BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Secondary Structure of a Conserved Domain in an Intron of Influenza A M1 mRNA

Tian Jiang,[†] Scott D. Kennedy,[‡] Walter N. Moss,[§] Elzbieta Kierzek,[∥] and Douglas H. Turner*,[†]

[†]Department of Chemistry and Center for RNA Biology, University of Rochester, Rochester, New York 14627, United States

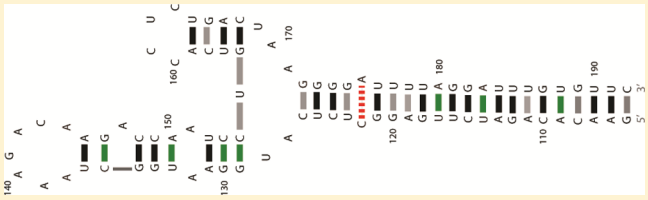
[‡]Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, New York 14627, United States

[§]Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06536, United States

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

Supporting Information

ABSTRACT: Influenza A virus utilizes RNA throughout infection. Little is known, however, about the roles of RNA structures. A previous bioinformatics survey predicted multiple regions of influenza A virus that are likely to generate evolutionarily conserved and stable RNA structures. One predicted conserved structure is in the pre-mRNA coding for essential proteins, M1 and M2. This structure starts 79 nucleotides downstream of the M2 mRNA 5′ splice site. Here, a combination of biochemical structural mapping, mutagenesis, and NMR confirms the predicted three-way multibranch structure of this RNA. Imino proton NMR spectra reveal no change in secondary structure when 80 mM KCl is supplemented with 4 mM $MgCl_2$. Optical melting curves in 1 M NaCl and in 100 mM KCl with 10 mM $MgCl_2$ are very similar, with melting temperatures ∼14 °C higher than that for 100 mM KCl alone. These results provide a firm basis for designing experiments and potential therapeutics to test for function in cell culture.



Influenza A virus is a member of the *Orthomyxoviridae* family of enveloped viruses with segmented, single-stranded, negative-sense RNA genomes. Each year, influenza A infects an estimated three to five million people worldwide, killing up to 500 000 people.[1] Moreover, influenza pandemics have occurred several times in the past 100 years. For example, the 1918−1920 "Spanish Flu" claimed more than 50 million lives.[2] Few diseases have had a greater impact than influenza on public health and global economic output. In 2012−2013, the United States had an unusually bad flu season, with overall vaccine effectiveness about 47 and 67% against influenza A and B, respectively.[3] Currently, few drugs treat influenza; neuraminidase inhibitors are essentially the only therapeutics available.[4] Moreover, sporadic cases of drug-resistant influenza viruses have been detected worldwide.[5−7] The once widely used adamantanes are now mostly ineffective toward currently circulating influenza (H3N2).[8] Thus, it is important to develop novel antiviral treatments as well as more effective vaccines.[9,10]

RNA structure plays key roles in many viruses, including influenza. For example, a panhandle/corkscrew structure in influenza genomic viral (v)RNA required for RNA transcription, replication, and packaging is formed by annealing the 5′ and 3′ ends of influenza vRNAs.[11] Beyond this vRNA, knowledge of influenza virus RNA structures is limited. vRNA is coated with viral nucleoprotein (NP) much of the time, which may melt secondary structure.[12] At various stages of infection, however, regions of vRNA are free of NP and may form functional RNA structures. Influenza positive-sense RNAs are predicted to contain extensive conserved and stable secondary structure. In particular, segments 7 and 8, both of which are alternatively spliced, are enriched in predicted structures.[13] A recent survey of predicted structure in influenza B and C discovered evidence of conserved structures in coding RNAs from spliced segments.[14] Interestingly, in all three viral species (influenza A, B, and C), predicted conserved structures occur at or near splice sites, suggesting common strategies for the regulation of splicing. Knowledge of influenza virus RNA structure can inform experiments to reveal function, enrich understanding of molecular mechanisms underlying the viral life cycle, and facilitate development of new therapeutics.

Segment 7 of influenza A encodes M1 protein and is alternatively spliced to produce M2, M3, and occasionally M4 mRNA (Figure 1A).[15,16] M1 and M2 proteins are essential in the viral life cycle. M1 (252 amino acids) is the most abundant influenza protein. It is the matrix protein that connects vRNPs to each other and to the viral envelope. It also determines the directionality of vRNP transport.[17] M2 (97 amino acids) is a transmembrane ion channel protein that permits the flow of protons from the endosome into the virion interior to facilitate

Received: May 19, 2014
Revised: July 10, 2014
Published: July 15, 2014

   dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

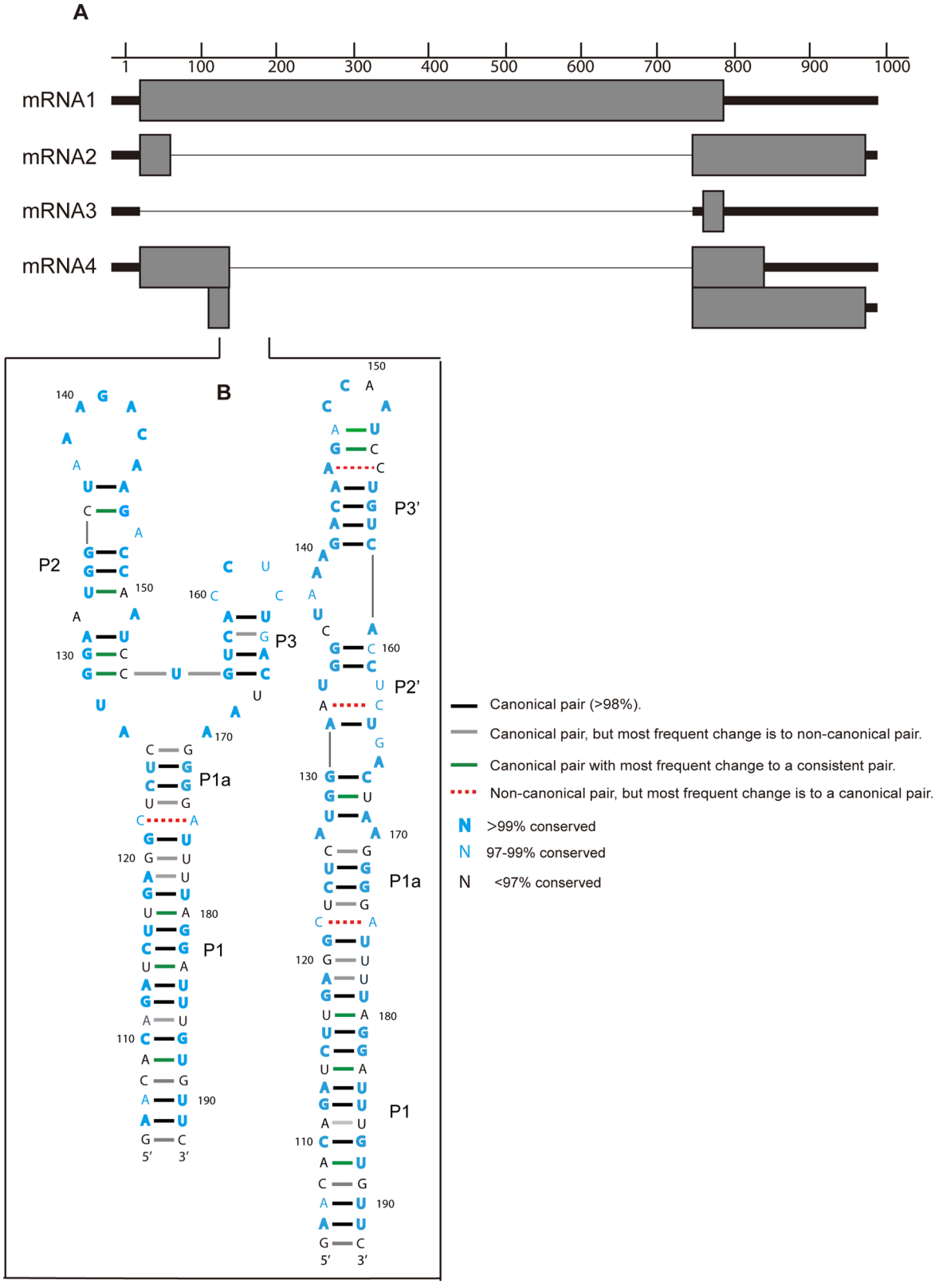


Figure 1. Location and predicted structures of the wild-type RNA. (A) Diagrammatic summary of mRNA splice variants. Gray boxes represent coding regions. Bold lines represent noncoding regions. Thin lines represent introns. M4 mRNA has two different open reading frames.[21] It was only observed in a few viral strains.[16] The M4 mRNA 5′ splice site, GAG/GUUCUC (nucleotides 118−126 with/as splice site), is not present in the consensus sequence. (B) Predicted multibranch model[13] and hairpin model. Base pairs are color-annotated with information from base pair counts from an alignment of all available unique sequences (see Supporting Information). Conserved nucleotides are shown in blue (see color annotation key).

