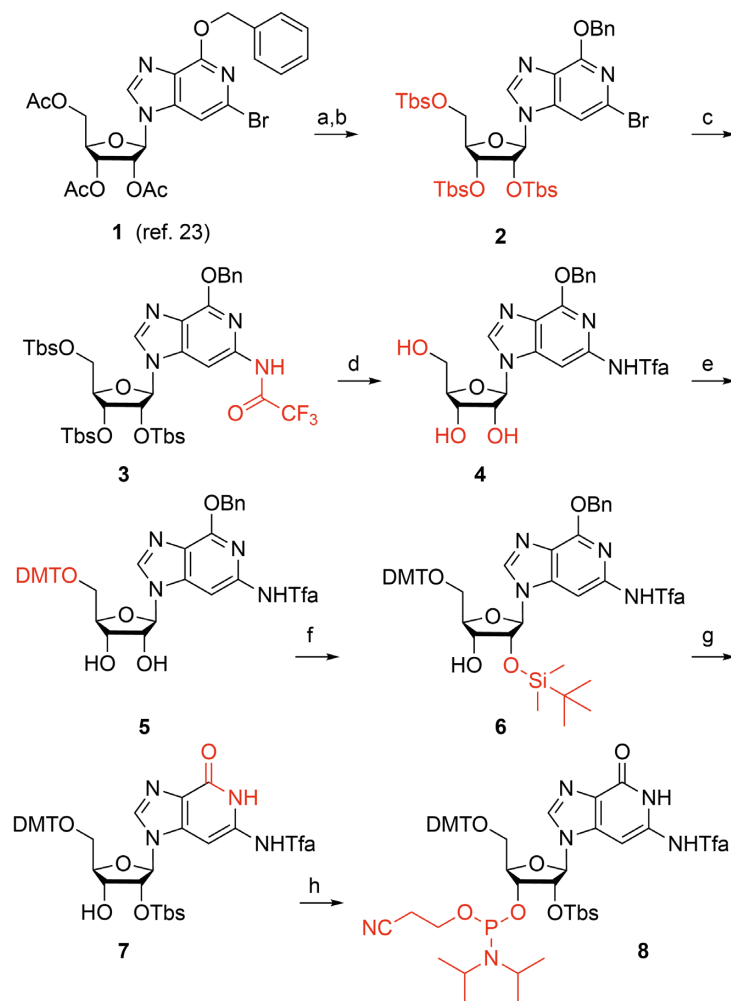
RESULTS AND DISCUSSION

Improved c³G building block with N²-Tfa and 2'-OTbs protection

For the solid-phase synthesis of c³A and c³G modified RNA, appropriate phosphoramidite building blocks are needed. Their syntheses have been a bottleneck for a long time (18,22), with new and more powerful routes coming up only recently (23,28).

A practical synthesis for c³G phosphoramidites from inexpensive starting materials has been introduced by our research group (23). Therein, the phenoxyacetyl (Pac) and 2-cyanoethoxymethyl (Cem) groups block the N²- and 2'-hydroxyl functionalities, while the O⁶ atom remains unprotected. This building block was successfully incorporated into a short 5 nt RNA strand. Deprotection of the N²-phenoxyacetyl group, however, requires rather long reaction times and high pH values which causes degradation when longer oligoribonucleotides are envisaged. To solve this problem, we here demonstrate the installation of a more base-labile protection group for the exocyclic amine of c³G, resulting in N²-trifluoroacetyl protected 3-deazaguanosine phosphoramidite **8**. Distinct to our previously introduced c³G building block, we used 2'-*O*-*tert*-butyldimethylsilyl (Tbs) protection (instead of the 2'-*O*-Cem group) to guarantee widest possible compatibility with standard RNA solid-phase synthesis.

The synthesis started from the previously reported 3-deazaguanosine key intermediate **1** (23) (Scheme 1). The application of the temporary *O*-*tert*-butyldimethylsilyl (Tbs) instead *O*-acetyl protection (compound **2**) was necessary because of otherwise partial acetyl transfer to the N² amino

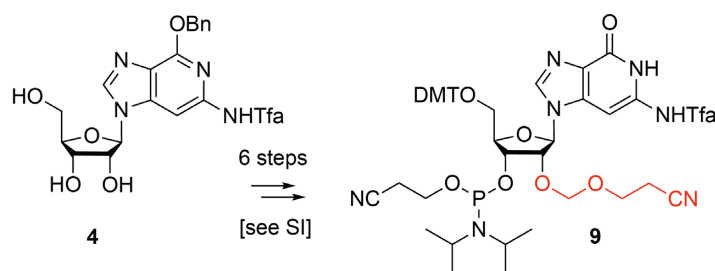


Scheme 1. Synthesis of the novel c^3G phosphoramidite building block **8**. Reagents and conditions: (a) MeNH_2 in ethanol, room temperature, 1 h, 99%; (b) 6 equiv *tert*-butyldimethylsilyl chloride (TbsCl), 7 equiv imidazole in DMF, room temperature, 72 h, 90%; (c) 1.2 equiv trifluoroacetamide, 1.4 equiv Cs_2CO_3 , 0.05 equiv $\text{Pd}_2(\text{dba})_3$, 0.15 equiv Xantphos in 1,4-dioxane, 110°C , 5 h, 90%; (d) 4.1 equiv tetra-*n*-butylammonium fluoride trihydrate in tetrahydrofuran, room temperature, 2 h, 87%; (e) 1.15 equiv 4,4'-dimethoxytriphenylmethyl chloride, 0.05 equiv 4-(*N,N*-dimethylamino)pyridine in pyridine, room temperature, 16 h, 74%; (f) 1.9 equiv TbsCl, 1.9 equiv AgNO_3 , in tetrahydrofuran, room temperature, 16 h, 39% (after two rounds of 2',3' equilibration); (g) 0.33 equiv Pd/C 10%, H_2 in tetrahydrofuran/ethanol 4:1, room temperature, 4 h, 79%; (h) 3.0 equiv 2-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphorodiamidite, 0.5 equiv 5-(benzylthio)-1*H*-tetrazole in dichloromethane, room temperature, 18 h, 79%.

group under the conditions required for the next step, namely the installation of the N^2 -trifluoroacetyl-amino (Tfa) functionality: Pd catalyzed cross coupling between trifluoroacetamide and bromo compound **2** by using Buchwald-Hartwig conditions gave derivative **3** in excellent yields. Deprotection of the Tbs groups proceeded smoothly with tetra-*n*-butylammonium fluoride (TBAF) in tetrahydrofuran to furnish compound **4**. Then, the functionalization of **4** into the desired c^3G phosphoramidite **8** required four more transformations. First, nucleoside **4** was converted into the dimethoxytritylated compound **5**, using 4,4'-dimethoxytriphenylmethyl chloride in pyridine. Then, Tbs-protection according to Ogilvie (39) with silver nitrate proceeded with rather low stereoselectivity (about 3:2), but in combination with base-induced equilibration the overall yield was increased to 39% for compound **6**. At this point, we note that we did not follow up our ini-

tial attempts to take advantage of the Beigelman 5',3'-silyl clamp and 2'-*O*-Tbs protection concept (40) for compound **1** because of low yields. The O^6 -benzyl group of nucleoside **6** was cleaved by Pd/C-catalyzed hydrogenation to provide the free lactam moiety in derivative **7**. Finally, phosphitylation was executed with 2-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphorodiamidite in the presence of 5-(benzylthio)-1*H*-tetrazole. Starting from nucleoside **1**, our route provides phosphoramidite **8** with 13% overall yield in eight steps involving eight chromatographic purifications; in total, 0.90 g of compound **8** was obtained during the course of this study.

