m1A and m1G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs

Huiqing Zhou1, Isaac J Kimsey1, Evgenia N Nikolova2, Bharathwaj Sathyamoorthy1,8, Gianmarc Grazioli3, James McSally3, Tianyu Bai1, Christoph H Wunderlich4, Christoph Kreutz5,6, Ioan Andricioaei3 & Hashim M Al-Hashimi1,7

The B-DNA double helix can dynamically accommodate G-C and A-T base pairs in either Watson–Crick or Hoogsteen configurations. Here, we show that G-C+ (in which + indicates protonation) and A-U Hoogsteen base pairs are strongly disfavored in A-RNA. As a result, N1-methyladenosine and N1-methylguanosine, which occur in DNA as a form of alkylation damage and in RNA as post-transcriptional modifications, have dramatically different consequences. Whereas they create G-C+ and A-T Hoogsteen base pairs in duplex DNA, thereby maintaining the structural integrity of the double helix, they block base-pairing and induce local duplex melting in RNA. These observations provide a mechanism for disrupting RNA structure through post-transcriptional modifications. The different propensities to form Hoogsteen base pairs in B-DNA and A-RNA may help cells meet the opposing requirements of maintaining genome stability, on the one hand, and of dynamically modulating the structure of the epitranscriptome, on the other.

The Watson–Crick (WC) double helix is the most common structural element in RNA and the dominant structure of genomic DNA. It provides the basis for templated replication, transcription, and translation, and it also serves as a scaffold that defines the 3D structure of DNA, RNA, and their protein complexes. The canonical double helices formed by RNA (A form) and DNA (B form) differ in several important respects (Fig. 1a). In B-form DNA (B-DNA), the five-membered deoxyribose ring is flexible and favors the C2'-endo sugar pucker (Fig. 1a). In contrast, owing to the sugar 2'–OH group, the sugar in A-RNA is more rigid and adopts an alternative C3'–endo conformation1,2 (Fig. 1a). This conformation in turn brings the oxygen atoms (O5' and O3') adjoining sequential nucleotides closer, thus effectively compressing and rigidifying the A-form helix, widening its helical diameter, and displacing base pairs away from the helical axis1,3 (Fig. 1a). In addition, B-DNA and A-RNA differ considerably with respect to their deformability, and B-DNA is generally more flexible4. The higher bendability of B-DNA than A-RNA is fundamentally important in many biochemical processes including the tight compaction of genome within the nucleus in higher-order organisms.

Recently, NMR studies have uncovered a new dynamic property in canonical B-DNA; WC dG–dC and dA–dT base pairs exist in a dynamic equilibrium with alternative Hoogsteen (HG) base pairs5,6. An HG base pair is created by rotation of a WC purine base ~180° around the glycosidic bond, so that it adopts a syn rather than anti conformation2 (Fig. 1b). The two bases are also brought ~2.0–2.5 Å closer, thus allowing formation of a unique set of hydrogen bonds (Fig. 1b). HG base pairs exist transiently (with typical lifetimes of 0.1–1 ms) and in low abundance (with populations typically <3%) in naked canonical B-DNA6,7. However, dA–dT and dG–dC+ HG base pairs can become the dominant configuration (reviewed in ref. 8) in DNA–protein9 and DNA–small molecule10 complexes, in which these base pairs contribute to DNA recognition; in damaged nucleotides, in which they contribute to damage accommodation and repair11–13; and in the active sites of translesion synthesis polymerases that use HG pairing to bypass damage during DNA replication14. Purine-purine HG base pairs have also been shown to play important roles in RNA replication errors and in DNA-damage accommodation and repair15,16.

Here, we set out to study WC–HG dynamics in canonical A-RNA duplexes. We show that, unlike the canonical B-DNA double helix, rA–rU and rG–rC+ HG base pairs are strongly disfavored in A-RNA duplexes. As a result, whereas the DNA double helix can absorb damaged nucleotides (such as N1-methyldeoxyadenosine (m1dA) and N1-methyldeoxyguanosine (m1dG)) that are incapable of forming WC base pairs by forming HG base pairs, the same methyl marks, N1-methyladenosine (m1RA) and N1-methylguanosine (m1RG), acting as a post-transcriptional modification in RNA, block base-pairing altogether. These phenomena provide a direct mechanism for potently modulating the structure of the epitranscriptome. Our results indicate

1Department of Biochemistry, Duke University School of Medicine, Durham, North Carolina, USA. 2Department of Molecular Biology, The Scripps Research Institute, La Jolla, California USA. 3Department of Chemistry, University of California Irvine, Irvine, California USA. 4Bachem Americas, Torrance, California USA. 5Institute of Organic Chemistry, University of Innsbruck, Innsbruck, Austria. 6Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria. 7Department of Chemistry, Duke University, Durham, North Carolina, USA. 8Present address: Department of Chemistry, Indian Institute of Science Education and Research Bhopal, Bhopal, India. Correspondence should be addressed to H.M.A. (hashim.al.hashimi@duke.edu) or I.A. (andricio@uci.edu).

Received 26 May; accepted 5 July; published online 1 August 2016; doi:10.1038/nsmb.3270
that HG-dependent DNA biochemical transactions may not be as readily supported in RNA duplexes, and they identify a unique dynamic property in B-DNA that may help enhance its ability to function as the repository of genetic information.

RESULTS

Absence of conformational exchange in A-RNA

We used NMR spin relaxation in the rotating frame ($R_{1p}^{\chi}$) to examine whether WC base pairs in A-RNA duplexes transiently adopt HG base pairs, as in B-DNA. A dynamic equilibrium between a dominant ground state (GS) and short-lived low-abundance ‘excited state’ (ES) conformation can lead to line-broadening of NMR resonances if the conformational exchange occurs on the microsecond-to-millisecond timescale. The $R_{1p}$ experiment measures this line-broadening contribution ($R_{1p}$) to the transverse relaxation rate ($R_2$) during a relaxation period in which a continuous radiofrequency (RF) field is applied with variable power ($\omega^2$). The resulting dependence of $R_2 + R_{1p}$ on $\omega_{gs}$ and $\omega_{rg}$ referred to as relaxation dispersion (RD), can be fitted to the Bloch–McConnell equations describing n-site exchange to extract exchange parameters of interest, including the population of the ES ($p_E$), the rate constant for conformational exchange ($k_{ex} = k_{forward} + k_{backward}$), and the difference between the chemical shifts of the ES and GS ($\Delta \delta = \delta_{ES} - \delta_{GS}$).

To date, RD studies have provided evidence for microsecond-to-millisecond conformational exchange in noncoding RNAs involving localized changes in secondary structure in and around noncanonical motifs (reviewed in ref. 18). The RD contributions from such chemical-exchange processes can mask the ability to detect WC $\Rightarrow$ HG exchange. To home in on WC $\Rightarrow$ HG exchange in A-RNA, we carried out $^{13}$C and $^{15}$N $R_{1p}$ RD experiments on an RNA duplex (hp-A$_6$-RNA) capped by a stabilizing apical loop lacking noncanonical motifs and containing the same sequence (A$_6$-DNA) for which we previously reported transient HG base pairs in B-DNA ($^6$).