viral uncoating.[18,19] Temporal control of splicing is required for generating various mRNA isoforms that must be present at differing abundances over the course of infection.[20] Furthermore, the alternative splicing of segment 7 is complex, and except for

dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

M2, the products of spliced mRNAs are not well characterized. For example, the M3 mRNA 5′ splice site more closely fits the consensus 5′ splice site motif than the M2 mRNA 5′ splice site, the latter of which has C rather than G at the 3′ end of the 5′ exon; this finding is surprising because M2 is essential for viral replication, while the M3 protein has yet to be observed.[20] Additionally, some viral strains have an M4 mRNA 5′ splice site.[16] Normally, M4 mRNA is not translated, but when the M2 mRNA 5′ splice site is disrupted, M4 mRNA can produce M42 protein, which can functionally replace M2 to support efficient replication in tissue culture cells and exhibit pathogenicity in an animal host.[21] M2, M3, and M4 mRNA share the same 3′ splice site. Multiple factors, including viral polymerase complex, NS1 protein, and cellular splicing factor SF2/ASF, have been implicated in helping to regulate alternative splicing of segment 7.[20,22,23]

A 63 nucleotide fragment in segment 7 containing the 3′ splice site, key residues of SF2/ASF binding site, and the polypyrimidine tract required for splicing exists in an equilibrium between a pseudoknot and a hairpin structure.[24] The conformational switch places each splicing element into different structural contexts. Interestingly, a similar pseudoknot/hairpin conformational switch was described at the 3′ splice site of segment 8.[25] In addition to the structure at the 3′ splice site of segment 7, a region 79 nucleotides downstream from the 5′ splice site for M2 mRNA (nucleotides 105−192) was predicted to have an especially stable and conserved secondary structure (Figure 1).[13] This sequence and predicted structure are conserved within influenza A viruses that infect human, swine, and avian species. A multibranch model was predicted[13] for the consensus sequence of all wild-type sequences available in the National Center for Biotechnology Information (NCBI) influenza virus resource.[26] A possible alternative hairpin structure was also hypothesized for this region, with equally conserved base pairing and similar predicted free energy. This putative alternative hairpin for segment 7 is attractive: a region 51 nucleotides downstream from the 5′ splice site of segment 8 folds into a hairpin conformation in solution.[27] Here, data from nondenaturing polyacrylamide gel electrophoresis, enzymatic/chemical mapping, isoenergetic microarray mapping, and NMR indicate that only the multibranch structure forms in solution.

EXPERIMENTAL METHODS

Preparing the Wild-Type and Mutant RNAs. The consensus sequence[13] for the intron region of M1 mRNA was deduced from all available unique sequences from the NCBI Influenza Virus Resource.[26] This sequence was also found in wild-type sequences of human, swine, and avian influenza viruses (GenBank accession nos. CY147367, GQ404614, CY037901, CY021382, CY009453, and M63525). The RNA was transcribed using the AmpliScribe T7 high yield transcription kit (Epicentre) from a double-stranded DNA oligonucleotide template. The antisense sequence is 5′-GAACACAAATCCTAAAATCC-CCTTAGTCAGAGGTGACAGGATTGGTCTTGTCTTT-AGCCATTCCATGAGAGCCTCAAGATCTGTGTTC *TATAGTGAGTCGTATTA*GAATTC-3′. The *italic* letters are complementary to the T7 promoter sequence. DNA templates were removed using DNase I after incubating the reactions at 37 °C for 4 h following Epicentre's protocol. RNA was purified using denaturing 8% PAGE and, when required, 5′-end-labeled with [γ^{32}P]-ATP. Two mutants were made and labeled in the same way.

Denaturing and Nondenaturing PAGE. Nondenaturing 6% PAGE was run for all full-length constructs. The gel was prepared and run with 1× THEM buffer (66 mM HEPES, 34 mM Tris, 0.1 mM disodium EDTA, and 10 mM $MgCl_2$, pH 7.4).[28] End-labeled RNAs were renatured by heating at 90 °C for 2 min and then slow cooling to 37 °C, at which point the RNAs were incubated for 20 min in 10 mM Tris (pH 7.0), 100 mM KCl, and different concentrations of $MgCl_2$. Electrophoresis was conducted at 15 W for 12 h at 4 °C. Dried gels were analyzed by exposing to a phosphorimager screen.

Denaturing 6% PAGE (7.5 M urea) was run for the same RNAs prior to end-labeling. The gel was prepared and run with 1× TBE buffer. RNAs were denatured by heating at 90 °C for 2 min in gel loading buffer II (Ambion) before loading. The gel was stained with 1× SYBR Green II (Invitrogen).

Enzymatic and Chemical Mapping. RNAs used in all mapping experiments were folded in 10 mM Tris (pH 7.0), 100 mM KCl, and 10 mM $MgCl_2$, as described for native gel analysis. All mapping experiments were carried out at room temperature.

Enzymatic mapping reactions were adapted from manufacturer's protocol[29] and carried out on 5′-end-labeled RNAs. Optimal enzyme concentration was determined by titration. The digestion reactions were stopped by ethanol precipitation.

DMS, CMCT, and SHAPE mapping reactions were adapted from published protocols[30,31] and carried out on unlabeled RNAs. Reactions were stopped by ethanol precipitation. Modifications were read out by primer extension using an LNA-modified primer (5′-GAA^{L}CA^{L}CA^{L}AA^{L}TC^{L}CTA^{L}AAA-3′) complementary to nucleotides 176−192. The primer was synthesized according to published method,[32,33] 5′-end-labeled with [γ^{32}P]-ATP, and purified by denaturing 8% PAGE.

DEPC reactions were run as described by Moss et al.[24] End-labeled RNAs were incubated with 0.69 mM DEPC, followed by $NaBH_4$ reduction and aniline cleavage. Reactions were stopped by ethanol precipitation.

Digestion and modification products were fractionated on a denaturing 8% polyacrylamide gel. All gels were dried, exposed to a phosphorscreen, and imaged with a Bio-Rad personal molecular imager. The intensities of the product bands were quantified using ImageJ,[34] and the bands were considered strong and medium when the integrated intensities were $\geq^{2}/_{3}$ and $\geq^{1}/_{3}$ of the strongest integrated intensity, respectively.

Isoenergetic Microarray Mapping. Microarray probing was conducted on the wild-type RNA with isoenergetic pentamer and hexamer 2′-O-methyl oligonucleotide probes with LNA and 2,6-diaminopurine substitutions.[35] Universal microarrays with 861 probes divided into two microarray slides were used. Negative internal controls were U, UUUUU, and spotted buffer. Microarrays were printed at the European Center of Bioinformatics and Genomics in Poznan, Poland.

Radioactively labeled RNA was folded in 10 mM Tris-HCl (pH 7.0), 300 (or 100) mM KCl, and 10 mM $MgCl_2$, as described above. RNA in folding buffer was incubated with the microarray at 4 °C for 18 h to allow hybridization. Then, buffer with RNA was poured out, and slides were washed in the same buffer for 1 min at 0 °C and dried by centrifugation. Hybridization was visualized by exposure to a phosphorimager screen, and quantitative analysis was performed with ArrayGaugeV2.1. Possible alternative binding sites were predicted using RNA−RNA thermodynamics[36] with the RNAstructure program.,[37]

NMR Sample Preparation. The 61 nucleotide sequence corresponding to the multibranch structure was transcribed from

a double-stranded DNA oligonucleotide template with antisense sequence 5′-TmGmGATCCCCTTAGTCAGAGGTG-ACAGGATTGGTCTTGTCTTTAGCCATTCCATGAG-AGCCC *TATAGTGAGTCGTATTA*GAATTC-3′. The *italic* letters are complementary to the T7 promoter sequence. The last two nucleotides of the 5′ end were modified with C2′-methoxyls (annotated by "m") to reduce nontemplated nucleotide addition at the 3′ end by the T7 polymerase.[38] Transcription was carried out in 40 mM Tris (pH 7.0), 10 mM spermidine, 0.01% Triton, 10 mM $MgCl_2$, 40 mM DTT, 4 mM each NTP, 1 U/mL inorganic pyrophosphatase, with 5 μM DNA template, and 120 μg/mL recombinant T7 RNA polymerase in 20 mL reaction volume. Recombinant T7 RNA polymerase was expressed and purified from BL21 cells.[39] The reaction was stopped after incubating 4 h at 37 °C upon adding 800 μL of 0.5 M EDTA. RNA was purified by denaturing 8% PAGE and dialyzed into 20 mM KH_2PO_4 (pH 6.5), 80 mM KCl, 0.05 mM EDTA using Millipore Amicon Ultra-15 centrifugal filter units (MWCO 3 kDa).