Additionally, we synthesized the N^2 -Tfa protected building block **9** to take advantage of the less bulky 2'-*O*-(2-cyanoethoxy)methyl protecting group if short coupling times during solid-phase synthesis would be needed (see the Supporting Information and Scheme 2).



Scheme 2. Synthesis of the novel c^3G phosphoramidite building block **9**. For individual reaction steps, reagents and conditions see the Supporting Information.

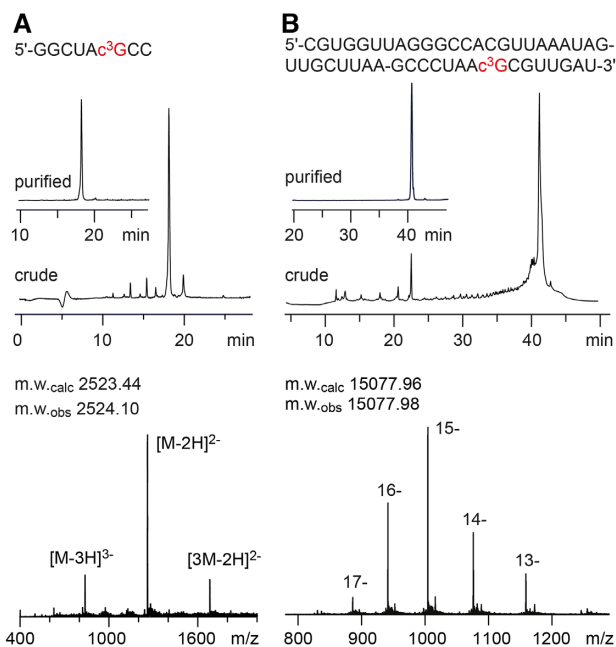


Figure 1. Characterization of c^3G -modified RNA. (A) Anion-exchange HPLC traces of 8 nt RNA (top) and LC-ESI iontrap mass spectrum (bottom); (B) Anion-exchange HPLC traces of 47 nt RNA (top) and ESI FT-ICR mass spectrum (bottom). HPLC conditions: Dionex DNAPac column (4×250 mm), 80°C , 1 ml min^{-1} , 0–40% buffer B (for 8 nt RNA) in buffer A within 30 min; 0–60% buffer B (for 47 nt RNA) in buffer A within 30 min; buffer A: Tris-HCl (25 mM), urea (6 M), pH 8.0; buffer B: Tris-HCl (25 mM), urea (6 M), NaClO₄ (0.5 M), pH 8.0. See the experimental for LC-ESI MS conditions.

Synthesis of c^3G containing RNA

The incorporation of building blocks **8** and **9** into oligoribonucleotides proceeds with high coupling rates and excellent yields using standard RNA solid-phase synthesis protocols and, as intended, significantly milder conditions can be used to reach complete deprotection of the c^3G containing oligonucleotides. Figure 1 exemplarily illustrates the synthesis of 8 and 47 nt long RNA oligonucleotides with a single c^3G modification (for a complete list of synthesized oligos see Supplementary Table S1). In short, the c^3G phosphoramidite building blocks **8** or **9** were applied in combination with standard *N*-acetylated 2'-*O*-[(triisopropylsilyl)oxy]methyl (TOM) phosphoramidites and the oligomers were assembled on controlled pore glass (CPG) supports (41). Cleavage from the solid support and

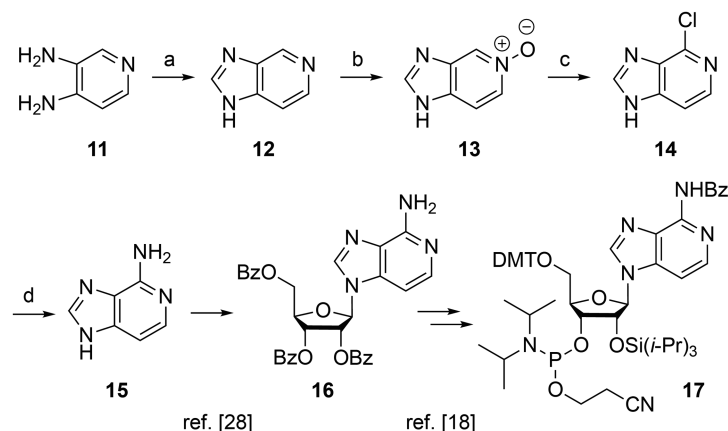
deprotecting of the base labile groups were accomplished by treatment with methylamine/ammonia in water (1:1 mixture of 40% aqueous methylamine and 30% aqueous ammonia (AMA) for 3 h at 37°C room temperature or 15 min at 65°C). Subsequently, deprotection of the 2'-*O*-silyl groups (and the 2'-*O*-Cem group in case of **9**) was carried out with tetra-*n*-butylammonium fluoride trihydrate in tetrahydrofuran for 14 hours at 37°C . The reaction was quenched by the addition of triethylammonium acetate buffer at pH 7.4. Salts were removed by size-exclusion chromatography. Analysis by anion exchange chromatography under strong denaturing conditions usually gave a major peak for the desired RNA which was further purified by anion exchange chromatography on a semipreparative column. The molecular weights of the purified RNAs were confirmed by LC-ESI-MS (Supplementary Table S1) and the sequences were confirmed by top-down mass spectrometry (MS) using a Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer (42) (Supplementary Table S2).

Synthesis of c^3A containing RNA

The synthetic path of the c^3A nucleoside building block has been merged from two routes that we published earlier (Scheme 3) (18,28). The key reaction is a silyl-Hilbert-Johnson nucleosidation using unprotected 6-amino-3-deazapurine **15** and benzoyl-protected 1-*O*-acetylribose (28). We then proceeded with *N*⁶-benzoyl protection (18) instead of *N*⁶-(di-*n*-butyl)amidine protection (28) for the following reason: The introduction of the 2'-*O*-triisopropylsilyl group is not regioselective and results in a mixture of 2'- and 3'-*O*-regioisomers; the required chromatographic purification of this mixture is more straightforward for the *N*⁶-benzoyl protected regioisomers (compared to the *N*⁶-(di-*n*-butyl)amidine counterparts) because of the significantly larger gap in retention times. Furthermore, in this work we also optimized the access to 6-amino-3-deazapurine **14** on large scale and starting from cheap materials (Scheme 3) (43,44). Solid-phase synthesis of c^3A -modified building block **17** into RNA was performed using standard protocols as described earlier (15).

Base pairing properties of c^3G and c^3A containing RNA

To the best of our knowledge, hardly anything is known about the impact of 3-deazapurine nucleosides on base pairing in regular RNA. We therefore set out to analyze the



Scheme 3. Synthesis of the novel c^3A phosphoramidite building block **16**. Reagents and conditions: (a) in formic acid, 140°C, 48 h, 97%; (b) 2.0 equiv 3-chloroperbenzoic acid in 2:1 (v/v) CH_2Cl_2/CH_3OH , room temperature, 16 h, 96%; (c) in $POCl_3$, 110°C, 4 h, 67%; (d) in 1:1 (v/v) 30% NH_3 in $H_2O/1,4$ -dioxane, 190°C, autoclave, 120 h, 87%.