To broaden the search for WC $\Rightarrow$ HG exchange in A-RNA duplexes, we carried out additional RD measurements over a wide range of conditions (pH 5.4–8.4 and 5–35 °C) for another eight rG-rC and four rA-rU WC base pairs, under low pH conditions (pH 5.4) that allow optimal RD detection of WC $\Rightarrow$ HG exchange in B-DNA ($^6$), after the temperature is increased (to 35 °C), and in the presence of 4 mM Mg$^{2+}$ (at pH 6.8 and 5 or 25 °C) ($^{1d}$).

Figure 1. Absence of detectable WC $\Rightarrow$ HG exchange in A-RNA by NMR relaxation dispersion. (a) Comparison of A-form RNA (violet) and B-form (blue) DNA double helices. (b) WC and HG base pairs in dynamic equilibrium in B-DNA. Sites used for RD measurements are in orange. (c) A$_6$-DNA and hp-A$_6$-RNA duplexes with base pairs targeted in RD measurements highlighted. (d) Off-resonance RD profiles showing $R_2 + R_{1p}$ as a function of spin-lock offset ($\Omega \times 2 \pi^{-1}$ Hz, where $\Omega = \omega_{gs} - \omega_{rg}$) and power ($\omega_{gs} \times 2 \pi^{-1}$ Hz, in insets). Error bars, experimental uncertainty (s.d.) estimated from monoeponential fitting of $n = 10$ (A$_6$-DNA) and $n = 6$ (hp-A$_6$-RNA) independently measured peak intensities by using a Monte Carlo–based method (Online Methods). The solid line represents a fit to two-state exchange.
TAR-UUCG(GU) that cannot undergo secondary-structure chemical exchange\textsuperscript{25} (Fig. 2a and Supplementary Fig. 2a). In all cases, we did not detect any signs of RD (Fig. 2b and Supplementary Fig. 2b). These results, together with those from our previous studies\textsuperscript{25–27}, reporting flat RD profiles for RNA WC base pairs near noncanonical motifs and mismatches (wtTAR in Supplementary Fig. 2a and the P5abc subdomain of the Tetrahymena thermophila intron RNA) and for the reverse-wobble rG\textsuperscript{979}–rU mispairs in apical loops\textsuperscript{28}, stand in striking contrast to results for canonical duplex DNA, in which we have robustly observed WC \textasciitilde HG exchange in all 35 dA–dT and dG–dC base pairs examined to date in a wide variety of positional and sequence contexts in eight different duplexes with varying lengths and stabilities\textsuperscript{6,7}.

The lack of detectable WC \textasciitilde HG exchange in A-RNA could in principle result from small differences between the WC and HG NMR chemical shifts (\(\Delta\omega<0.5\) p.p.m. for carbon chemical shifts). However, on the basis of density functional theory calculations (DFT)\textsuperscript{6,29} and a survey of syn purine-base chemical shifts in the Biological Magnetic Resonance Data Bank\textsuperscript{30}, it is highly unlikely that such a large transformation in base-pairing would result in such small changes in chemical shifts for the different sugar (C1\textsuperscript{′}) and base (C8, C6 and N1/N3) sites targeted for RD measurements (Supplementary Note). The absence of RD is unlikely to be a result of the exchange rate falling outside the detection limits of the RD experiment, given that we observed flat profiles over a wide range of temperatures and pH conditions (Fig. 2b and Supplementary Fig. 2b) known to substantially alter the rate of WC \textasciitilde HG exchange in B-DNA\textsuperscript{6,7}.

A more likely explanation is that HG base pairs are energetically disfavored in A-RNA duplexes and have an abundance that falls below the detection threshold of the RD experiment (population \textless{}0.01%). Indeed, a survey of X-ray structures of RNA duplexes in the Protein Data Bank (PDB)\textsuperscript{31} did not identify a single rG–rC\textsuperscript{′} or rA–rU HG base pair within continuous A-RNA duplexes out of a total of 123,935 rG–rC and rA–rU base pairs (Online Methods); in sharp contrast, a similar survey conducted recently on B-DNA duplexes\textsuperscript{32} has identified 54 dG–dC\textsuperscript{′} or dA–dU HG base pairs out of a much smaller set of 51,485 base pairs. The current survey of HG base pairs in RNA identified a single rA–rU HG base pair (for example, PDB 1GID\textsuperscript{33}) within an RNA duplex that fell well outside the A-form structural context, being surrounded by a bulge and internal loop. The survey

did identify several examples of long-range rG–rC\textsuperscript{′} and rA–rU HG base pairs forming tertiary contacts; these included rG–rC\textsuperscript{′} and rA–rU HG base pairs in triplexes and reverse rA–rU HG base pairs within duplexes typically near rG–rA mismatches, where purines adopt anti rather than syn conformation, as well as several examples of HG mispairs in A-RNA duplexes (for example, rG\textsuperscript{979}–rU and rG\textsuperscript{981}–rA) (Supplementary Note).

\textbf{m\textsuperscript{1}A- and m\textsuperscript{1}G-modified A-RNA} 

If rG–rC\textsuperscript{′} and rA–rU HG base pairs are indeed thermodynamically disfavored in A-RNA, they should prove more difficult to trap by using chemical modifications known to stabilize dG–dC\textsuperscript{′} and dA–dT HG base pairs in B-DNA\textsuperscript{6}. We therefore examined whether HG base pairs could be stably trapped in A-RNA duplexes by using m\textsuperscript{1}A and m\textsuperscript{1}G. These modified bases block WC pairing because of steric collisions with the methyl group and because the methylation prevents one of the WC hydrogen bonds from forming (Fig. 3a). Both m\textsuperscript{1}dA and m\textsuperscript{1}dG occur in DNA after alkylhydation damage\textsuperscript{12,13}. In B-DNA, m\textsuperscript{1}dA and m\textsuperscript{1}dG are accommodated as m\textsuperscript{1}dA–dT and m\textsuperscript{1}dG–dC\textsuperscript{′} HG base pairs\textsuperscript{6,11,34} (Fig. 3a), which can in turn be recognized and repaired by damage-repair enzymes\textsuperscript{12,13} and m\textsuperscript{1}rA and m\textsuperscript{1}rG can also occur as a form of alklylation damage in RNA, but they are also highly conserved post-transcriptional modifications in rRNAs and rRNAs, and they play critical structural and functional roles, often by blocking WC base-pairing\textsuperscript{35–39}. m\textsuperscript{1}rG and m\textsuperscript{1}rA have been shown to induce duplex-to-hairpin transitions in palindromic RNA sequences in which the modified base favors an unpaired conformation within neighboring WC base pairs, as evidenced by HG-specific chemical shifts, NOESY cross-peaks, and imino resonances (Fig. 3a). In contrast, we did not observe any NMR evidence for HG base pairs or syn purine bases in the corresponding A\textsubscript{6}-RNA duplex containing m\textsuperscript{1}rA or m\textsuperscript{1}rG\textsubscript{10} (Fig. 3b–e and Supplementary Fig. 3). These observations were in spite of the highly similar thermodynamic stabilities of A\textsubscript{6}-DNA and A\textsubscript{6}-RNA duplexes. Instead, the rA–C1\textsuperscript{′} chemical
shifts fell in a region consistent with A-form helical residues (Fig. 3c and Supplementary Note). We also observed continuous NOE distance-based connectivity between H8 of the m^1rA and H1' of its preceding residue, thus suggesting an anti conformation for the purine base (Online Methods and Supplementary Fig. 3). These data, together with the absence of strong H1'-H8 NOEs expected for syn base (Fig. 3d) and imino and amino resonances indicative of hydrogen-bonding (Fig. 3e), suggest that in A_C-RNA, m^1rA adopts a predominantly unpaired anti conformation, although we cannot rule out transient formation of syn base conformations. The resonances belonging to m^1rG in A_C-RNA were broadened out of detection, thus suggesting extensive conformational exchange at the microsecond-to-millisecond