NMR Experimental Conditions. The RNA was renatured by heating at 90 °C for 2 min and slow cooling to 37 °C. Then, it was put in a Shigemi NMR tube with 10% D_2O. NMR spectra were acquired on a Varian Inova 600 MHz spectrometer. One-dimensional imino proton spectra were recorded at temperatures ranging from 0 to 25 °C. Two-dimensional NOESY spectra were recorded at 0, 5, 12, 20, and 25 °C with mixing time ranging from 100 to 250 ms, processed with NMRpipe,[40] and analyzed with Sparky.[41] Subsequently, up to 4 mM final concentration $MgCl_2$ was added in NMR buffer. RNA was renatured by incubating at 37 °C for 20 min after adding $MgCl_2$. 2D-NOESY spectra were then measured under the conditions described above.

Optical Melting Curves. Absorbance versus temperature melting curves were measured at 280 nm with a heating rate of 0.5 °C min^{-1} in (A) 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, and 100 mM KCl; (B) 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, 100 mM KCl, and 10 mM $MgCl_2$; (C) 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, 300 mM KCl, and 10 mM $MgCl_2$; and (D) 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, and 1 M NaCl on a Beckman Coulter DU 640 spectrophotometer. Melting curves were normalized and analyzed with MeltWin 3.5.[42]

RESULTS

Native Gel Electrophoresis and Enzymatic/Chemical Mapping. The two potential secondary structures of M1 mRNA are identical at the basal stem, but nucleotides between positions 127 and 170 are folded differently. Comparative sequence analysis does not favor one structure over the other (Figures 1B and Table S1). To distinguish between the structures, native PAGE and enzymatic/chemical mapping experiments were performed on *in vitro* transcribed wild-type RNA and on mutants designed to stabilize the multibranch (MBmutant) or the hairpin (HPmutant) conformation.

Figure 2 shows native PAGE of the wild-type RNA and mutants. All RNA ran as a single band, indicating that the wild-type RNA and mutants each fold into a single conformation. The mobility of the wild-type RNA is the same as that for MBmutant but slower than that for HPmutant. The three constructs are of the same size (see denaturing gel in Figure 2B) This suggests that the wild-type RNA folds into the multibranch conformation.

Figure 3 shows enzymatic/chemical mapping results for the wild-type RNA. Enzymatic mapping[43] used RNase A (cleaves after unpaired pyrimidines), RNase T1 (cleaves after unpaired

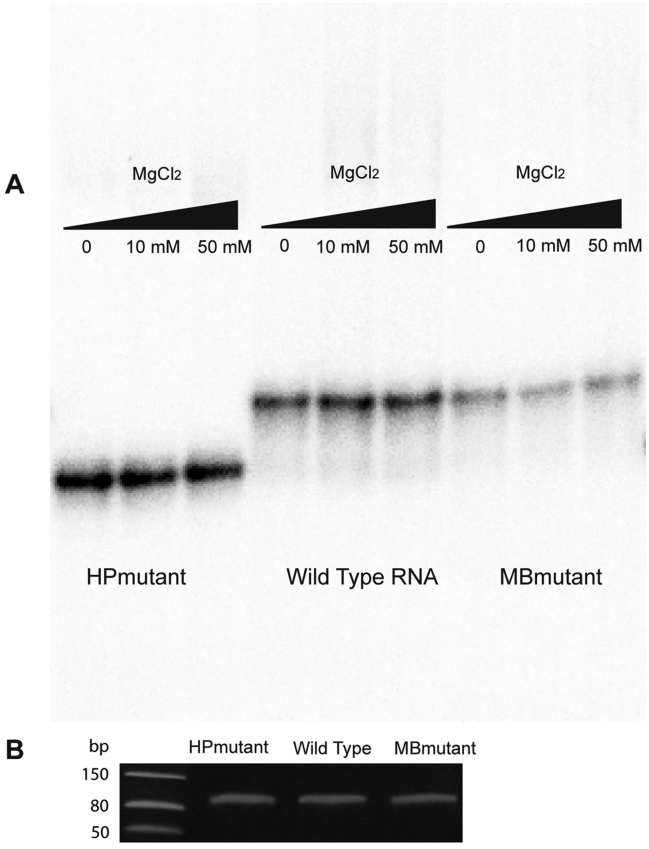


Figure 2. Native gel of the wild-type RNA and mutants. (A) The native gel. RNAs were folded in 10 mM Tris (pH 7.0) and 100 mM KCl with increasing $MgCl_2$ concentration (0, 10, and 50 mM). (B) Denaturing gel of wild-type and mutant RNAs.

G), RNase I_f (cleaves after single-stranded residues),[44] and RNase V1 (cleaves after double-stranded or structured regions). Chemical mapping[43] used DMS (modifies N1 of A and N3 of C when unpaired), CMCT (modifies N3 of U and N1 of G when unpaired), and DEPC (modifies an exposed N7 of A). SHAPE mapping[45] was used to identify flexible regions. The results agree with the predicted multibranch structure but not with the hairpin model. In particular, A142 and C143 were heavily hit by single-strand sensitive probes, which is inconsistent with the hairpin model. Also, U133, C148, C149, C154, and G165, which are single-stranded in the hairpin model but paired in the multibranch model, were not hit by any single-strand sensitive probes.

Mapping results of two mutants provide additional evidence for the multibranch structure. MBmutant was designed to stabilize the multibranch conformation (and forbid the hairpin) by changing the A132−A151 pair into a CG pair (Figure 4). This mutation results in a predicted free energy, ΔG°_{37}, of −35.3 kcal/mol for the multibranch conformation, a gain of −6.2 kcal/mol in stability compared to that of the multibranch structure of the wild-type RNA, without a change in predicted ΔG°_{37} for the hairpin (see captions to Figures 3 and 4). Thus, the equilibrium constant of the mutated sequence for folding to the multibranch conformation is predicted to be 8×10^6-fold more favorable than to the hairpin. As shown in Figure 4, the mapping results of MBmutant are very similar to that of the wild-type RNA. There were significant differences only in P2: in the wild-type RNA, nucleotides A131, A132, A150, and A151 were hit by several single-strand sensitive probes (Figure 3), whereas in MBmutant,

dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

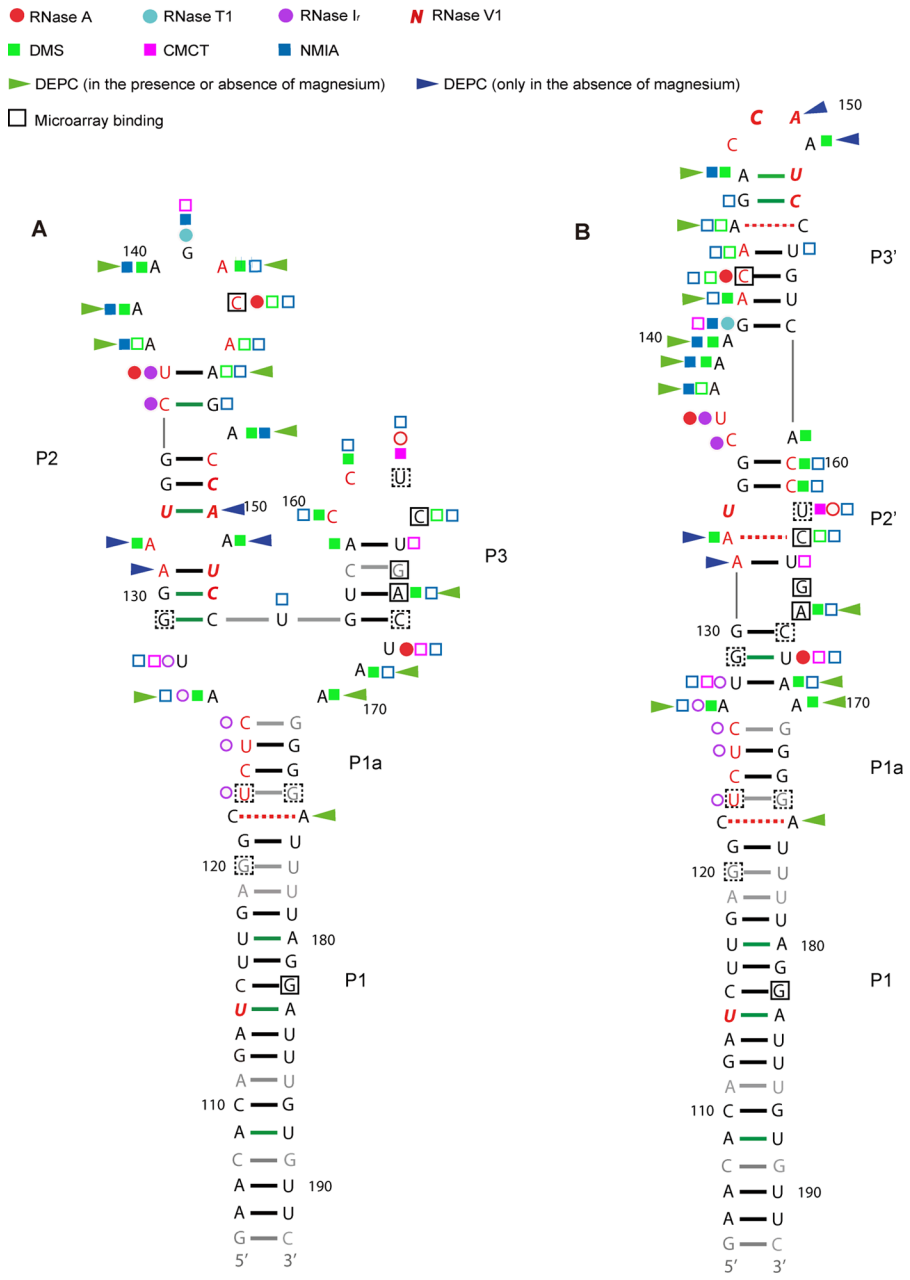


Figure 3. Mapping results of the wild-type RNA. (A) The multibranch model. (B) The hairpin model. The color annotation key is given at the top. For single-strand sensitive nucleases and small molecule probes, filled and open shapes indicate strong and medium hits, respectively. For RNase V1, bold *italic* letters and regular letters indicate strong and medium hits, respectively. For microarray probing, solid and dashed boxes indicate the center of strong and medium binding sites, respectively. The predicted free energies of folding, ΔG°_{37},[54] for the multibranch and hairpin models are −29.1 and −25.5 kcal/mol, respectively.

this region was unreactive to single-strand sensitive probes, as expected for the multibranch conformation but not the hairpin conformation. Thus, the enzymatic/chemical mapping results for MBmutant are consistent with the multibranch model but not the hairpin model.

HPmutant was designed to fold into the hairpin conformation by changing the CCAA tetraloop in P3′ into GCAA, predicted to be more stable by −0.4 kcal/mol at 37 °C. Also, three mismatches along the stem were mutated to canonical base pairs by A132 → G, A145 → G, and U162 → A changes (Figure 5). These mutations make a very stable hairpin conformation with −40.6 kcal/mol predicted free energy at 37 °C, making the equilibrium constant for folding to the hairpin predicted to be 8 × 10^{10}-fold more favorable than to the multibranch conformation (see captions to Figures 3 and 5). The enzymatic/chemical mapping results for this mutant are consistent with the hairpin model but not the multibranch model. As shown in Figure 5, G141, A142, and C143 were not hit by single-strand sensitive probes in HPmutant as they were in the wild-type RNA (Figure 3). Additionally, nucleotides that are expected to be single-stranded in the hairpin model of both HPmutant and the wild-type RNA, including nucleotides 148 and 165, were only hit by single-strand sensitive probes in HPmutant. These observations suggest that the wild-type RNA does not fold into the hairpin model.

dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

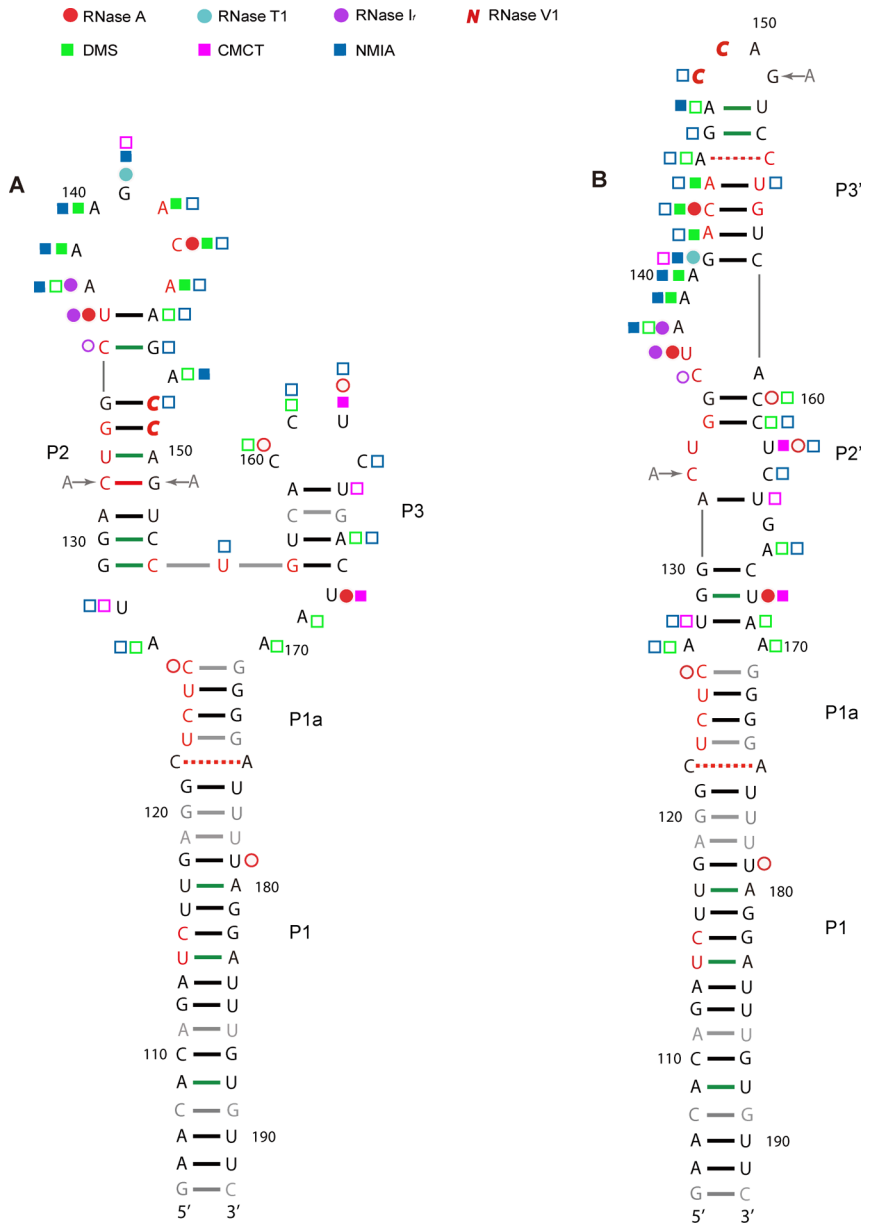


Figure 4. Mapping results of MBmutant. (A) The multibranch model. (B) The hairpin model. Color annotation is the same as in Figure 3. Mutations are indicated by arrows. The predicted free energies of folding, ΔG°_{37}, for the multibranch and hairpin models are −35.3 and −25.5 kcal/mol, respectively.

Isoenergetic Microarray Probing of the Wild-Type RNA. A complementary method for probing RNA structure, isoenergetic microarray mapping,[35] was also applied. The isoenergetic microarray uses pentamer and hexamer 2′-O-methyl RNA probes modified by inclusion of LNA and 2,6-diaminopurine to provide similar free energies of binding to unfolded complementary RNA, independent of sequence, and to stabilize the binding compared with unmodified probes.[46,47] Thus, probe binding interrogates primarily RNA folding rather than differences in thermodynamic binding affinity. The average predicted free energy of binding for the modified library to complementary sites on the wild-type RNA is −9.2 ± 0.8 kcal/mol at 37 °C, compared with −2.4 ± 1.6 kcal/mol for an unmodified DNA library. The stability enhancement averages 60 000-fold, and just as important, the sequence dependence of free energies, compared to unmodified probes, is reduced; this greatly simplifies interpretation of binding results.

The isoenergetic microarray binding results on the wild-type RNA are consistent with enzymatic/chemical mapping results and also support the multibranch structure (Figures 3 and S1 and Table S2). Five unambiguous strong binding sites were revealed: C143, C163, G165, A166, and G182 (center nucleotides of probe binding sites). Except for G182, most nucleotides at binding sites were also hit by single-strand sensitive probes, confirming their accessibility in the structure. Strong binding sites with possible alternative binding sites include nucleotides C126, A142, and A147, which also react with single-strand sensitive probes.