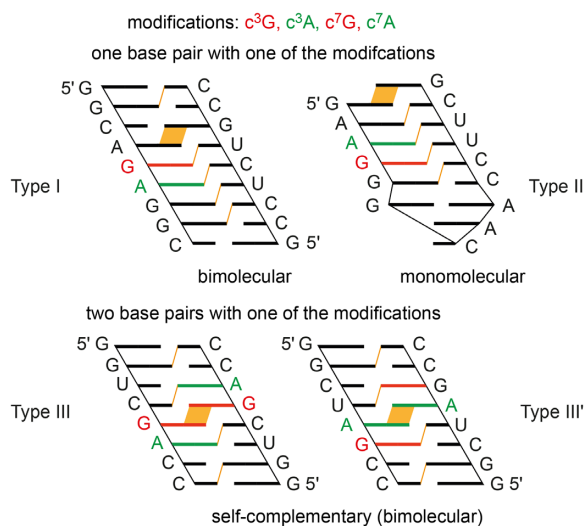


Figure 2. Sequence design for thermodynamic analysis of base pairing of deazapurine modified RNAs. Cartoon presentation to highlight inter-strand stacking interactions (in orange).

influence of c^3G and c^3A on the stability of RNA double helices, and additionally, to compare them to the c^7G and c^7A modified counterparts. Figure 2 illustrates the sequence design of the double helices investigated. The first motif constitutes an asymmetric bimolecular duplex of nine base pairs with a single deazapurine modification in the center (Type I). The second motif forms a hairpin with an extra-stable GNRA loop (GCAA), a five base pair stem with the modification again in the center, and additionally a 3'-dangling guanosine to reduce fraying of the terminal base pair (Type II). The third RNA motif covers two palindromic RNAs of eight base pairs and identical pyrimidine–purine stacking patterns (Type III and III') with the purine modification either side-by-side in the center (thereby enabling direct inter-strand stacking), or separated by two base pairs. The latter sequence design is highly sensitive for the thermodynamic impact arising from a modification because only

two and three regular Watson–Crick base pairs can form next to them: we remind that nucleation of a bimolecular double helix of oligonucleotides becomes thermodynamically favorable only when at least (three to) four continuous Watson–Crick base pairs can form (45,46). Therefore, these palindromic RNAs are expected to markedly respond to a nucleobase modification expressed in significant alterations of their thermal stabilities (T_m) and thermodynamic parameters (ΔG , ΔH , ΔS).

Table 1 summarizes the thermal and thermodynamic data we obtained for the four RNA systems by UV-spectroscopic melting profile measurements (for melting profiles see the Supplementary Figures S1 to S16). The native type I RNA melts at 66.7°C. Both 3-deazapurines exhibit a destabilization, by 3.9 and 2.9°C, for c^3G (in **Ia**) and c^3A (in **Ib**) respectively. Interestingly, in this sequence context, the corresponding 7-deazapurines (c^7G in **Ic** and c^7A in **Id**) have negligible effect on the T_m and thermodynamic parameters. For the hairpin RNA (type II), the destabilization of both 3-deazapurines is significant, reflected in a decrease of the T_m -value by 9.1 and 7.9°C, for c^3G (in **IIa**) and c^3A (in **IIb**) respectively, when compared to the native hairpin **II**. In contrast to type I RNA, the monomolecular system (type II) reveals a slight destabilizing effect of 7-deazapurines as well (c^7G in **IIc** and c^7A in **II'd**).

To further elucidate the impact of 3- and 7-deazapurines on base pairing, we investigated the thermodynamic base pairing properties of 3- and 7-deazapurines in short palindromic RNAs. As mentioned above, the effects become intensified because the self-pairing results in two modified base pairs with regularly Watson–Crick–paired regions next to them that are below the required number of base pairs needed for successful double helix nucleation (45,46). Indeed, for c^3G the reduction in both palindromic RNAs was dramatic, reflected in –13.5/–13.0°C reduced T_m values (**IIIa** and **III'a**), accounting for –6.8/–6.5°C destabilization per single modification which is more than twice of the destabilization that we observed for a single c^3G -C base pair in the bimolecular 9 bp duplex **Ia**. The destabilizing nature was also very pronounced for c^3A in both palindromic RNAs (–9.8°C in **IIIb** and –13.0°C in **III'b**). In contrast,

only little destabilization was found for both 7-deazapurines in both palindromic RNAs (c^7G : -2.4°C in **IIIc** and -2.1°C in **III'e**; c^7A : -2.7°C in **III'd** and -1.0°C in **III'f**).

To the best of our knowledge, our data provides the first insights into thermal and thermodynamic stabilities of 3-deazaadenine and 3-deazaguanine containing RNA. Thus far, c^3G was investigated only within the context of 2'-OCH₃-RNA; c^3G caused destabilization when paired to C (-6°C with complementary RNA; -9°C with complementary DNA in a Type I like sequence design) (47). Interestingly, when an acetyl group resided at the exocyclic amino group (*N*²-acetyl-3-deazaguanine) the native pairing strength was restored (48).

Furthermore, a single paper points at the strong destabilizing effect of 3-deazaguanine on DNA base pairing (36). In contrast, in several early studies, the slight destabilizing effect of 3-deazaadenine within DNA double helices has been recognized (31,32,35,36). 7-Deazaguanine and 7-deazaadenine modified DNAs have been investigated most intensively with respect to base pairing stabilities; generally, they act slightly destabilizing compared to their native counterparts (29,30,34–38).

NMR spectroscopy of 3-deazapurine modified RNA

The NMR resonances for the hydrogen-bonded protons of Watson–Crick base-pairs ('imino protons') directly reflect the double helical segments within folded RNA. The chemical shifts of these signals are characteristic for A-U (>14 ppm) and C-G base pairs (~ 12 – 13 ppm), and the linewidths reflect proton exchange with the solvent. Clearly, they are very sensitive to modifications, in particular, if they concern the nucleobases. Figure 3A depicts ¹H NMR spectra of the palindromic duplex 5'-GGUCGACC **III** that were collected as a function of temperature, over the range 5–25°C. The signal assignment was based on 2D ¹H/¹H-NOESY spectroscopy (Supplementary Figure S17). With increasing temperature, the signal of the base pair G1-C becomes broader and loses intensity, consistent with base pair fraying of the duplex termini; the internal base pairs remain and reflect stable duplex formation at ambient temperatures.

In comparison, the c^3G imino proton resonance of the modified duplex **IIIa** was shifted highfield (Figure 3B), consistent with the increased pK_a of c^3G (compared to G) of almost 3 pK_a units (Table 2). Interestingly, the neighboring U3-A base pair senses the perturbation caused by c^3G , reflected in a decrease in intensity of the U3 imino proton signal at higher temperature. The enhanced exchange with the solvent indicates a dynamic hot spot of base pair opening at this particular position.

Concerning the c^3A -modified duplex **IIIb**, the thermal destabilization is reflected in the imino proton ¹H NMR spectra as well (Figure 3C), with a tiny signal barrow at 5°C in the typical ppm region of A–U base pairs; no signal was observed at higher temperature. This stands for intensive proton exchange with the solvent of the uridine N3-H in the c^3A -U base pair and is reminiscent to ¹H NMR spectroscopy of 5F-U–A base pairs (52).