Figure 3  m^1A and m^1G disrupt A-RNA structure and do not form HG base pairs. (a) N^2-methylated purines trap HG base pairs in B-DNA. NMR chemical-shift probes of HG base pairs are in orange, and those of the purine-methylation state are in cyan. Arrows indicate characteristic HG NOE cross-peaks. (b) Duplexes containing m^1A or m^1G (turquoise circles). Syn or anti purines deduced by NMR are shown as open and filled letters, respectively. HG and partially melted base pairs, as deduced by NMR, are indicated by open and dashed lines, respectively. Residues showing significant chemical-shift perturbations or line-broadening due to m^1A or m^1G are orange and gray, respectively. (c) m^1A- or m^1G-induced purine-C1' chemical-shift perturbations (Δω = ω_ modified − ω_unmodified) in A-RNA (violet) and B-DNA (blue). Shown for comparison are Δω = δ_arginine − δ_lysine measured for transient dA-dT HG base pairs by RD in unmodified DNA duplexes (error bars, one s.d. estimated from monoeponential fitting of n = 10 independently measured peak intensities) and computed for adenine residues by using DFT (Online Methods). (d) NOESY H1'-H8 cross-peaks showing syn purine bases in B-DNA but not A-RNA. Shown for reference is the cytosine base H5-H6 NOE with interatomic distance ≈ 2.5 Å. (e) 1D ^1H spectra showing the imino and amino resonances expected for HG-type hydrogen bonds at low temperatures in A_C-DNA/m^1A (5 °C) and A_C-DNA/m^1G (15 °C), but not in A_C-DNA/m^1A (5 °C) and A_C-DNA/m^1G (15 °C). (f) Example showing m^1G-induced loss of a WC imino resonance (circled) in A_C-RNA but not A_C-DNA in 2D NMR spectra. (g) Example of downfield-shifted carbon chemical shifts induced by m^1A. (h) Free energy (ΔG) and enthalpy (ΔH) destabilization due to m^1A and m^1G in DNA (blue) and RNA (violet) duplexes, as measured by UV melting experiments. Error bars, s.d. (n = 3 independent measurements) (details in Online Methods and Supplementary Table 3).
timescale with no NMR evidence for HG pairing, given the absence of downfield-shifted rC-H4 (Supplementary Fig. 1b). However, we cannot exclude microsecond-to-millisecond exchange between syn and anti conformations for the m^1rG base, because the resonances were broadened out of detection.

Compared with A6-DNA, m^1A and m^1G also induced more pronounced structural perturbations in A6-RNA (Fig. 3b and Supplementary Table 2). We did not observe some of the imino resonances belonging to WC base pairs neighboring the modified site (Fig. 3f), thus suggesting a loss of hydrogen bonds and the melting of these base pairs. The modifications also induced more extensive chemical-shift perturbations (Fig. 3g, Supplementary Fig. 1b; orange in Fig. 3b) and line broadening (Supplementary Fig. 1b; gray in Fig. 3b) in the sugar and base resonances that extended to the partner strand. The direction of the perturbations was consistent with deviations from a helical conformation (Online Methods and Supplementary Fig. 1b). The perturbations were particularly pronounced for m^1rG, which broadened all imino resonances out of detection at 35 °C, a result consistent with substantial melting of the entire duplex (Supplementary Fig. 1c). Thus, HG base pairs are so sufficiently disfavored in A-RNA that m^1rA and m^1rG prefer to adopt predominantly nonhelical conformations that disrupt the duplex structure.

We obtained similar results in GC-rich (gc^m^1A) and scrambled (A6^m^1A)^pf (B.S., H.Z., Y. Xue and H.M.A., unpublished data) duplexes, in which m^1A consistently formed HG base pairs or adopted a syn conformation in B-DNA but not in A-RNA (Fig. 3b,d and Supplementary Figs. 4 and 5), and the modification perturbed the structure of A-RNA more than that of B-DNA (Fig. 3b,g). The structural perturbations induced by m^1rA varied with sequence and either were distributed across many WC base pairs (A6-RNA and A2-RNA) or were more severe but localized to the modified and partner base (gc-RNA) (Fig. 3b and Supplementary Figs. 1b and 4). In all cases, we did not observe any evidence that m^1rA or m^1rG induced duplex-to-hairpin transitions, on the basis of spectral overlays with the unmodified counterparts (Supplementary Figs. 1b and 4).

We corroborated the more potent destabilization of A-RNA than B-DNA duplexes by m^1A and m^1G, by using UV melting experiments. m^1DA destabilized A6-DNA, A2-DNA, and gc-DNA duplexes by free energy (ΔΔG) = 1.8–3.4 kcal mol−1 (Fig. 3h), a result in good agreement with the relative stability of transient HG base pairs measured by NMR RD (2.1–4.3 kcal mol−1) and prior UV-melting studies of m^1DA-containing DNA duplexes (~2 kcal mol−1)44. By comparison, m^1A and m^1G destabilized the corresponding A-RNA duplexes by a larger amount ΔΔG = 4.3–6.5 kcal mol−1. Interestingly, this greater destabilization was comparable to the relative stability of the base-opened state45. This result suggests that in A-RNA, the modification results in a conformation similar to that of the base-opened state, in agreement with the NMR evidence for local melting. We observed greater destabilization (by ~1.1–4.7 kcal mol−1) of A-RNA than B-DNA across different duplex and hairpin contexts in the presence or absence of Mg^2+ and with principally enthalpic destabilization (Fig. 3h and Supplementary Table 3). The potent m^1rA-induced destabilization of duplex RNA is notable, given the recent studies showing that it is a dynamic mRNA modification with roles in post-transcriptional gene regulation42,43. In comparison, the other well-studied mRNA modification N6^-methyladenosine (mA) which affects mRNA localization, stability, translation, and splicing, destabilizes A-RNA by only 0.5–1.7 kcal mol−1 (ref. 47).

The more potent m^1A and m^1G destabilization of A-RNA than B-DNA is unlikely to be due to differences in steric contacts involving the methyl group in an HG base-pair configuration (Supplementary Note). Whereas the positive charge on m^1rA may affect stacking and hydrogen-bonding interactions, we also observed pronounced destabilization with the neutral m^1rG, and the m^1A destabilization was greater for A-RNA than B-DNA (Fig. 3h). Instead, the greater destabilization observed in A-RNA is probably due to the higher energetic cost of forming HG base pairs in A-RNA than B-DNA.
Why are HG base pairs disfavored in A-RNA?