Interestingly, when microarray probing was done in 100 mM KCl/10 mM $MgCl_2$ instead of 300 mM KCl/10 mM $MgCl_2$, no detectable binding was observed. This suggests that the wild-type RNA folds into a very compact and stable structure at 100 mM KCl/10 mM $MgCl_2$. Evidently, the 300 mM KCl makes the binding sites in the RNA more accessible. This increased

dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

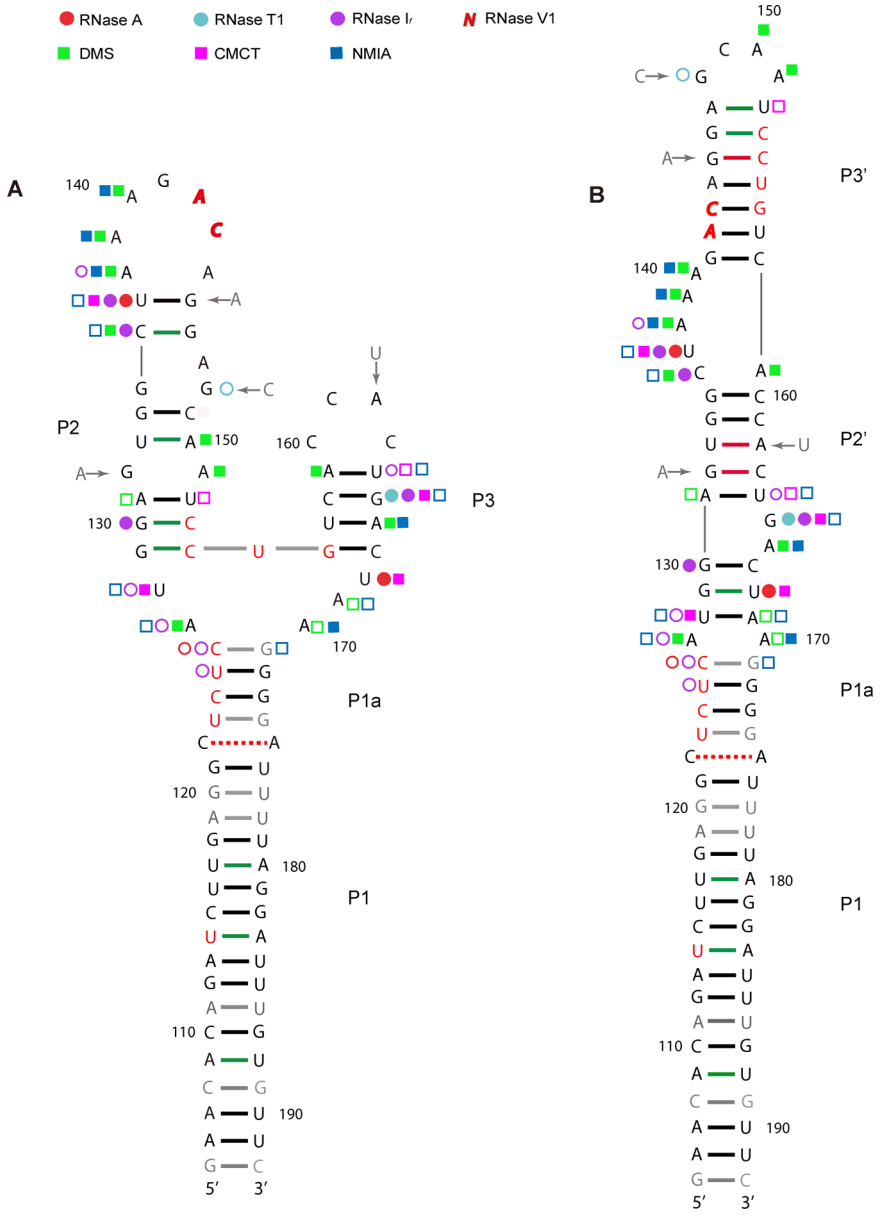


Figure 5. Mapping results of HPmutant. (A) The multibranch model. (B) The hairpin model. Color annotation is the same as in Figure 3. Mutations are indicated by arrows. The predicted free energies of folding, ΔG°_{37}, for the multibranch and hairpin models are −25.1 and −40.6 kcal/mol, respectively.

accessibility of RNA induced by higher KCl concentration was also observed in a conserved RNA structure in the NS gene.[27]

NMR Spectra Are Consistent with the Multibranch Structure. The consensus sequence of the wild-type RNA was cut after G121−U176, and two GC pairs plus a 3′ dangling A were added after G121−U176 to stabilize P1a (Figure 6A). NMR spectra were taken for this 61 nucleotide RNA in 20 mM KH_2PO_4 (pH 6.5), 80 mM KCl, and 0.05 mM EDTA. Figure 6B shows the imino region of a 2D-NOESY spectrum acquired at 12 °C, with 100 ms mixing time. Base pair types of the imino resonances were identified on the basis of their proton chemical shifts as well as typical NOE cross peaks to amino and nonexchangeable protons. More definitive peak assignments would require isotope labeling, which is beyond the scope of this work.

NOE cross peaks are formed between imino protons in adjacent base pairs. Four helical walks consistent with the multibranch structure were observed for this RNA. No NOE connections to G119−C178, U123−G174, or U137−A145 were observed. These pairs are adjacent to loops or the end of the structure. In P1a, the imino cross peak between G121 and U176 appeared to be split, suggesting conformational exchange in this helix. The signal for G171 was weaker than that of G173, consistent with G171−C126 closing the multibranch junction. In P2, the imino cross peak between U152 and G130 was weak but could be observed with the contour level lowered. A131H2 and G130H1 gave a clear NOE peak (Figure S2), confirming the connection between G130−C153 and A131−U152. The imino cross peak between U133 and G134 was broad. In P3, no NOE connections were seen, which is consistent with A166 being hit by single-strand sensitive probes (Figure 3). P3 may be a dynamic region. Adding up to 4 mM $MgCl_2$ in the NMR buffer did not significantly affect the spectrum (Figure S3). Some NOE peaks appeared weaker, possibly because $MgCl_2$ made the multibranch structure more rigid.

dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

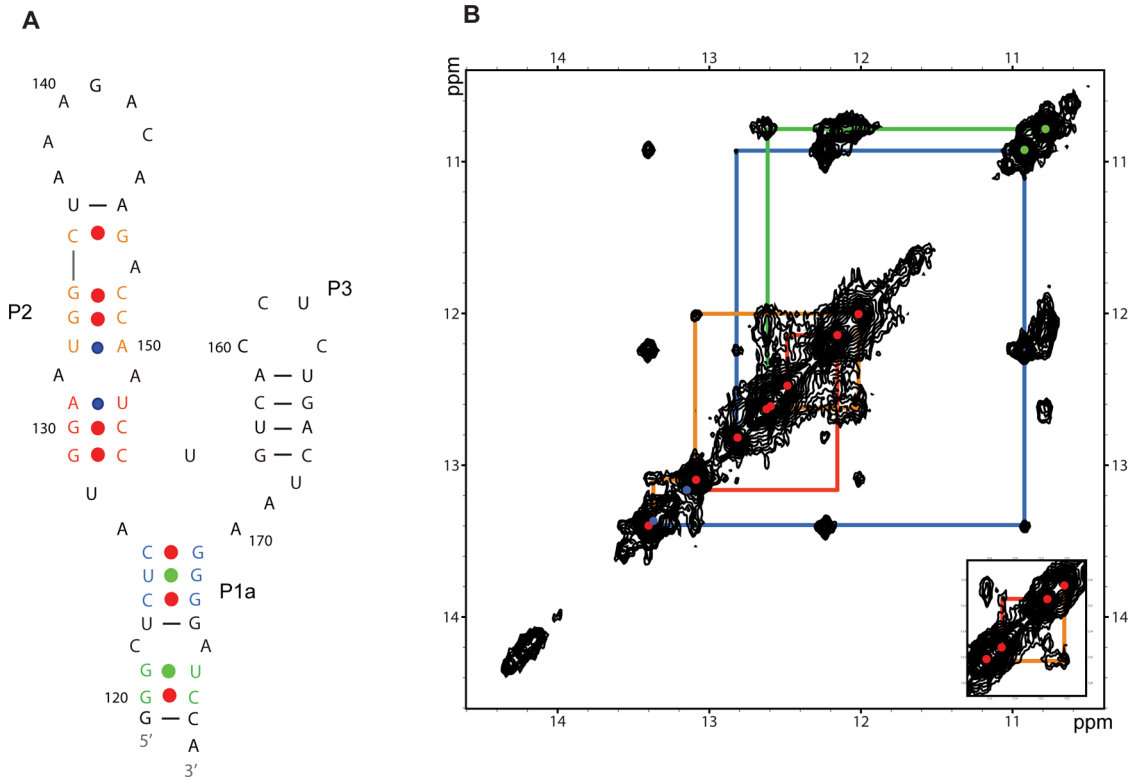


Figure 6. 2D-NOESY spectrum of the multibranch structure. (A) The 61 nucleotide construct of the multibranch structure. (B) The imino region of a 2D-NOESY spectrum recorded at 12 °C, with 100 ms mixing time. RNA was folded in 20 mM KH_2PO_4 (pH 6.5), 80 mM KCl, and 0.05 mM EDTA. Base pairs involved in observed helical walks are shown by colored dots (red = GC, blue = AU, green = GU). Four helical walks are shown by different color, which corresponds to the color of base pairs in panel A. The inset shows the region between 11.8−12.8 ppm with increased contour level to show the imino cross peaks of G129 to G130 and G135 to G146. The imino cross peak between G130 and U152 is not apparent from this plot but can be seen with contour level lowered. The evidence of connection between G130−C153 and A131−U152 is shown in Figure S2.