To investigate these effects in more detail, we quantified the water imino proton exchange kinetics of a c^3G -C base pair containing duplex (**IIIa**) by the application of

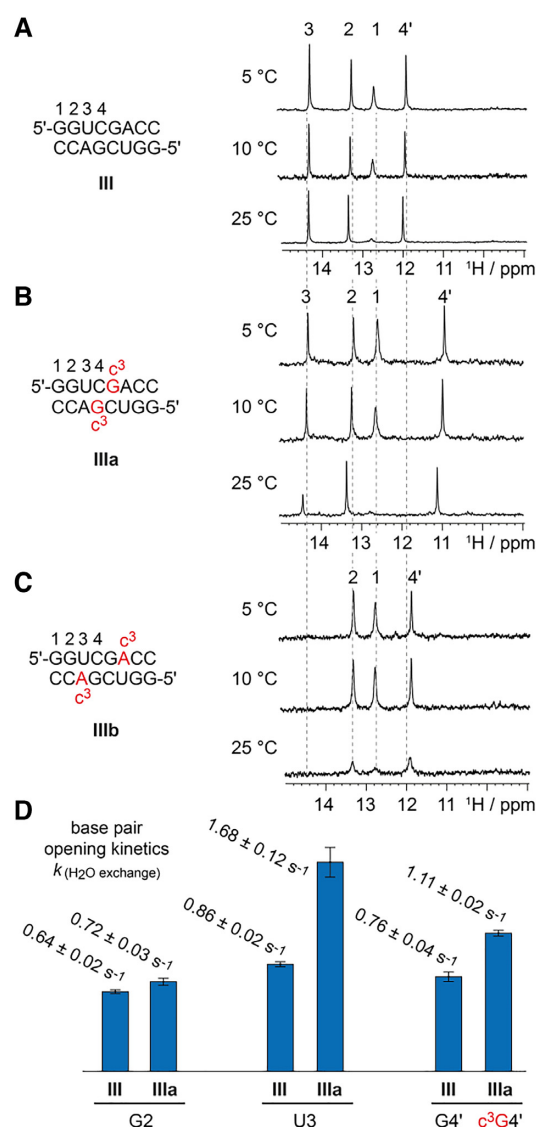


Figure 3. Comparative ¹H NMR imino proton spectra of (A) unmodified, (B) c^3G -, and (C) c^3A -modified oligoribonucleotides; conditions: $c_{\text{RNA}} = 0.1$ mM, 15 mM Na[AsO₂(CH₃)₂] \cdot 3H₂O, 25 mM NaCl, 3 mM NaN₃, in H₂O/D₂O 9/1, pH 6.5; (D) the graph illustrates the imino proton exchange rates determined by 1D CLEANEX-PM experiments (see Supporting Information for details).

Table 2. pK_a values of nucleobases and deaza derivatives^a

	A•H ⁺ (N1)	c^3A •H ⁺ (N1)	c^7A •H ⁺ (N1)	U (N3)
pK_a	3.7	6.8	5.3	9.2
	G (N1)	c^3G (N1) ^b	c^7G (N1)	C•H ⁺ (N3)
pK_a	9.5	12.3	10.3	4.1

^aValues are summarized from references 7 and 51. Atom position of protonation/deprotonation are provided in brackets.

^bFor an independent NMR-based pK_a determination of c^3G see Supplementary Figure S22.

1D Phase-Modulated CLEAN Chemical EXchange Spectroscopy (CLEANEX-PM) NMR experiments (Figure 3D, and Supplementary Figure S18) (53–55). Interestingly, the largest exchange rate $k(\text{H}_2\text{O ex.})$ of 1.68 s^{-1} was observed for

the A-U base pair next to the c^3G -C modification. Also for the c^3G -C base pair itself the rate $k_{(H_2O\ ex.)}$ was increased compared to the unmodified duplex ($1.11\ s^{-1}$ versus $0.76\ s^{-1}$), while the G-C base pair two positions away almost the same exchange rates $k_{(H_2O\ ex.)}$ for modified and unmodified duplexes were observed ($0.72\ s^{-1}$ versus $0.62\ s^{-1}$). This is consistent with a local structural perturbation induced by the 3-deazapurine hotspot.

Crystal structures of a c^3A containing RNA

To further shed light on 3-deazapurine containing base pairs in RNA we put efforts into a high resolution X-ray crystallographic analysis. We utilized the 27 nt fragment of the *E. coli* 23S rRNA sarcin-ricin loop (SRL) which is a robust and well-behaved crystallization scaffold (56,57). For the incorporation of c^3A , we deemed nucleotide A2670 appropriate, which forms a Watson-Crick base pair with C2650 in the regular A-form double helical region. Crystallization trials with c^3A -modified SRL RNA indeed provided crystals that diffracted to subatomic resolution (Supplementary Table S3). X-ray structure determination showed that the c^3A nucleobase is well-defined in the electron density maps for the c^3A -modified RNA (Figure 4). Superimpositions of the c^3A -modified RNA structure with the unmodified RNA showed a root-mean-square deviation (rmsd) of $0.12\ \text{\AA}$ (within the errors on coordinates of $0.13\ \text{\AA}$). Overall, the structure revealed that the 3-deazanucleobase does not significantly affect the base pair geometry and the SRL RNA fold. However, what is different is the local hydration network in the minor groove of the c^3A -U base pair (Figure 4A). While three ordered water molecules are found in the minor groove for the unmodified base pair, four are observed for the modified base pair. An obvious distinction is that the N3 atom of A2670 is hydrogen-bonded to one of the water molecules ($2.8\ \text{\AA}$ distance) while for the c^3A 2670/U2650 base pair, a continuous hydrogen-bonded water chain spans from the 2'-OH of c^3A 2670 to the 2'-OH of U2650 without involving the c^3A nucleobase; the closest water is in $3.3\ \text{\AA}$ -distance from its C3 atom (Figure 4D, Supplementary Figure S19). Although these structural differences in the hydration pattern appear minor in the static X-ray structure, they may contribute to increased base pair opening dynamics as reflected in the imino proton NMR spectrum of the c^3A modified SRL RNA by significant signal broadening (Supplementary Figure S20). The destabilizing effect of the c^3A modification in the SRL RNA hairpin is further reflected in the UV-melting profiles measured under diverse conditions (Supplementary Figure S21).

Towards a rationale for the destabilization by c^3G and c^3A

For all four modifications— c^3G , c^7G and c^3A , c^7A —the base pair geometry and hydrogen bonding pattern to C and A respectively, are the same as encountered for the native Watson-Crick G-C and A-U base pairs. Therefore, other factors such as altered nucleobase dipole moments that can affect nucleobase stacking (58), or altered pK_a values of the modified nucleobases (7,59–61), as well as altered hydration of minor groove (3-deazapurines) and major groove (7-