Why are HG base pairs disfavored in A-RNA? The HG base pair could in principle be disfavored in RNA because of the sugar 2′-OH at the purine residue. The 2′-OH helps bias the sugar pucker toward the C3′-endo conformation (Fig. 1a), owing to unfavorable steric contacts between O2′ and O3′ and electronic effects involving the 2′-OH group. This conformation in turn disfavors the syn purine base conformation even in nucleosides and single-stranded polynucleotides, because of unfavorable base-sugar steric contacts (N3-H3′ and N3-O4′). The syn purine base conformation may also destabilize water-bridged interactions involving the 2′-OH and N3 of the anti purine base. To examine whether the presence of a 2′-OH group on the ribose moiety of the flipping purine base might be sufficient to suppress WC = HG exchange, we carried out Rd experiments on site- or strand-specifically labeled A6-DNA duplexes containing a single ribonucleotide, rA16 or rG10 (Supplementary Fig. 6a,b and Online Methods). These RD measurements were also of interest, given that single ribonucleotides are frequently incorporated in DNA during replication and can have important biological consequences through mechanisms that are not fully understood.

Both rA16 and rG10 formed the expected rA16-dT9 and rG10-dC15 WC base pairs and exhibited RD consistent with WC = HG exchange (Fig. 4a). The larger R2 value in A6-DNA-A16-C8 than in A6-DNA probably reflects decreased flexibility in rA16 (Fig. 4a). The lower R2 contribution observed for the rA16- and rG10-substituted samples relative to the unmodified DNA duplex can be attributed to an exchange rate approximately four-fold faster (kex = 2.325 s⁻¹ versus 595 s⁻¹) in the case of rA16 and a combination of a slightly smaller Δν (1.8 versus 2.1 p.p.m.) and a transient HG population (0.8% versus 1.3%) in the case of rG10 (Fig. 4a, Supplementary Table 4 and Supplementary Fig. 6c). Neither rA16 nor rG10 significantly affected the abundance of the transient HG base pairs relative to the unmodified A6-DNA duplex (Fig. 4a and Supplementary Table 4), thus indicating that the purine sugar 2′-OH group alone cannot account for the lack of observable WC = HG exchange in A-RNA duplexes. We confirmed these findings by analyzing A6-DNA duplexes containing N5′-methylated single ribonucleotide, m5′rA16 or m5′rG10. In both cases, we observed stably formed m5′rA16-dT9 and m5′rG10-dC15° HG base pairs (Supplementary Fig. 6b). These data suggest that the destabilization of HG base pairs requires the broader A-form RNA helical context.

Next, we examined whether there were unique steric clashes that might disfavor syn purine bases within the compact A-RNA helix context and that are absent in the more capacious B-form DNA helix. Indeed, flipping the purine base around the glycosidic χ-angle through a range of angles (160°–200°) that span syn base conformations found in RNA helices (Supplementary Note) resulted in greater steric clashes in A-RNA than in B-DNA. The additional base-sugar (N3-H3′ and N3-O4′) and base-backbone (N3-O5′) clashes observed in A-RNA arose as a result of both the C3′-endo sugar pucker and the unique phosphodiester backbone conformation at the syn purine residue (Fig. 4b).

To further examine the energetics of the WC = HG transition, we carried out biased molecular dynamics (MD) simulations on the A6-DNA duplex and hp-A6-RNA hairpin, as well as a 3′→5′ inverted sequence of the hp-A6-RNA hairpin. We applied a bias on dA16 or rA16 starting in a WC base pair configuration to force purine base flipping and a transition to a target HG configuration (Online Methods and Supplementary Movies 1 and 2). The computed mean interaction energy (averaged over an ensemble of biased trajectories) as a function of the χ-angle along the WC = HG transition (Online Methods) revealed a clear two-state transition in the case of B-DNA, in agreement with previous results, whereas in the case of A-RNA the resultant HG base pair was markedly destabilized in comparison to its WC base pair counterpart; for A-RNA, the energy profile in the syn region had much higher relative energies than in the case of DNA (Fig. 4c). In accord with this energetic destabilization, the simulations revealed that flipping the purine base in A-RNA is accompanied by major structural disruption of the surrounding base pairs (Supplementary Movie 2), in agreement with m1rA-induced NMR chemical-shift perturbations, which were more pronounced for residues 3′ to the modified nucleotide (Supplementary Fig. 1b). The extent of the disruption for the neighboring base pair was far less pronounced in B-DNA (Supplementary Movie 1).

We corroborated these findings by using unbiased MD simulations, which began with an HG base pair embedded in various duplex and hairpin contexts (Online Methods). The HG hydrogen-bonding remained stable during the course of the simulation in the case of B-DNA, B-DNA containing a single rA, and B-DNA containing m1rA (Supplementary Table 5 and Supplementary Movies 3–5). In contrast, for A-RNA, we observed strong disruption of the N7→H3-N3 HG hydrogen bond between A16 and U9 in cases of the hp-A6-RNA hairpin (Supplementary Table 5). In ~35% of the trials, in the case of 3′→5′ sequence hp-A6-RNA, the HG base pair transitioned rapidly after equilibration back to a WC base pair. Strikingly, in the case of m1rA embedded in A-RNA, the HG base pair caused melting of the A-form helix (Fig. 4d and Supplementary Movie 6). Together, these results indicate that HG base pairs are disfavored in the more compact
A-RNA helix, owing to steric contacts that are difficult to alleviate without substantially perturbing the A-form-helix structure.

**DISCUSSION**

Duplex B-DNA can stably accommodate dA-dT and dG-dC+ HG base pairs, which can in turn play roles in sequence-specific DNA recognition, damage induction and repair, and DNA replication. In contrast, our results indicate that rA-rU and rG-rC+ HG base pairs are so unstable in the more compressed A-RNA that melting is preferred over the HG base-pair conformation. It remains to be seen whether the greater instability of HG base pairs in A-RNA than in B-DNA extends to purine-purine HG mispairs (Supplementary Note), which play important roles in replication4,5,5 and translation errors5,5, mismatch repair, and translational reprogramming5,5,5,5,7.

The markedly different stability of the A-T/U and G-C+ HG base pairs in RNA and DNA duplexes provides a basis for achieving opposing functions at the genome and transcriptome levels (Fig. 5). If DNA did not have a capacity to form HG base pairs, and instead behaved similarly to RNA, lesions, such as m1dA and m1dG, that block canonical WC base pairing could greatly destabilize the double helix and potentially cause genomic instability (Fig. 5). The ability to form HG base pairs therefore endows DNA with an additional layer of chemical stability than is present in its RNA counterpart; this stability goes beyond resistance to hydrolysis, owing to the absence of the sugar 2′-OH group. The greater instability of HG base pairs in A-RNA gives rise to a chemical switch in the form of m1rA and m1rG that can potently modulate RNA structure (Fig. 5). Although it has long been recognized that m1A and m1G can modulate the structure and function of tRNA, rRNA, and other noncoding RNA5,5,37–39,5,8, this functionality hinges on the unique instability of HG base pairs in A-RNA uncovered in this work.

For example, m1rA9 has been shown to stabilize the native structures of human mitochondrial tRNAs by blocking helical rA-rU WC base pairs that would otherwise stabilize alternative secondary structures5 (Fig. 6a). Likewise, m1rG37 next to the anticodon loop, which is highly conserved in most tRNAs that read the CNN codon, has been shown to prevent +1 frameshifting by blocking base-pairing between G37 and the first rC in the codon sequence5,39 (Fig. 6b). If RNA behaved similarly to DNA, such post-transcriptional modifications would simply create HG base pairs and fail to block base-pairing and to have their proper functional consequences (Fig. 5).