Optical Melting Experiments Reveal Similar Melting Profiles in 1 M NaCl and in 10 mM $MgCl_2$. Figure 7 shows melting profiles of the wild-type RNA in 100 mM KCl with/without 10 mM $MgCl_2$, in 300 mM KCl with 10 mM $MgCl_2$, and in 1 M NaCl without $MgCl_2$. Addition of 10 mM Mg^{2+} to 100 mM KCl increased the melting temperature by ∼14 °C. Increasing the K^+ concentration in the presence of Mg^{2+} had little effect. The melting profile of wild-type RNA in 1 M NaCl is similar to that in 100 (or 300) mM KCl with 10 mM $MgCl_2$. Similar agreement was found for melting of a cyclized group I intron.[48] Chemical mapping of the cyclized group I intron[48] and of the specificity domain of RNase P RNA[49] in 1 M NaCl and in 10 mM $MgCl_2$ have also shown good agreement. Evidently, 1 M NaCl is a reasonable approximation for buffers containing Mg^{2+}. This supports the common practice of using thermodynamic parameters measured in 1 M NaCl[50−52] to predict RNA secondary structures in the presence of Mg^{2+}.[53−55] For the RNA sequence studied here, free energy minimization with 1 M NaCl parameters and no experimental restraints predicts all the base pairs shown in Figure 3A when slippage of ±1 nucleotide is allowed.

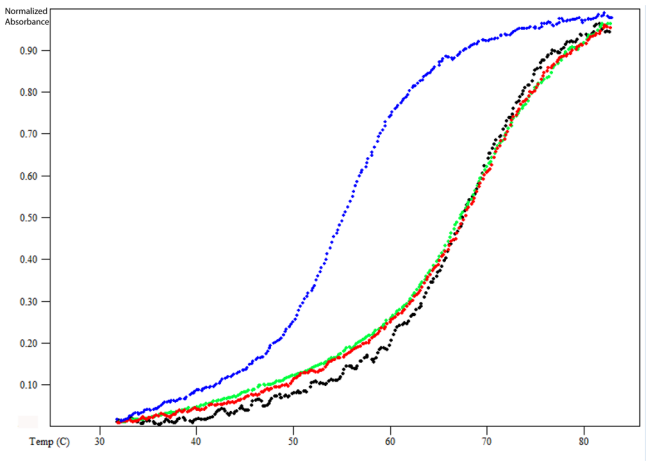


Figure 7. Optical melting curves for the wild-type RNA. The absorbance was measured at 280 nm in 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, and 100 mM KCl (blue line); 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, 100 mM KCl, and 10 mM $MgCl_2$ (black line); 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, 300 mM KCl, and 10 mM $MgCl_2$ (green line); and 20 mM sodium cacodylate (pH 7.0), 0.5 mM EDTA, and 1 M NaCl (red line).

DISCUSSION

Functional Implications of the Multibranch Structure. On the basis of predicted thermodynamics, sequence comparison, and bioinformatics, the multibranch structure was predicted to be one of the most thermodynamically stable and conserved structures in M1 mRNA.[13] The experimental data reported here confirms these predictions. This agreement contrasts with a similar 68 nucleotide region that is 51 nucleotides downstream from the 5′ splice site of segment 8. In that case, the bioinformatics approach predicted a multibranch model, but chemical and isoenergetic mapping revealed a hairpin structure in solution.[27]

dx.doi.org/10.1021/bi500611j | *Biochemistry* 2014, 53, 5236−5248

The multibranch structure in segment 7 is typically only 79 nucleotides downstream from the 5′ splice site for M2 mRNA (Figure 1). Intriguingly, in all of the spliced segments in influenza A, B, and C, unusually stable and conserved structures were predicted at or near the splice sites,[13,14,25] which indicates that these structures may play common roles in the regulation of influenza splicing.

The average canonical base pair conservation of this multibranch structure is 93.6% (Figure 1 and Table S1). In particular, the hairpin loops of P2 and P3 have most nucleotides with greater than 97% conservation. Although the base pairs in P1 and P1a are less conserved than those in P2 and P3, roughly one-third of the mutations retain canonical pairing. When mutations lead to noncanonical pairs, about one-third are CA pairs (Table S1). CA pairs are isosteric with GU wobble pairs, can have a high pK_a,[56] and preserve the A form RNA helix.[57] In general, variations in sequences and structures of mRNA are highly restricted in coding areas with cis-acting functions, presumably because of the necessity to maintain both protein and RNA structures.[58] The special conservation of this multibranch structure suggests that this motif is functionally important.

mRNA secondary structures can play roles in the modulation of splicing, for example, by hiding or revealing splice sites and other regulatory elements[59] or by modifying the spatial distance between cis-acting elements.[60] Splicing can also be regulated via protein-induced RNA conformational switching[61,62] or small molecule binding.[63,64] Many alternative splicing processes are associated with RNA secondary structure formation in pre-mRNA.[65−68] Previous studies have postulated roles for RNA secondary structure in the regulation of splicing in influenza and other viruses.[69,70]

While not present in the consensus sequence studied here, the basal helix of the multibranch structure in many influenza A strains contains the M4 mRNA 5′ splice site, GAG/GUUCUC (nucleotides 118−126), where the slash represents the splice site.[16] Additionally, there is a putative human intronic splicing enhancing sequence,[71] GGGGAUU (nucleotides 171−177), within this structure. In both cases, the regulatory sequences are embedded in secondary structure, where they would be expected to be less accessible to splicing factors.[68,72] Interestingly, at the 3′ splice site of the M2 mRNA, a pseudoknot is formed, which also buries splicing regulatory sequences in structure.[24,73] Thus, likely inhibitory cis-regulatory structures are at or near both splice sites of the M2 mRNA. Such structures could facilitate temporal control of M2 production, which occurs late in infection.[20]

Mutagenesis studies carried out on segment 7 of influenza A virus in the Shih group[74] are supportive of the proposed role for the multibranch structure in modulating alternative splicing of segment 7. The sequence used in their study contains the M4 mRNA 5′ splice site GAG/GUUCUC (nucleotides 118−126). When G120 was mutated to A, the virus was still viable, but the amounts of alternatively spliced products were changed, and the viral growth rate was attenuated. Similar results were observed when G121 was mutated to A. In contrast, when G121 was mutated to C or U, the virus could not be rescued. These results are consistent with the structure shown in Figure 3A, where G120 → A and G121 → A mutations maintain base pairing, but G121 → C/U disrupts base pairing. The mutation results thus also support the proposal that this multibranch conformation is important for regulating alternative splicing of segment 7. Changing this RNA structure may affect the alternative splicing process or even diminish viral viability. This has significance for designing attenuated live-virus vaccines.[75] The structure presented in Figure 3A provides a basis for further tests of functions for this motif.

Potential Tertiary Interactions and Dynamics in the Multibranch Structure. Three-way multibranch loops are common RNA secondary structural motifs, which are important in organizing tertiary interactions in large molecules.[76] They have been extensively studied in Varkud satellite ribozymes,[77] the hepatitis C virus internal ribosome entry site,[78−80] hammerhead ribozymes,[81,82] the P4−P6 domain of a group I intron,[83] and so forth. Comparison of the secondary structure in Figure 3A with 3D structures of other three-way multibranch loops suggests possible tertiary interactions.

Three-way multibranch loops usually fold by coaxial stacking of two of the three arms.[84,85] It is possible that small rearrangements in the multibranch structure (Figure 3A) would allow P1a to stack with P2 and connect with P3 by a U168 A169 A170 triloop closed by C167−G171. Both A residues in the triloop could form base triples with adjacent G129−C154 and G130−C153 pairs (Figure S4). CG/A base triples are sterically possible.[86] Such a long-range interaction was observed in RNase P and rRNAs[87−89] and is characteristic of UAA triloops.[90]

Extensive enzymatic and chemical mapping data are consistent with potential tertiary interactions and dynamics in this multibranch structure. For instance, the 3′ side of P3 was extensively hit by single-strand sensitive probes and bound by oligonucleotides (Figure 3A). Moreover, no NOESY helical walks were detected for P3 (Figures 6 and S3). These observations indicate that this helix is relatively flexible. It agrees with the potential 3D stacking, in which G156−C167 does not form, P3 becomes a relatively weak helix, and A166 becomes reactive to single-strand sensitive probes (Figure S4).