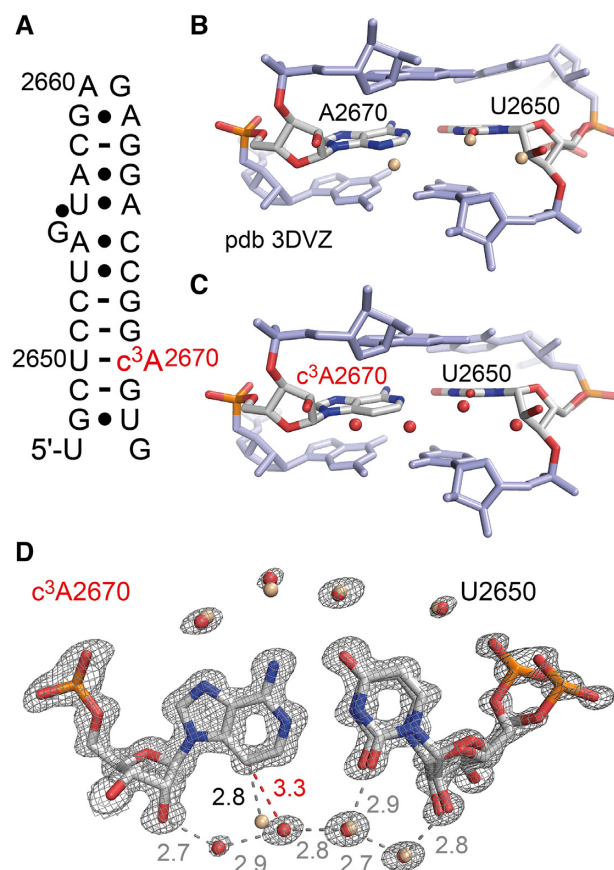


Figure 4. X-ray structure of 3-deazaadenosine modified RNA at atomic resolution. (A) Secondary structure of the *E. coli* Sarcin-ricin stem-loop (SRL) RNA used for crystallization. The c^3A nucleotide is labeled in red. (B) Side view on the U2650/A2670 base pair in the unmodified duplex with ordered water molecules in the minor groove highlighted in beige color (PDB ID 3DVZ). (C) Side view on the U2650/ c^3A 2670 base pair with ordered water molecules in the minor groove highlighted in red color (PDB ID 7L3R). (D) $2F_{obs} - F_{calc}$ electron density map contoured at $1.5\ \sigma$ level showing the U2650/ c^3A 2670 base pair. Water molecules are shown as red spheres (PDB ID 7L3R) and are superposed to the water molecules observed in the unmodified RNA (PDB ID 3DVZ). Numbers are distances in Angstrom (\AA). We note that the alternative conformations of U2650 (and C2649, G2648) are also observed in the unmodified SRL RNA scaffold (PDB ID 3DVZ).

deazapurine) are of potential origin for these differences in thermal stabilities and thermodynamic parameters (30).

It has been proposed that the pK_a of a nucleobase is one significant factor that contributes to base-pairing strength (7,58–60). Table 2 summarizes the pK_a values of the nucleobases investigated in this study, and the following trend becomes obvious: The pK_a decreases in the order of $A\bullet H^+ < c^7A\bullet H^+ < c^3A\bullet H^+$, and the same is true for $G < c^7G < c^3G$. This trend is consistent with the notion that 7-deazapurines are less destabilizing than 3-deazapurines compared to A and G in standard Watson-Crick pairs in RNA double helices (Table 1). We furthermore note that c^3A with a pK_a of 6.8 is partially protonated at physiological pH values around 7.0 which may affect base pairing in addition.

Another aspect that was documented earlier is that natural Watson-Crick base pairs A-U and C-G are characterized by a pK_a gap between the H-donor (U, G; N3-H, N1-

H) and the H-acceptor (A, C; N1, N3) of ΔpK_a 5.5 and 5.4, respectively, which appears optimal for efficient base pairing (7,58). This gap becomes significantly altered for isomeric c^7A-U (ΔpK_a 3.9), c^3A-U (ΔpK_a 2.4), c^7G-C (ΔpK_a 6.2) and c^3G-C (ΔpK_a 8.2) deazapurine base pairs.

Clearly, caution has to be taken because many more factors in addition to the pK_a of the heterocycles (as listed above) contribute to macroscopically observed double helix stabilities. This might be the reason why concepts for predicting hydrogen-bond strengths from acid–base molecular properties (e.g. proton affinity/ pK_a equalization, pK_a slide rule, etc.) cannot be directly applied to nucleic acid base pairing strengths, although being promising for an integrated theoretical and experimental approach to solve this long-standing problem in the future (62).

Finally, we point at the hydration of RNA grooves which is distinct between 3-deazapurines (c^3G and c^3A) affecting the minor groove, and 7-deazapurines (c^7G and c^7A) affecting the major groove. In particular, the first solved crystal structure of a c^3A modified RNA and the comparison to the unmodified counterpart reveal that the characteristic water molecule that always docks on the purine N3 atom and often spans further to the ribose 2'-OH in native RNA double helices (63,64), is lacking in the 3-deaza-modified counterparts. The altered hydration pattern is a clear structural distinction that should be taken into account when searching for the concurrent causes of the decrease in stability.

Role of deazapurines in RNA atomic mutagenesis studies

Deazapurine modified RNAs have been frequently applied in biochemical mutagenesis studies of ribozymes to verify or falsify mechanistic proposals for the chemical reactions they catalyze (65,66). Several examples are found in the literature where atomic mutagenesis experiments led to an in-depth understanding of the chemical mechanism, revealing the functionally crucial imino groups of purines that participate in general acid-base catalysis of small nucleolytic ribozymes for phosphodiester cleavage (14–16). For instance, this concerns the twister ribozyme where proton transfer from the (protonated) N3 of the conserved adenine (A6) at the cleavage site to the 5'-O leaving group had been postulated to contribute to reaction catalysis (i.e. δ -catalysis; for explanation of α , β , γ , δ -catalysis see reference 67). Indeed, replacement of this adenine by c^3A or c^1c^3A rendered the twister ribozyme inactive (14,67). A follow-up NMR spectroscopic analysis revealed a pK_a shift of the very adenine N3 from 3.7 to 5.1 (14). As a second example, the pistol ribozyme is pointed out. At first, distance analysis and interpretation of the three-dimensional structure of this ribozyme implicated the N3 of a conserved purine (A32 or G32, respectively) as a possible candidate for general acid–base catalysis, being involved in leaving group stabilization (δ -catalysis). However, upon replacement of this purine by c^3A , the ribozyme retained full activity. Later, it was found that the 2'-OH of this purine nucleoside plays a crucial role, and not the N3 (15).

Additionally, deaza-nucleobase substitutions have identified crucial metal binding sites that are responsible for structuring of riboswitch aptamers and ribozyme active sites (13,67), as well as have disclosed key coordination sites

of hydrated Mg^{2+} ions that critically participate in proton transfer to release the 5'-O leaving group in the course of phosphodiester cleavage (15,16,68,69). Recent prominent examples refer to the NAD^+ class-I riboswitch (68) and the pistol ribozyme (15,16). In pistol RNA, a highly conserved guanine (G33) interacts with a Mg^{2+} cation through inner-sphere coordination with its N7 atom. Replacement of this guanine by c^7G renders the ribozyme inactive (15). Together with distance analysis of pistol ribozyme crystal structures, this suggests that a water molecule of the hydrated Mg^{2+} ion assists in stabilizing the 5'-O leaving group during the course of the reaction (15,16). In the NAD^+ class-I riboswitch, a highly conserved adenine (A10) exhibits N7-innersphere coordination to a Mg^{2+} cation that additionally coordinates to the oxygen atom of a sequence-distant phosphate backbone unit and thus molds a critical part of the binding pocket. Replacement of this adenine by c^7A abolishes binding to NAD^+ (68).

With respect to the ribosome, atomic mutagenesis involving 3-deazaadenine in the ribosomal peptidyl transferase center (PTC) demonstrated that an initial suspect (N3 of A2451) did not actively participate in proton transfer to catalyze peptide bond formation, but instead, the 2'-OH of the very adenosine assists in positioning of the substrate (19; see also reference 17,18,20,70).