In eukaryotic cells, including yeast and mammals, m1rA has recently been shown to be a reversible mRNA modification that responds to changes in physiological conditions5,2,5,3. It is enriched in the 5′ untranslated region (UTR) near start codons and has been shown to promote translation through mechanisms that are not yet understood5,2,5,3. The formation of stable mRNA secondary structure around start codons has been shown to reduce translational efficiency5,6,5. Although it is unclear whether these m1rA modifications target adenine nucleotides involved in WC base-pairing, it is possible that m1rA enhances translation in part by destabilizing secondary structure at the 5′ UTR near the start codons. Indeed, according to our results, m1rA should also be capable of stabilizing alternative RNA secondary structures that feature bulged adenosines, even if such structures are disfavored by as much as ~5 kcal mol−1 in the absence of the modification. Furthermore, placement of m1rA in an unpaired bulged conformation could make the nucleotide accessible to demethylases for achieving efficient reversible control at the epitranscriptomic level (Fig. 6c). Further studies are needed to test this proposed mechanism for m1rA-enhanced translation.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Note:** Any Supplementary Information and Source Data files are available in the online version of the paper.

**ACKNOWLEDGMENTS**

We thank M. Juen (University of Innsbruck, Austria), N. Orlowsky, Y. Xue, A. Shakya, M. Clay, A. Rangadurai, E. Szymanski, members of the Al-Hashimi laboratory, and T. Mustoe (UNC–Chapel Hill) for assistance and critical input. We acknowledge technical support and resources from the Duke Magnetic Resonance Spectroscopy Center, the Duke Compute Cluster and the Shared Materials Instrumentation Facility at Duke University. This work was supported by NIH grants (R01GM088846 to I.A. and H.M.A.; 5P50GM103297 to H.M.A.) and Austrian Science Fund (FWF) grants (P26550 and P28725 to C.K.). H.Z. acknowledges support from the China Scholarship Council.
ARTICLES

AUTHOR CONTRIBUTIONS
H.Z., E.N.N., and H.M.A. conceived the project and experimental design. H.Z. prepared NMR samples, with assistance from I.J.K. and E.N.N., and performed NMR experiments and analyzed NMR data, with assistance from I.J.K. and B.S. H.Z. performed DFT calculations and modeling of steric analysis. H.Z., I.J.K., and E.N.N. performed the structure-based analysis of RNA Hoogsteen base pairs. I.A.G., G.G., and J.M. performed and analyzed the MD simulations. C.H.W. and C.K. carried out the NMR labeling experiments and performed the data analysis. H.M.A., H.Z., and I.A. wrote the manuscript with critical input from I.J.K., B.S., E.N.N., G.G., J.M., T.B., C.H.W., and C.K.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Neidle, S. Principles of Nucleic Acid Structure (Academic Press, 2010).
2. Brameld, K.A. & Goddard, W.A. Ab initio quantum mechanical study of the structures and energies for the pseudodotation of 5'-deoxy-2'-deoxyribonucleic and ribose sugars. J. Am. Chem. Soc. 121, 985–993 (1999).
3. Seeman, N.C., Rosenberg, J.M., Suddath, F.L., Kim, J.J.P. & Rich, A. RNA double-helix structure: synthetic resolution. I. The crystal and molecular structure of sodium adenyl-3',5'-uridylate hexahydrate. J. Mol. Biol. 104, 109–144 (1976).
4. LipPERT, J. & et al. Double-stranded RNA under force and torque: similarities to and striking differences from double-stranded DNA. Proc. Natl. Acad. Sci. USA 111, 1541–1546 (2014).
5. Hoogsteen, K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-acteoline. Acta Crystallogr. 12, 822–823 (1959).
6. Nikolova, E.N. et al. Transient Hoogsteen base pairs in canonical duplex DNA. Nature 470, 498–502 (2011).
7. Alvey, H.S., Gottardo, F.L., Nikolova, E.N. & Al-Hashimi, H.M. Widespread transient Hoogsteen base pairs in canonical duplex DNA with variable energetics. Nat. Commun. 5, 4786 (2014).
8. Nikolova, E.N. et al. A historical account of Hoogsteen base-pairs in duplex DNA. Biopolymers 99, 955–968 (2013).
9. Kitayner, M. et al. Diversity in DNA recognition by p53 revealed by crystal structures of a RNA double-helical DNA and duplex-structured G-DNA: an ab initio quantum mechanical study of the structures and energies for the pseudodotation of 5'-deoxy-2'-deoxyribonucleic and ribose sugars. J. Am. Chem. Soc. 121, 985–993 (1999).
10. Xue, Y.
11. Palmer, A.G. III. A dynamic look backward and forward. Nat. Rev. Mol. Cell Biol. 13, 859–863 (2003).
12. Aas, P.A. et al. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-acteoline. Acta Crystallogr. 12, 822–823 (1959).
13. Chen, F.
14. Asa, S.A. et al. The authors declare no competing financial interests.

Page 1 of 810
ONLINE METHODS

Sample preparation. NMR buffer. All RNA and DNA samples were buffer-exchanged at least three times with a centrifugal concentrator (EMD Millipore) until they contained >99.9% of the desired buffer, which, unless stated otherwise, consisted of 15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 5.4 or 6.8, and 10% D2O.

Uniformly $^{13}$C,$^{15}$N-labeled RNA and DNA samples. bp-A6-RNA and single strands of the E-gc and TAR-UUGC$^{15}$N were prepared through in vitro transcription with uniformly $^{13}$C,$^{15}$N-labeled ribonucleotide triphosphates (Cambridge Isotope Laboratories), T7 polymerase (Takara Mirus Bio) and synthetic DNA templates (Integrated DNA Technologies); samples were purified through 20% (v/v) denaturing PAGE and electroeluted as described above. Subsequently, the derived RNA was oxidized under conditions, as previously described[26], uniformly labeled 2′-deoxyribonucleotide triphosphates (Silantes) as previously described[6]. m1A- and m6G-containing oligonucleotides. Oligonucleotides were purchased from Glen Oligo Synthesis (W.M. Keck Foundation) with GlenPak DNA/RNA cartridge purification (A5′-RNA$m^{15}$A, A5′-RNA$m^{16}$A, A5′-RNA$m^{14}$A, A5′-RNA$m^{13}$A, A5′-RNA$m^{12}$A, A5′-RNA$m^{11}$A, A5′-RNA$m^{10}$A, bp-A6-RNA$m^{15}$A, bp-A6-RNA$m^{14}$A, bp-A6-RNA$m^{13}$A, bp-A6-RNA$m^{12}$A, bp-A6-RNA$m^{11}$A, bp-gc-RNA$m^{15}$A, and bp-gc-RNA$m^{14}$A). Midland Certified Reagents with reverse-phase (RP) HPLC purification (A5′-RNA$m^{16}$A and A5′-RNA$m^{15}$A), and GE Healthcare Drimaco with RP-HPLC purification (A5′-RNA$m^{14}$A and bp-A6-RNA$m^{15}$A). To minimize Drinmo rearrangement of m1A into m6A, all DNA and RNA oligonucleotides containing m1A were synthesized and deprotected with the UltraMild protocol (http://www.glenresearch.com/Technical/TR_UltraMild_Deprotection.pdf; Glen Research Corporation).