The other potentially flexible region of the multibranch structure is P1a, where most nucleotides on one side of the stem were hit by both RNases V1 and I_f (Figure 3). Also, no evidence for U123−G174 was observed, and more than one conformation was seen for G121−U176 in NOESY spectra (Figures 6 and S3). P1a may form a weak and dynamic helix, by shifting the pairing between CUCUC (nucleotides 122−126) and GGGGA (nucleotides 171−175), thus placing the putative intronic splicing enhancer GGGGAUU in different structural contexts. This dynamic region could serve as a good target for protein, small molecule, or oligonucleotide binding, as illustrated by medium microarray binding on both sides of P1a (Figure 3A). More sophisticated NMR techniques can be used to measure RNA dynamics,[91] as exemplified by some studies conducted on riboswitches.[92,93] For instance, $^{1}H/^{15}N$-heteronuclear exchange-sensitive NMR allows the detection of structural changes occurring within the time frame of ^{15}N-longitudinal relaxation.[94]

In P2, A131, A132, A150, and A151 were hit by DEPC only when the RNA was folded without Mg^{2+}. None reacted with NMIA to a detectable level, but A132 and A151 reacted with DMS in the presence and absence of Mg^{2+} (Figure 3A). High reactivities with DMS and negligible reactivities with NMIA of adenosines suggest that the Hoogesteen edge of the adenosines are buried and nucleobases are stacked on both faces.[95] It is possible that A132 and A151 pair in a sheared configuration (*trans* Hoogsteen/Sugar-edge), with two hydrogen bonds forming from the two amino protons of one adenosine to N3 and O2′ of the other adenosine, respectively.[83,96] In this way, the N 1s of A132 and A151 are available to be modified by DMS, but the ribose groups may not sample a conformation necessary for

reactions with NMIA. In the absence of Mg^{2+}, A132 and A151 can be accessed by DEPC, possibly because of lack of tertiary interaction and relatively weak stacking with adjacent AU pairs. In the presence of Mg^{2+}, however, long-range contacts with other nucleotides may occur in the position of the sheared AA pair, thus protecting N7 of A residues in this region. The A131H2 to U152H3 and the U133H3 to A150H2 cross peaks are both weak in NOESY spectra (Figures 6, S2, and S3), suggesting that the imino protons of the four residues are exchanging with water. It is possible that the two adenosines in the sheared AA pair are rapidly exchanging positions.[97,98]

An imino cross peak between G146 and G135 was observed in NOESY spectra (Figures 6 and S3). A147 may form a base triple with C136−G146 by interacting with its minor groove. Three hydrogen bonds may form to stabilize this interaction, including from N7 of A to H22 of G and from two amino protons of A to O2 and O2′ of C, respectively.[99] This kind of CG/A triple has been observed in rRNAs and a group I intron.[100−103]

Implications for Therapeutics against Influenza. Isoenergetic microarrays identified oligonucleotide binding sites in this multibranch conformation. Most of them are in flexible regions of the structure. Microarray probes can also bind to regions that are not hit by single-strand sensitive probes,[49,104] as apparent for nucleotide G182 (Figures 3 and S1 and Table S2). G182 is flanked by eight non-GC pairs on its 5′ side and four on its 3′ side. Evidently, the four modifications of probe 182 allow it to invade the P1 helix. Following strand invasion, U114 and U117 may further enhance hybrid stability by forming base triple interactions with A183 and A180, respectively. Thus, microarrays can reveal binding sites not apparent from secondary structure prediction or from enzymatic/chemical mapping.

Several approaches have been developed to target RNA secondary structure for therapeutics, including RNAi,[105] antisense RNA,[106] and aptamers.[107] More and more RNA structures that can be used for therapeutic targets in major human diseases have been found in recent years.[108,109] Influenza A is a significant public health threat, with evolved resistance to current vaccines and treatments. The potential biological function and sequence/structure conservation of the multibranch structure identified in this work makes this region a new attractive therapeutic target. Probe binding sites found in the microarray study are therefore promising for use in a chemical genetics approach to test for function and for designing oligonucleotides as potential therapeutics against influenza A infection. The loops identified may also be targetable with small molecules.[110]

■ ASSOCIATED CONTENT

Supporting Information

Hybridization results on isoenergetic microarrays for the wild-type RNA; table of NMR chemical shifts; NMR evidence of connection between G130−C153 and A131−U152 in the multibranch structure; imino region of 2D-NOESY spectrum of the multibranch structure in the presence of Mg^{2+}; putative tertiary interaction in the wild-type RNA; gel results of enzymatic and chemical mapping of the wild-type RNA; base pairing frequencies and nucleotide composition by alignment position; and microarray hybridization results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochester.edu.

Funding

This work was supported by NIH grant GM22939 (D.H.T.), NIH grant 1R03TW008739-01 (E.K. and D.H.T.), and Polish National Science Centre grants UMO-2013/08/M/NZ/01062 (E.K.) and NN301788440 (E.K.).

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

CMCT, 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-toluene sulfonate; DEPC, diethylpyrocarbonate; DMS, dimethyl sulfate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HEPES, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; LNA, locked nucleic acid; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; PAGE, polyacrylamide gel electrophoresis; SHAPE, selective 2′-hydroxyl acylation and primer extension

■ REFERENCES

(1) (2009) *Influenza (Seasonal) fact sheet N°211*, World Health Organization, Geneva, Switzerland.
(2) Johnson, N., and Mueller, J. (2002) Updating the accounts: global mortality of the 1918−1920 "Spanish" influenza pandemic. *Bull. Hist. Med. 76*, 105−115.
(3) CDC (2013) Interim adjusted estimates of seasonal influenza vaccine effectiveness—United States, February 2013. *Morb. Mortal. Wkly. Rep. 62*, 119−123.
(4) Hayden, F. G. (2010) Clinical aspects of pandemic 2009 influenza A (H1N1) virus infection. *N. Engl. J. Med. 362*, 1708−1719.
(5) Baz, M., Abed, Y., Papenburg, J., Bouhy, X., Hamelin, M. È., and Boivin, G. (2009) Emergence of oseltamivir-resistant pandemic H1N1 virus during prophylaxis. *N. Engl. J. Med. 361*, 2296−2297.
(6) Mai, L. Q., Wertheim, H. F. L., Duong, T. N., van Doorn, H. R., Hien, N. T., and Horby, P. (2010) A community cluster of oseltamivir-resistant cases of 2009 H1N1 influenza. *N. Engl. J. Med. 362*, 86−87.
(7) Stephenson, I., Democratis, J., Lackenby, A., McNally, T., Smith, J., Pareek, M., Ellis, J., Bermingham, A., Nicholson, K., and Zambon, M. (2009) Neuraminidase inhibitor resistance after oseltamivir treatment of acute influenza A and B in children. *Clin. Infect. Dis. 48*, 389−396.
(8) Bright, R. A., Shay, D. K., Shu, B., Cox, N. J., and Klimov, A. I. (2006) Adamantane resistance among influenza A viruses isolated early during the 2005−2006 influenza season in the United States. *JAMA, J. Am. Med. Assoc. 295*, 891−894.
(9) Beigel, J., and Bray, M. (2008) Current and future antiviral therapy of severe seasonal and avian influenza. *Antiviral Res. 78*, 91−102.
(10) Hayden, F. (2009) Developing new antiviral agents for influenza treatment: what does the future hold? *Clin. Infect. Dis. 48*, S3−S13.
(11) Cheong, H., Cheong, C., and Choi, B. (1996) Secondary structure of the panhandle RNA of influenza virus A studied by NMR spectroscopy. *Nucleic Acids Res. 24*, 4197−4201.
(12) Baudin, F., Bach, C., Cusack, S., and Ruigrok, R. W. (1994) Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J. 13*, 3158−3165.
(13) Moss, W. N., Priore, S. F., and Turner, D. H. (2011) Identification of potential conserved RNA secondary structure throughout influenza A coding regions. *RNA 17*, 991−1011.
(14) Dela-Moss, L. I., Moss, W. N., and Turner, D. H. (2014) Identification of conserved RNA secondary structures at influenza B and C splice sites reveals similarities and differences between influenza A, B, and C. *BMC Res. Notes 7*, 22.