For atomic mutagenesis experiments using deazanucleobases, a profound knowledge of the intrinsic properties of these heterocycles in an RNA context is imperative to avoid misinterpretations. Functional assays to evaluate ribozyme or riboswitch activity in response to a specific purine-deazapurine replacement are usually designed to impair a potential H-bond or inner-sphere coordination to a divalent metal ion, as suggested by the observed distances in the three-dimensional structure. Provided such an interaction is crucial for activity, it can thus be easily identified and the mechanistic interpretation becomes straightforward. Importantly, in all of the examples mentioned above, the purine/deazapurine finds itself in an exposed and unpaired position. Clearly, in the light of our present study, functional assays that would involve atomic mutagenesis with 3-deazanucleobases in double helical regions would encounter severe limitations because of thermodynamic destabilization of the stem containing the modification. This aspect is, however, less critical for atomic mutagenesis using 7-deazanucleobases which do not or only minimally destabilize double helices.

CONCLUDING REMARKS

In this study, we synthesized and comprehensively analyzed the properties of 3-deazaguanine and 3-deazaadenine modified RNA. When these nucleobases are embedded in a double helical environment we found that they can decrease the stability of base pairing. Interestingly, the effects were much more pronounced for 3-deazapurines compared to the isomeric 7-deazapurines. Moreover, destabilization is more severe in short stems interfering with double helix nucleation, while a more modest decrease is observed if located in the center of an extended double helix. We furthermore found that the imino proton exchange for deazapurine-modified Watson Crick base pairs is significantly higher compared

to their native counterparts and simultaneously induces enhanced opening of the very neighbor base pair. These perturbations in local structural dynamics provide a rationale for the observed reduced base pairing strength. Additionally, we solved the X-ray structure of a c³A modified RNA to visualize the hydration pattern of the minor groove. The structure points out that the hydrophobic minor groove face of c³A impairs hydration because the water molecule that is usually hydrogen bonded to the N3 atom becomes extruded from the 3-deaza-modified base pair. Taken together, our comparative study sheds light on Watson–Crick base pairing of 3- and 7-deazapurines and thus represents a first important step towards a comprehensive understanding of the intrinsic properties of these RNA modifications that are critical for their proper application in atomic mutagenesis experiments of biologically relevant RNA.

DATA AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession numbers 7L3R.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Jonas Kottersteger for synthetic contributions, Sarah Heel for assistance in FT-ICR mass spectrometry, and Daniel Fellner (Innsbruck) for technical support.

FUNDING

Austrian Science Fund FWF [P31691, F8011-B to R.M., P32773 to C.K., P30087 to K.B.]; Austrian Research Promotion Agency FFG [West Austrian BioNMR 858017]. Funding for open access charge: Austrian Science Fund FWF.

Conflict of interest statement. None declared.

REFERENCES

- Polacek, N. (2013) Atomic mutagenesis of the ribosome: towards a molecular understanding of translation. *Chimia – Int. J. Chem.*, **67**, 322–326.
- Hoernes, T.P., Clementi, N., Juen, M.A., Shi, X., Faserl, K., Willi, J., Gasser, C., Kreutz, C., Joseph, S., Lindner, H. *et al.* (2018) Atomic mutagenesis of stop codon nucleotides reveals the chemical prerequisites for release factor-mediated peptide release. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E382–E389.
- Hu, T., Suter, S.R., Mumbleau, M.M. and Beal, P.A. (2018) TLR8 activation and inhibition by guanosine analogs in RNA: Importance of functional groups and chain length. *Bioorg. Med. Chem.*, **26**, 77–83.
- Forconi, M., Benz-Moy, T., Gleitsman, K.R., Ruben, E., Metz, C. and Herschlag, D. (2012) Exploring purine N7 interactions via atomic mutagenesis: the group I ribozyme as a case study. *RNA*, **18**, 1222–1229.
- Seela, F., Debalak, H., Usman, N., Burgin, A. and Beigelman, L. (1998) 1-Deazaadenosine: synthesis and activity of base-modified hammerhead ribozymes. *Nucleic Acids Res.*, **26**, 1010–1018.
- Kapinos, L.E., Operschall, B.P., Larsen, E. and Sigel, H. (2011) Understanding the acid-base properties of adenosine: the intrinsic basicities of N1, N3 and N7. *Chem. Eur. J.*, **17**, 8156–8164.
- Krishnamurthy, R. (2012) Role of pK_a of nucleobases in the origins of chemical evolution. *Acc. Chem. Res.*, **45**, 2035–2044.
- Acharya, P., Cheruku, P., Chatterjee, S., Acharya, S. and Chattopadhyaya, J. (2004) Measurement of nucleobase pK_a values in model mononucleotides shows RNA–RNA duplexes to be more stable than DNA–DNA duplexes. *J. Am. Chem. Soc.*, **126**, 2862–2869.
- Bande, O., Braddick, D., Agnello, S., Jang, M., Pezo, V., Schepers, G., Rozenski, J., Lescrier, E., Marlière, P. and Herdewijn, P. (2016) Base pairing involving artificial bases in vitro and in vivo. *Chem. Sci.*, **7**, 995–1010.
- Fuchs, E., Falschlunger, C., Micura, R. and Breuker, K. (2019) The effect of adenine protonation on RNA phosphodiester backbone bond cleavage elucidated by deaza-nucleobase modifications and mass spectrometry. *Nucleic Acids Res.*, **47**, 7223–7234.
- Spitale, R.C., Volpini, R., Heller, M.G., Krucinska, J., Cristalli, G. and Wedekind, J.E. (2009) Identification of an imino group indispensable for cleavage by a small ribozyme. *J. Am. Chem. Soc.*, **131**, 6093–6095.
- Spitale, R.C., Volpini, R., Mungillo, M.V., Krucinska, J., Cristalli, G. and Wedekind, J.E. (2009) Single-atom imino substitutions at A9 and A10 reveal distinct effects on the fold and function of the hairpin ribozyme catalytic core. *Biochemistry*, **48**, 7777–7779.
- Zheng, L., Falschlunger, C., Huang, K., Mairhofer, E., Yuan, S., Wang, J., Patel, D.J., Micura, R. and Ren, A. (2019) Hatchet ribozyme structure and implications for cleavage mechanism. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 10783–10791.
- Košutić, M., Neuner, S., Ren, A., Flür, S., Wunderlich, C., Mairhofer, E., Vušurović, N., Seikowski, J., Breuker, K., Höbartner, C. *et al.* (2015) A mini-twister variant and impact of residues/cations on the phosphodiester cleavage of this ribozyme class. *Angew. Chem. Int. Ed.*, **54**, 15128–15133.
- Neuner, S., Falschlunger, C., Fuchs, E., Himmelstoss, M., Ren, A., Patel, D.J. and Micura, R. (2017) Atom-specific mutagenesis reveals structural and catalytic roles for an active-site adenosine and hydrated Mg²⁺ in pistol ribozymes. *Angew. Chem. Int. Ed.*, **56**, 15954–15958.
- Teplova, M., Falschlunger, C., Krasheninina, O., Egger, M., Ren, A., Patel, D.J. and Micura, R. (2020) Crucial roles of two hydrated Mg²⁺ ions in reaction catalysis of the pistol ribozyme. *Angew. Chem. Int. Ed.*, **59**, 2837–2843.
- Erlacher, M.D., Lang, K., Shankaran, N., Wotzel, B., Hüttenhofer, A., Micura, R., Mankin, A.S. and Polacek, N. (2005) Chemical engineering of the peptidyl transferase center reveals an important role of the 2'-hydroxyl group of A2451. *Nucleic Acids Res.*, **33**, 1618–1627.
- Erlacher, M.D., Lang, K., Wotzel, B., Rieder, R., Micura, R. and Polacek, N. (2006) Efficient ribosomal peptidyl transfer critically relies on the ribose 2'-OH at A2451 of 23S rRNA. *J. Am. Chem. Soc.*, **128**, 4453–4459.
- Lang, K., Erlacher, M., Wilson, D.N., Micura, R. and Polacek, N. (2008) The role of 23S ribosomal RNA residue A2451 in peptide bond synthesis revealed by atomic mutagenesis. *Chem. Biol.*, **15**, 485–492.
- Polikanov, Y.S., Steitz, T.A. and Innis, C.A. (2014) A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nat. Struct. Mol. Biol.*, **21**, 787–793.
- Das, S.R. and Piccirilli, J.A. (2005) General acid catalysis by the hepatitis delta virus ribozyme. *Nat. Chem. Biol.*, **1**, 45–52.
- Seela, F., Debalak, H., Andrews, L. and Beigelman, L. (2003) Synthesis and enzymic hydrolysis of oligoribonucleotides incorporating 3-deazaguanosine: the importance of the nitrogen-3 atom of single conserved guanosine residues on the catalytic activity of the hammerhead ribozyme. *Helv. Chim. Acta*, **86**, 2726–2740.
- Mairhofer, E., Flemmich, L., Kreutz, C. and Micura, R. (2019) Access to 3-deazaguanosine building blocks for RNA solid-phase synthesis involving Hartwig-Buchwald C–N cross-coupling. *Org. Lett.*, **21**, 3900–3903.
- Perlíková, P. and Hocek, M. (2017) Pyrrolo[2, 3-d]pyrimidine (7-deazapurine) as a privileged scaffold in design of antitumor and antiviral nucleosides. *Medicinal Res. Rev.*, **37**, 1429–1460.
- Richardson, K.E. and Znosko, B.M. (2016) Nearest-neighbor parameters for 7-deaza-adenosine-uridine base pairs in RNA duplexes. *RNA*, **22**, 934–942.
- Ingale, S.A., Leonard, P. and Seela, F. (2018) Glycosylation of pyrrolo[2,3-d]pyrimidines with 1-O-acetyl-2, 3, 5-tri-O-benzoyl-β-d-ribofuranose: substituents and protecting groups