Assessing purity of m1A- and m6G-containing oligonucleotides. Samples were assessed with 20% denaturing PAGE, MALDI mass spectrometry, liquid chromatography–mass spectrometry (LC-MS), and NMR spectroscopy. For hairpin constructs, bp-gc-RNA$m^{15}$A and bp-A6-RNA$m^{15}$A, we obtained evidence for incomplete base deprotection during synthesis on the basis of observation of additional imino proton and acetyl group (the N4 protecting group on the cytosine) resonances and NOE cross-peak between the two. Evidence for the acetyl group was also obtained by LC-MS. We suspect that incomplete deprotection arose because of the formation of stable secondary structure in these hairpin constructs during the UltraMild deprotection step. These impurities were effectively eliminated by synthesizing individual single strands of duplex versions of the hairpin sequence (gc-RNA$m^{15}$A and gc-RNA$m^{14}$A). In all cases, the NMR chemical shifts of the N6-methyl group and base moieties (A-C2, N1C, and N1H) were consistent with m1A, and there was no evidence for Drinmo rearrangements[62], which lead to the formation of m6A (Supplementary Figs. 1b and 4). In particular, we observed an ~4 p.p.m. downfield shift in m1A-C2 and m1G-C2 of the gc-RNA$m^{15}$A and gc-RNA$m^{14}$A, respectively. Resonances were assigned with conventional 2D chemical-shift perturbation (CSP) analyses and analyzed with NMRpipe[68] and SPARKY (http://www.cgl.ucsf.edu/home/sparky), respectively. Resonances were assigned with conventional 2D HSQC, HMQC, NOE, and HCN experiments.

Chemical-shift perturbations (CSPs) induced by m1A or m6G for each residue ($\Delta\text{residue}_i$) were calculated with equation (1) (ref. 69) from the average Euclidean distance of all measured CSP ($\Delta\text{residue}_i$) and ($\Delta\text{residue}_j$):

$$\Delta\text{residue}_i = \frac{1}{N} \sum_{i,j} \frac{g}{T} \frac{\Delta\varphi_i}{\Delta\varphi_j}$$

where $g$ is the gyromagnetic ratio of the $i$th nucleus (C, H, or N); $N$ is the total number of CSPs measured for each residue, and $\Delta\varphi$ is the difference in chemical shifts (in p.p.m.) for the $i$th nucleus between the m1A or m6G-modified and unmodified duplexes. Residues with $\Delta\text{residue}_i$ <0.1 p.p.m. are highlighted on the duplexes in Figure 3b and Supplementary Figures 3 and 5. An average CSP ($\Delta\text{residue}_A$) was calculated for each duplex by averaging $\Delta\text{residue}_i$ for two base pairs above and below the modified base pair.

$^{13}$C and $^{15}$N relaxation dispersion. The $^{13}$C and $^{15}$N relaxation dispersion were measured at 600 MHz (14.1 Tesla) and 700 MHz (16.4 Tesla) with Bruker spectrometers as previously described[21,22] with spinlock powers ($\varphi_{\text{SL}}$, 2π $\times$ 1 Hz) and offset frequencies ($\varphi_{\text{EF}}$, 2π $\times$ 1 Hz) listed in Supplementary Table 1. Magnetization of the spins of interest was allowed to relax under an applied spinlock for the following durations: 0–120 ms for N1/N3 in bp-A6-RNA and E-gc; 0–60 ms for C8/C1 in bp-A6-RNA, E-gc, TAR-UUGC$^{15}$N, A5′-DNA, A5′-DNA$^{14}$A, and A5′-DNA$^{13}$A.$^{66}$

Analysis of $R_{1p}$ data. Fitting of $^{13}$C and $^{15}$N $R_{1p}$ data. Experimental $R_{1p}$ relaxation rate constants were calculated by fitting peak intensities versus relaxation delay durations to a single exponential decay[71]. Uncertainty in the fitted $R_{1p}$ values (one s.d.) were derived with a Monte Carlo method[18]. $R_{1p}$ data were fitted to simulated $R_{1p}$ values given by the solution to the Bloch–McConnell (BM) equations[72] at each given $\varphi_{\text{SL}}$ and $\varphi_{\text{EF}}$. Combination. Residual sums of squares were minimized with a bounded least-squares algorithm[73] yielding best-fit exchange parameters. The uncertainty in the chemical-exchange parameters was calculated as the standard error of the fit (1). A two-state-exchange model was used to fit the $R_{1p}$ data profiles of A5′-DNA, A5′-DNA$^{14}$A, and A5′-DNA$^{13}$A with the initial magnetization aligned either along the effective field of the ground (for slow exchange with $k_{\text{ex}} < 1$) or average (for fast exchange with $k_{\text{ex}} > 1$) state[74]. For the $\Delta\text{A6-C8}$ RD data measured in A5′-DNA at low temperatures, both protocols yielded acceptable fits but resulted in different exchange parameters, given the slower exchange rate (Supplementary Table 4). The exchange parameters obtained from the average alignment protocol were selected on the basis of a van’ Hoff analysis[6] (Supplementary Fig. 6d). For the dC15-C6 RD data measured in A6-DNA$^{66}$,
a three-state chemical-exchange model without minor exchange with average alignment was statistically favored over two-state models (Supplementary Fig. 6c). In all cases, the Akaike information criterion (AIC)59 and Bayesian information criterion (BIC)27 were used to select the models.

Analysis of chemical shift and NOE data. Chemical shifts and NOE cross-peaks were used to characterize WC versus HG base pairs. The NOE cross-peaks unique to HG base pairs include strong intranucleotide H1′-H8 NOE for syn purine, (i) A-H2–(i-1) H1′/H2′ and (i) A-H2–(i-1) H6/H8 for syn-adenosine, and H8-H3, A-H6/C–H4–H3, (i) H3–(i+1)–(i-1) H1/H3, and (i) A-H6/C–H4–(i+1)–(i-1) H1/H3 NOEs for connectivity involving imino or amino protons in both G-C and A-U/T HG base pairs73,74. Absence of the canonical sequential (i-1) H1′-i H8 NOE was also expected for syn purines, owing to the base flip. As described previously6, HG base pairs are also characterized by a unique set of chemical shifts relative to WC base pairs, including downfield-shifted purine-C8 and purine-C1′, protonated cytosine-C6 (3 p.p.m.) and upfield-shifted protonated cytosine-C5, guanine-N1 and thymine-N3 (1–2 p.p.m.). The C1′ chemical shifts are also sensitive to sugar pucker. In A-RNA, deviations from the A-form C3′-endo toward a C2′-endo sugar pucker lead to an upfield shift (~4 p.p.m.)75,76. In addition, deviations from the A-form conformation due to loss of stacking and bulging out of nucleotides result in a downfield shift in the base C6/C8 and cytosine-C5 and upfield shift on sugar-C1′18.