- effecting the synthesis of 7-deazapurine ribonucleosides. *J. Org. Chem.*, **83**, 8589–8595.
27. Seela, F., Mersmann, K., Grasby, J.A. and Gait, M.J. (1993) 7-Deazaadenosine – oligoribonucleotide building-block synthesis and autocatalytic hydrolysis of base-modified hammerhead ribozymes. *Helv. Chim. Acta*, **76**, 1809–1820.
 28. Mairhofer, E., Fuchs, E. and Micura, R. (2016) Facile synthesis of a 3-deazaadenosine phosphoramidite for RNA solid-phase synthesis. *Beilstein J. Org. Chem.*, **12**, 2556–2562.
 29. Gold, B., Stone, M.P. and Marky, L.A. (2014) Looking for Waldo: a potential thermodynamic signature to DNA damage. *Acc. Chem. Res.*, **47**, 1446–1454.
 30. Kowal, E.A., Ganguly, M., Pallan, P.S., Marky, L.A., Gold, B., Egli, M. and Stone, M.P. (2011) Altering the electrostatic potential in the major groove: thermodynamic and structural characterization of 7-deaza-2'-deoxyadenosine:dT base pairing in DNA. *J. Phys. Chem. B*, **115**, 13925–13934.
 31. Ganguly, M., Wang, R.W., Marky, L.A. and Gold, B. (2010) Thermodynamic characterization of DNA with 3-deazaadenine and 3-methyl-3-deazaadenine substitutions: the effect of placing a hydrophobic group in the minor groove of DNA. *J. Phys. Chem. B*, **114**, 7656–7661.
 32. Lever, C., Li, X., Cosstick, R., Edell, S. and Brown, T. (1993) Thermodynamic stability and drug-binding properties of oligodeoxyribonucleotide duplexes containing 3-deazaadenine:thymine base pairs. *Nucleic Acids Res.*, **21**, 1743–1746.
 33. Cosstick, R., Li, X., Tuli, D.K., Williams, D.M., Connolly, B.A. and Newman, P.C. (1990) Molecular recognition in the minor groove of the DNA helix. Studies on the synthesis of oligonucleotides and polynucleotides containing 3-deaza-2'-deoxyadenosine. Interaction of the oligonucleotides with the restriction endonuclease EcoRV. *Nucleic Acids Res.*, **18**, 4771–4778.
 34. Minakawa, N., Kawano, Y., Murata, S., Inoue, N. and Matsuda, A. (2008) Oligodeoxynucleotides containing 3-bromo-3-deazaadenine and 7-bromo-7-deazaadenine 2'-deoxynucleosides as chemical probes to investigate DNA-protein interactions. *ChemBioChem*, **9**, 464–470.
 35. Seela, F. and Grein, T. (1992) 7-Deaza-2'-deoxyadenosine and 3-deaza-2'-deoxyadenosine replacing dA within d(A6)-tracts: differential bending at 3'- and 5'-junctions of d(A6).d(T6) and B-DNA. *Nucleic Acids Res.*, **20**, 2297–2306.
 36. Grein, T., Lampe, S., Mersmann, K., Rosemeyer, H., Thomas, H. and Seela, F. (1994) 3-Deazapurines and 7-deazapurines - duplex stability of oligonucleotides containing modified adenine or guanine bases. *Bioorg. Med. Chem. Lett.*, **4**, 971–976.
 37. Ganguly, M., Wang, F., Kaushik, M., Stone, M.P., Marky, L.A. and Gold, B. (2007) A study of 7-deaza-2'-deoxyguanosine 2'-deoxycytidine base pairing in DNA. *Nucleic Acids Res.*, **35**, 6181–6195.
 38. Ganguly, M., Wang, R.-W., Marky, L.A. and Gold, B. (2009) Introduction of cationic charge into DNA near the major groove edge of a guanine cytosine base pair: characterization of oligodeoxynucleotides substituted with 7-aminomethyl-7-deaza-2'-deoxyguanosine. *J. Am. Chem. Soc.*, **131**, 12068–12069.
 39. Hakimelahi, G.H., Proba, Z.A. and Ogilvie, K.K. (1981) Nitrate ion as catalyst for selective silylations of nucleosides. *Tetrahedron Lett.*, **22**, 4775–4778.
 40. Serebryany, V. and Beigelman, L. (2002) An efficient preparation of protected ribonucleosides for phosphoramidite RNA synthesis. *Tetrahedron Lett.*, **43**, 1983–1985.
 41. Micura, R. (2002) Small interfering RNAs and their chemical synthesis. *Angew. Chem. Int. Ed.*, **41**, 2265–2269.
 42. Glasner, H., Riml, C., Micura, R. and Breuker, K. (2017) Label-free, direct localization and relative quantitation of the RNA nucleobase methylations m⁶A, m⁵C, m³U, and m⁵U by top-down mass spectrometry. *Nucleic Acids Res.*, **45**, 8014–8025.
 43. Crey-Desbiolles, C., Lhomme, J., Dumy, P. and Kotera, M. (2004) 3-Nitro-3-deaza-2'-deoxyadenosine as a versatile photocleavable 2'-deoxyadenosine mimic. *J. Am. Chem. Soc.*, **126**, 9532–9533.
 44. Maiti, M., Persoons, L., Andrei, G., Snoeck, R., Balzarini, J. and Herdewijn, P. (2013) Synthesis and anti-herpetic activity of phosphoramidate ProTides. *Chem. Med. Chem.*, **8**, 985–993.
 45. Saenger, W. (1984) In: *Principles of Nucleic Acid Structure*, Springer, Berlin.
 46. Majlessi, M. and Becker, M.M. (2008) Formation of the double helix: a mutational study. *Nucleic Acids Res.*, **36**, 2981–2989.
 47. Seio, K., Sasami, T., Ohkubo, A., Ando, K. and Sekine, M. (2007) Highly selective recognition of cytosine over uracil and adenine by a guanine analogue, 2-N-acetyl-3-deazaguanine, in 2'-O-methyl-RNA/RNA and DNA duplexes. *J. Am. Chem. Soc.*, **129**, 1026–1027.
 48. Seio, K., Sasami, T., Tawarada, R. and Sekine, M. (2006) Synthesis of 2'-O-methyl-RNAs incorporating a 3-deazaguanine, and UV melting and computational studies on its hybridization properties. *Nucleic Acids Res.*, **34**, 4324–4334.
 49. Marky, L.A. and Breslauer, K.J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*, **26**, 1601–1620.
 50. Petersheim, M. and Turner, D.H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry*, **22**, 256–263.
 51. Siraiwa, S., Suzuki, A., Katoh, R. and Saito, Y. (2016) Design and synthesis of a novel fluorescent benzofurimidazo[4, 5-c]quinoline nucleoside for monitoring base-pair-induced protonation with cytosine: distinguishing cytosine via changes in the intensity and wavelength of fluorescence. *Org. Biomol. Chem.*, **14**, 3934–3942.
 52. Puffer, B., Kreutz, C., Rieder, U., Ebert, M.O., Konrat, R. and Micura, R. (2009) 5-Fluoro pyrimidines: labels to probe DNA and RNA secondary structures by 1D ¹⁹F NMR spectroscopy. *Nucleic Acids Res.*, **37**, 7728–7740.
 53. Strebitzer, E., Rangadurai, A., Plangger, R., Kremser, J., Juen, M.A., Tollinger, M., Al-Hashimi, H.M. and Kreutz, C. (2018) 5-Oxyacetic acid modification destabilizes double helical stem structures and favors anionic Watson-Crick like cmo⁵U-G base pairs. *Chem. Eur. J.*, **24**, 18903–18906.
 54. Hwang, T.-L., van Zijl, P.C.M. and Mori, S. (1998) Accurate Quantitation of Water-amide Proton Exchange Rates Using the Phase-Modulated CLEAN Chemical EXchange (CLEANEX-PM) Approach with a Fast-HSQC (FHSQC) Detection Scheme. *J. Biomol. NMR*, **11**, 221–226.
 55. Guéron, M. and Leroy, J.L. (1995) Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.*, **261**, 383–413.
 56. Olieric, V., Rieder, U., Lang, K., Serganov, A., Schulze-Briese, C., Micura, R., Dumas, P. and Ennifar, E., 2009, A fast selenium derivatization strategy for crystallization and phasing of RNA structures. *RNA*, **15**, 707–715.
 57. Correll, C.C., Wool, I.G. and Munishkin, A. (1999) The two faces of the Escherichia coli 23 S rRNA sarcin/ricin domain: the structure at 1.11 Å resolution. *J. Mol. Biol.*, **292**, 275–287.
 58. Sponer, J., Leszczynski, J. and Hobza, P. (2001) Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. *Biopolymers*, **61**, 3–31.
 59. Krishnamurthy, R. (2009) A search for structural alternatives of RNA. *J. Mex. Chem. Soc.*, **53**, 22–33.
 60. Thaplyal, P. and Bevilacqua, P.C. (2014) Experimental approaches for measuring pK_a's in RNA and DNA. *Methods Enzymol.*, **549**, 189–219.
 61. Acharya, P., Cheruku, P., Chatterjee, S., Acharya, S. and Chattopadhyaya, J. (2014) The measurement of nucleobase pK_a of the model mononucleotides shows why RNA-RNA duplex is more stable than DNA-DNA duplex. *J. Am. Chem. Soc.*, **126**, 2862–2869.
 62. Gilli, P., Pretto, L., Bertolasi, V. and Gilli, G. (2009) Predicting hydrogen-bond strengths from acid–base molecular properties. The pK_a slide rule: toward the solution of a long-lasting problem. *Acc. Chem. Res.*, **42**, 33–44.
 63. Wahl, M.C., Rao, S.T. and Sundaralingam, M. (1996) The structure of r(UUCGCG) has a 5'-UU-overhang exhibiting Hoogsteen-like trans U • U base pairs. *Nat. Struct. Biol.*, **3**, 24–31.
 64. Adamiak, D.A., Milecki, J., Adamiak, R.W. and Rypniewski, W. (2010) The hydration and unusual hydrogen bonding in the crystal structure of an RNA duplex containing alternating CG base pairs. *New J. Chem.*, **34**, 903–909.

65. Micura,R. and Höbartner,C. (2020) Fundamental studies of functional nucleic acids: aptamers, riboswitches, ribozymes and DNazymes. *Chem. Soc. Rev.*, **49**, 7331–7353.
66. Scheitl,C.P.M, Ghaem Maghami,M., Lenz,A.K. and Höbartner,C. (2020) Site-specific RNA methylation by a methyltransferase ribozyme. *Nature*, **587**, 663–667.
67. Gebetsberger,J. and Micura,R. (2017) Unwinding the twister ribozyme: from structure to mechanism. *Wiley Interdiscip. Rev. RNA*, **8**, e1402.
68. Chen,H., Egger,M., Xu,X., Flemmich,L., Krasheninina,O., Sun,A., Micura,R. and Ren,A. (2020) Structural distinctions between NAD⁺ riboswitch domains 1 and 2 determine differential folding and ligand binding. *Nucleic Acids Res.*, **48**, 12394–12406.
69. Fuchs,E., Falschlunger,C., Micura,R. and Breuker,K. (2019) The effect of adenine protonation on RNA phosphodiester backbone bond cleavage elucidated by deaza-nucleobase modifications and mass spectrometry. *Nucleic Acids Res.*, **47**, 7223–7234.
70. Erlacher,M.D., Chirkova,A., Voegelé,P. and Polacek,N. (2011) Generation of chemically engineered ribosomes for atomic mutagenesis studies on protein biosynthesis. *Nat. Protoc.*, **6**, 580–592.
71. Kabsch,W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.*, **D66**, 133–144.
72. Vagin,A. and Teplyakov,A. (2010) Molecular replacement with MOLREP. *Acta Cryst.*, **D66**, 22–25.
73. Afonine,P.V., Grosse-Kunstleve,R.W., Echols,N., Head,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst.*, **D68**, 352–367.
74. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Cryst.*, **D66**, 486–501.