Density functional theory geometry optimizations and CS calculations. Density functional theory (DFT) calculations57 with Gaussian 09c (Gaussian) were performed as previously described6 to compute chemical shifts for WC and HG base pairs in A-RNA and B-DNA. In all cases, protons were added with PMol (https://www.pymol.org/), and the phosphate backbone was truncated, thus leaving only the nucleoside motifs for each base pair6. Calculations were performed on an rA-rU HG base pair obtained from snapshots of an rA16-rU9 HG base pair in the X-ray structure of the P4–P6 domain of group I intron RNA (PDB 1L8V78), a tertiary rG-rC HG base pair in the structure of the 23S ribosomal RNA-protein complex (PDB 3U56) and the rG′-rG′ mispair in a duplex RNA structure (PDB 3CZW79). Reference rA-rU or rG-rC WC base pairs were taken from MD snapshots or from the same X-ray structures used to obtain HG base pairs (Supplementary Fig. 6e). Two runs of geometry optimizations were carried out with the B3LYP functional with 3-21G and 6-311 + G(2d,p) basis sets, with all heavy atoms (C, N, and O) frozen. Carbon chemical shifts were computed out with the B3LYP functional with 3-21G and 6-311 + G(2d,p) basis sets, with all heavy atoms (C, N, and O) frozen. Carbon chemical shifts were computed for reference WC base pairs were subtracted from those computed for HG base pairs, thus yielding chemical-shift changes after HG formation (∆ω = ω_{HC} – ω_{WC}).

Analysis of UV melting data. The UV absorbance at 260 nm (A_{260}) as a function of temperature was measured on a Shimadzu UV-3600 UV-vis-NIR spectrophotometer with an eight-cell sample holder with a Fisher Isotemp refrigerated circulator to regulate sample temperature. All DNA and RNA oligonucleotides were diluted directly from NMR samples with the same NMR buffer (15 mM phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 5.4 or 6.8) unless stated otherwise, and triplicate measurements were carried out for each oligonucleotide simultaneously with a sample volume of 125 μL in each cell and an additional reference cell containing the same amount of buffer. The temperature was varied between 5 °C and 90 °C at a ramping rate of 1 °C min⁻¹. The melting temperature (T_m) and enthalpy (∆H) for duplex association and hairpin folding was obtained by fitting the melting curves to equations (2) and (3)50, respectively,

\[ f = 1 + 4e^{(1/T_m-1/T)\Delta H/\Delta R} = 1 + 4e^{(1/T_m-1/T)\Delta H/\Delta R} \]  
\[ f = \frac{e^{(1/T_m-1/T)\Delta H/\Delta R}}{1 + e^{(1/T_m-1/T)\Delta H/\Delta R}} \]  

\[ \Delta S = \Delta H/T_m - R\ln(C_T/2); \Delta G = \Delta H - T\Delta S \]  
\[ \Delta S = \Delta H/T_m; \Delta G = \Delta H - T\Delta S \]  

The fitting was carried out with nonlinear model fitting with Mathematica 10.0 (Wolfram Research). Errors in T_m and ΔH represent the s.d. from triplicate measurements. The stabilization effects due to m^2 A and m^2 G in DNA or RNA were calculated by taking the difference in free energy for folding, i.e.:  

\[ \Delta G = \Delta G_{mod} - \Delta G_{unmod} \]  
\[ \Delta H = \Delta H_{mod} - \Delta H_{unmod} \]  
\[ \Delta S = \Delta S_{mod} - \Delta S_{unmod} \]  

Because the Dimroth rearrangement can occur for m^2 A in both DNA and RNA under basic conditions60 and high temperatures61, melting experiments were repeated for all m^2 A-containing duplexes when the temperature was restricted to <85 °C, under both neutral (pH 6.8) and acidic (pH 5.4) conditions. These control experiments yielded reproducible melting curves, and the fitted thermodynamic parameters at neutral or acidic pH conditions were within experimental error (Supplementary Table 3). 1H 1D NMR spectra recorded for the A6 DNA sample after melting showed insignificant changes and no evidence of Dimroth rearrangements (data not shown).

Steric analysis and survey of HG base pairs in RNA. A-T/U and G-C WC base pairs were obtained from idealized B-DNA and A-RNA helices built with 3DNA51 (Fig. 4b). 146 and 159 WC base pairs surrounded by at least one WC base pair above and below were obtained from high-resolution (<2 Å) X-ray structures of A-RNA and B-DNA duplexes, respectively, in the PDB. Purine bases were flipped around the glycosidic bond, and interatomic distances were measured with an in-house Python script. The proximity of the exocyclic amino group on guanine to the phosphate group during the base flip was not considered to be a steric clash, given the potential for hydrogen-bonding. The survey of HG base pairs in RNA was carried out according to the same protocol reported for B-DNA32. Briefly, all RNA X-ray structures with resolution ≤5 Å were downloaded from the PDB on August 31 2014. The same in-house program was used to identify rA-rU and rG-rC base pairs in RNA structures with three HG criteria (hydrogen-bonding, constricted C1′-C1′ distance, and syn purine)52. Redundancies defined as base pairs surrounded by the same sequence contexts and from the same RNA or RNA–protein or RNA–ligand complexes were removed by manual inspection, as described previously52. The survey identified a single rA-rU HG base pair that reoccurs in four distinct X-ray structures of the P4–P6 domain of the Tetrahymena thermophila group 1 intron RNA (PDB 1GID33, 1L8V78, 1HR2 (ref. 82), and 2R8S53). The RNA HG survey also identified several examples of long-range HG base pairs forming tertiary contacts, HG base pairs in triplexes, and reverse rA-rU HG base pairs in duplexes typically near rG-rA mismatches.

Biased and unbiased molecular dynamics simulations. Structure generation for MD simulation. hp-Arc,RNA, hp-Arc,RNA 3′→5′, and Arc,RNA helices were built with make-nas44 with all bases in WC conformation. In the case of hp-Arc,RNA, a duplex structure was generated with make-na, and the UUCG loop was attached and annealed with the CHARMM simulation package85. Structures with HG conformation at A16 were created by rotation along the glycosidic-bond angle χ by 180°.

Unbiased MD equilibrium simulations. All structures were simulated with constant-temperature MD with the CHARMM36 forcefield86 and a generalized Born molecular volume (GBMV) implicit solvent57; parameters for m^2 A were taken from Xu et al.88. Integration used a velocity-Verlet algorithm with a time step of 1 fs. The cutoff for nonbonded list generation was 21 Å, the cutoff for nonbonded interactions was 18 Å, and the onset of switching for nonbonded interactions occurred at 16 Å. The SHAKE algorithm was used to constrain the
covalent bonds to hydrogen atoms involved. Each structure was heated to 300.0 K with harmonic constraints on all nonhydrogen atoms, heating occurred in 1-ps increments of 1.0 K for a total of 300 ps steps, and this was followed by 200-ps equilibration at 300.0 K. Harmonic constraints were then gradually removed during a sequence of four reductions for 50 ps each. Unbiased production-run simulations were then run for 3 ns without constraints for each system. Ten independent simulations with hp-AC2RNA and AC2-DNA, and with A16 in HG conformation were produced from independent conformations obtained during the heating and equilibration method described above. The simulation for AC2-DNA in HG was repeated twice.

Global r.m.s. deviations were calculated from the single 3-ns trajectories of m3A starting in HG for both hp-AC2RNA and AC2-DNA,

\[ \text{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} (r_i(t) - r_i^0)^2}{N}} \]

in which \( r_i(t) \) is the instantaneous coordinate of an atom, and \( r_i^0 \) is the position of the reference structure. Hydrogen-bond presence was evaluated with CHARMM's COOR HBOND module for each trajectory with cutoff distance and angle of 3.6 Å, and 120°, per Goldsmith COOR HBOND module for each trajectory with cutoff distance and angle of 3.6 Å, and 120°, per Goldsmith

Biased MD simulations. The protocols for minimization, heating, and solvation were identical to those used for the unbiased simulations. The biased molecular dynamics method implemented in the CHARMM package was used to force conformational transitions between WC and HG states with a biasing potential \( W(p(t)) \), applied according to equation (7)

\[
W(p(t)) = \begin{cases} 
\alpha / 2(p(t) - p_{g}(t))^2, & \text{if } p(t) < p_{g}(t) \\
0, & \text{if } p(t) \geq p_{g}(t)
\end{cases}
\]

where

\[
p(t) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} (r_i(t) - r_j(t))^2
\]

\( p(t) \) is a collective distance between the instantaneous \( \{r_i(t)\} \) and the reference structure \( \{r_i^0\} \), and \( \alpha \) is the strength of the half-harmonic bias. In all cases, biases were placed between pairs of atoms that share a hydrogen bond in the target structure, ensuring that the adenine base would not only perform the ~180° flip but also form the definitive hydrogen-bonding structure of the desired WC or HG configuration. After the biased trajectories were generated, they were post-processed in CHARMM, outputting the \( \chi \)-angle dependence of the relative interaction energy in the absence of the bias. Only successfully flipping trajectories were used, thus resulting in 40 trajectories for AC2-DNA, 24 for hp-AC2RNA, and 25 for hp-AC2RNA 3′→5′. The relative interaction energy was calculated for the base pair that includes the flipping base as well as the base pairs above and below the flipping base. Angle-energy pairs were binned into 50 bins, and the mean of the energy was evaluated within each bin. Plots of relative interaction energy as a function of the \( \chi \)-angle were thus generated.

61. Zimmer, D.P. & Crothers, D.M. NMR of enzymatically synthesized uniformly 13C15N-labeled DNA oligonucleotides. Proc. Natl. Acad. Sci. USA 92, 3091–3095 (1995).
62. Macon, J.B. & Wolfenden, R. 1-Methyladenosine: Dimroth rearrangement and reversible reduction. Biochemistry 7, 3453–3458 (1968).
63. Timofeev, E.N. et al. Oligodeoxynucleotides containing 2′-deoxy-1-methyladenosine and Dimroth rearrangement. Helv. Chim. Acta 90, 928–937 (2007).
64. Pivovarova, V.B., Stepianan, S.G., Reva, I.D., Sheina, G.G. & Blagoy, Y.P. Infrared spectra and the structure of 1-methyladenine in an argon matrix and solutions. Spectrochim. Acta A Mol. Biomol. Spectrosc. 51, 843–853 (1995).
65. Longhin, A.P. et al. Chemo-enzymatic synthesis of site-specifically isotope-labeled nucleotides for use in NMR resonance assignment, dynamics and structural characterizations. Nucleic Acids Res. 44, e52 (2016).
66. Wenter, P., Raymond, L., Azuwer, S.D., Allain, F.H.T. & Pitsch, S. Short, synthetic and selectively 13C-labeled RNA sequences for the NMR structure determination of protein-RNA complexes. Nucleic Acids Res. 34, e79 (2006).
67. Wunderlich, C.H. et al. Synthesis of G(13C)pyrimidine nucleotides as spin-labels for RNA dynamics. J. Am. Chem. Soc. 127, 7558–7569 (2005).
68. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR. 6, 277–293 (1995).
69. Williamson, P.M. Using chemical shift perturbation to characterise ligand binding. Prog. Nucleic. Reson. Spectrosc. 73, 1–16 (2013).
70. Branch, M., Coleman, T. & Li, Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. SIAM J. Sci. Comput. 21, 1–23 (1999).
71. Bothe, J.R., Stein, Z.W. & Al-Hashimi, H.M. Evaluating the uncertainty in exchange parameters determined from off-resonance R1p relaxation dispersion for systems in fast exchange. J. Magn. Reson. 244, 18–29 (2014).
72. Wagenmakers, E.-J. & Farrell, S. AIC model selection using Akaike weights. Psychon.
73. Sklenár, V. & Feigon, J. Formation of a stable triplex from a single DNA strand. Nature 345, 836–838 (1990).
74. de los Santos, C., Rosen, M. & Patel, D. NMR studies of DNA (R+H)·Y·(Y+H) triple helices in solution and amino proton markers of T·A.T and G·C·C base-triple formation. Biochemistry 28, 7282–7289 (1989).
75. Fürtig, B., Richter, C., Bermel, W. & Schwalbe, H. New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUUC tetraloop. J. Biomol. NMR 28, 69–79 (2004).
76. Fonville, J.M. et al. Chemical shifts in nucleic acids studied by density functional theory calculations and comparison with experiment. Chem. Biology 18, 12372–12387 (2012).
77. Xu, X.-P. & Au-Yeung, S.C.F Investigation of chemical shift and structure relationships in nucleic acids using NMR and density functional theory methods. J. Phys. Chem. B 104, 5641–5650 (2000).
78. Battle, D.J. & Doundia, J.A. Specificity of RNA-RNA helix recognition. Proc. Natl. Acad. Sci. USA 99, 11676–11681 (2002).
79. Rytniewski, W., Adamik, D.A., Milecki, J. & Adamik, R.W. Noncanonical G(syn)-G(anti) base pairs stabilized by sulphate anions in two X-ray structures of the (GUGGCUGAUAGGGC) RNA duplex. RNA 14, 1845–1851 (2008).
80. Crothers, D.M., Bloomfield, V.A. & Trinco, I. Nucleic Acids: Structures, Properties, and Functions (University Science Books Sausalito, 2000).
81. Lu, X.J. & Olson, W.K. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat. Protoc. 3, 1213–1227 (2008).
82. Joneset, K., Pedell, E., Harrington, D.J. & Cech, T.R. Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA-solvent interactions. Structure 9, 221–231 (2001).
83. Ye, J.-D. et al. Synthetic antibodies for specific recognition and crystallization of structured RNA. Proc. Natl. Acad. Sci. USA 105, 82–87 (2008).
84. Macke, T.J. & Case, D.A. in Molecular Modeling of Nucleic Acids Vol. 682 (eds. Leontes, N.B. and SantaLucia, J.) Ch. 24, 379–393 (American Chemical Society, 1998).
85. Brooks, B.R. et al. CHARMM: the biomolecular simulation program. J. Comput. Chem. 30, 1545–1614 (2009).
86. Denning, E.J., Priyakumar, U.D., Nilsson, L. & Mackere, A.D. Jr. Impact of 2′-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. J. Comput. Chem. 32, 1929–1943 (2011).
87. Lee, M.S., Feig, M., Salsbury, F.R. Jr. & Brooks, C.L. III. New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. J. Comput. Chem. 24, 1348–1356 (2003).
88. Xu, Y., Yanommeslieghe, K., Aleksandrov, A., Mackere, A.D. Jr. & Nilsson, L. Additive CHARMM force field for naturally occurring modified ribonucleotides. J. Comput. Chem. 37, 896–912 (2016).
89. Goldsmith, M., Raghtinawala, T. & Yathindra, N. Selective preference of parallel DNA triplexes Is due to the disruption of Hoogsteen hydrogen bonds caused by the severe nonisostericity between the G·GC and T·AT Triplets. PLoS One 11, e0152102 (2016).
90. Paci, F. & Karplus, M. Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. J. Mol. Biol. 288, 441–459 (1999).