(15) Lamb, R. A., Lai, C. J., and Choppin, P. W. (1981) Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. *Proc. Natl. Acad. Sci. U.S.A. 78*, 4170.
(16) Shih, S. R., Suen, P. C., Chen, Y. S., and Chang, S. C. (1998) A novel spliced transcript of influenza A/WSN/33 virus. *Virus Genes 17*, 179−183.
(17) Martin, K., and Helenius, A. (1991) Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. *Cell 67*, 117−130.
(18) Holsinger, L. J., Nichani, D., Pinto, L. H., and Lamb, R. A. (1994) Influenza A virus M2 ion channel protein: a structure−function analysis. *J. Virol. 68*, 1551−1563.
(19) Pinto, L. H., Holsinger, L. J., and Lamb, R. A. (1992) Influenza virus M2 protein has ion channel activity. *Cell 69*, 517−528.
(20) Shih, S. R., Nemeroff, M. E., and Krug, R. M. (1995) The choice of alternative 5′ splice sites in influenza virus M1 mRNA is regulated by the viral polymerase complex. *Proc. Natl. Acad. Sci. U.S.A. 92*, 6324−6328.
(21) Wise, H. M., Hutchinson, E. C., Jagger, B. W., Stuart, A. D., Kang, Z. H., Robb, N., Schwartzman, L. M., Kash, J. C., Fodor, E., and Firth, A. E. (2012) Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathog. 8*, e1002998.
(22) Robb, N. C., and Fodor, E. (2012) The accumulation of influenza A virus segment 7 spliced mRNAs is regulated by the NS1 protein. *J. Gen. Virol. 93*, 113−118.
(23) Shih, S. R., and Krug, R. M. (1996) Novel exploitation of a nuclear function by influenza virus: the cellular SF2/ASF splicing factor controls the amount of the essential viral M2 ion channel protein in infected cells. *EMBO J. 15*, 5415−5427.
(24) Moss, W. N., Dela-Moss, L. I., Kierzek, E., Kierzek, R., Priore, S. F., and Turner, D. H. (2012) The 3′ splice site of influenza A segment 7 mRNA can exist in two conformations: a pseudoknot and a hairpin. *PLoS One 7*, e38323.
(25) Gultyaev, A., Heus, H., and Olsthoorn, R. (2007) An RNA conformational shift in recent H5N1 influenza A viruses. *Bioinformatics 23*, 272−276.
(26) Bao, Y., Bolotov, P., Dernovoy, D., Kiryutin, B., Zaslavsky, L., Tatusova, T., Ostell, J., and Lipman, D. (2008) The influenza virus resource at the National Center for Biotechnology Information. *J. Virol. 82*, 596−601.
(27) Priore, S. F., Kierzek, E., Kierzek, R., Baman, J. R., Moss, W. N., Dela-Moss, L. I., and Turner, D. H. (2013) Secondary structure of a conserved domain in the intron of influenza A NS1 mRNA. *PLoS One 8*, e70615.
(28) Woodson, S. A., and Koculi, E. (2009) Analysis of RNA folding by native polyacrylamide gel electrophoresis. *Methods Enzymol. 469*, 189−208.
(29) *RNA structure−function protocols: alkaline hydrolysis, RNA sequencing and RNA structure analyses with nucleases*, Invitrogen, Carlsbad, CA.
(30) Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2006) Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc. 1*, 1610−1616.
(31) Ziehler, W. A., Engelke, D. R. (2001) Probing RNA structure with chemical reagents and enzymes, *Current Protocols Nucleic Acid Chemistry* Chapter 6, Unit 6.1., Wiley, New York.
(32) Fratczak, A., Kierzek, R., and Kierzek, E. (2009) LNA-modified primers drastically improve hybridization to target RNA and reverse transcription. *Biochemistry 48*, 514−516.
(33) Fratczak, A., Kierzek, R., and Kierzek, E. (2011) Isoenergetic microarrays to study the structure and interactions of DsrA and OxyS RNAs in two- and three-component complexes. *Biochemistry 50*, 7647−7665.
(34) Abràmoff, M. D., Magalhaes, P., and Ram, S. (2004) Image processing with ImageJ. *Biophoton. Int. 11*, 36−43.
(35) Kierzek, E., Kierzek, R., Moss, W. N., Christensen, S. M., Eickbush, T. H., and Turner, D. H. (2008) Isoenergetic penta- and hexanucleotide microarray probing and chemical mapping provide a secondary structure model for an RNA element orchestrating R2 retrotransposon protein function. *Nucleic Acids Res. 36*, 1770−1782.
(36) Turner, D. H., and Mathews, D. H. (2010) NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res. 38*, D280−282.
(37) Reuter, J., and Mathews, D. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinf. 11*, 129.
(38) Kao, C., Zheng, M., and Rüdisser, S. (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA 5*, 1268−1272.
(39) Rio, D. C. (2013) Expression and purification of active recombinant T7 RNA polymerase from *E. coli*, *Cold Spring Harbor Protocols*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
(40) Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(41) Goddard, T. D., and Kneller, D. G. (2004) *SPARKY 3*, University of California, San Francisco, CA.
(42) McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)2 by two-dimensional NMR and simulated annealing. *Biochemistry 35*, 14077−14089.
(43) Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J., and Ehresmann, B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res. 15*, 9109−9128.
(44) Meador, J., Cannon, B., Cannistraro, V. J., and Kennell, D. (1990) Purification and characterization of *Escherichia coli* RNase I. *Eur. J. Biochem. 187*, 549−553.
(45) Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc. 127*, 4223−4231.
(46) Kierzek, E., Ciesielska, A., Pasternak, K., Mathews, D. H., Turner, D. H., and Kierzek, R. (2005) The influence of locked nucleic acid residues on the thermodynamic properties of 2′-O-methyl RNA/RNA heteroduplexes. *Nucleic Acids Res. 33*, 5082−5093.
(47) Pasternak, A., Kierzek, E., Pasternak, K., Fratczak, A., Turner, D. H., and Kierzek, R. (2008) The thermodynamics of 3′-terminal pyrene and guanosine for the design of isoenergetic 2′-O-methyl-RNA-LNA chimeric oligonucleotide probes of RNA structure. *Biochemistry 47*, 1249−1258.
(48) Jaeger, J. A., Zuker, M., and Turner, D. H. (1990) Melting and chemical modification of a cyclized self-splicing group I intron: similarity of structures in 1 M sodium, in 10 mM magnesium and in the presence of substrate. *Biochemistry 29*, 10147−10158.
(49) Liang, R., Kierzek, E., Kierzek, R., and Turner, D. H. (2010) Comparisons between chemical mapping and binding to isoenergetic oligonucleotide microarrays reveal unexpected patterns of binding to the *Bacillus subtilis* RNase P RNA specificity domain. *Biochemistry 49*, 8155−8168.
(50) Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol. 288*, 911−940.
(51) Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
(52) Xia, T., SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson−Crick base pairs. *Biochemistry 37*, 14719−14735.
(53) Jaeger, J. A., Turner, D. H., and Zuker, M. (1989) Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
(54) Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for

prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.
(55) Zuker, M., and Stiegler, P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res. 9*, 133−148.
(56) Wilcox, J. L., and Bevilacqua, P. C. (2013) pK_a shifting in double-stranded RNA is highly dependent upon nearest neighbors and bulge positioning. *Biochemistry 52*, 7470−7476.
(57) Lima, S., Hildenbrand, J., Korostelev, A., Hattman, S., and Li, H. (2002) Crystal structure of an RNA helix recognized by a zinc-finger protein: an 18-bp duplex at 1.6 A resolution. *RNA 8*, 924−932.
(58) Gog, J. R., Afonso, E. D. S., Dalton, R. M., Leclercq, I., Tiley, L., Elton, D., von Kirchbach, J. C., Naffakh, N., Escriou, N., and Digard, P. (2007) Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res. 35*, 1897−1907.
(59) Muro, A. F., Caputi, M., Pariyarath, R., Pagani, F., Buratti, E., and Baralle, F. E. (1999) Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. *Mol. Cell. Biol. 19*, 2657−2671.
(60) Deshler, J. O., and Rossi, J. J. (1991) Unexpected point mutations activate cryptic 3′ splice sites by perturbing a natural secondary structure within a yeast intron. *Gene. Devel. 5*, 1252−1263.
(61) Hönig, A., Auboeuf, D., Parker, M. M., O'Malley, B. W., and Berget, S. M. (2002) Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72. *Mol. Cell. Biol. 22*, 5698−5707.
(62) Warf, M. B., Diegel, J. V., Von Hippel, P. H., and Berglund, J. A. (2009) The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc. Natl. Acad. Sci. U.S.A. 106*, 9203−9208.
(63) Cheah, M. T., Wachter, A., Sudarsan, N., and Breaker, R. R. (2007) Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature 447*, 497−500.
(64) Winkler, W., Nahvi, A., and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature 419*, 952−956.
(65) Gahura, O., Hammann, C., Valentová, A., Půta, F., and Folk, P. (2011) Secondary structure is required for 3′ splice site recognition in yeast. *Nucleic Acids Res. 39*, 9759−9767.
(66) Meyer, M., Plass, M., Pérez-Valle, J., Eyras, E., and Vilardell, J. (2011) Deciphering 3′ ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing. *Mol. Cell 43*, 1033−1039.
(67) Shepard, P. J., and Hertel, K. J. (2008) Conserved RNA secondary structures promote alternative splicing. *RNA 14*, 1463−1469.
(68) Warf, M. B., and Berglund, J. A. (2010) Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem. Sci. 35*, 169−178.
(69) Abbink, T. E. M., and Berkhout, B. (2008) RNA structure modulates splicing efficiency at the human immunodeficiency virus type 1 major splice donor. *J. Virol. 82*, 3090−3098.
(70) Nemeroff, M. E., Utans, U., Kramer, A., and Krug, R. M. (1992) Identification of cis-acting intron and exon regions in influenza virus NS1 mRNA that inhibit splicing and cause the formation of aberrantly sedimenting presplicing complexes. *Mol. Cell. Biol. 12*, 962−970.
(71) McCarthy, E. M., and Phillips, J. A. (1998) Characterization of an intron splice enhancer that regulates alternative splicing of human GH pre-mRNA. *Hum. Mol. Genet. 7*, 1491−1496.
(72) Buratti, E., and Baralle, F. E. (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell. Biol. 24*, 10505−10514.
(73) Moss, W. N., Dela-Moss, L. I., Priore, S. F., and Turner, D. H. (2012) The influenza A segment 7 mRNA 3′ splice site pseudoknot/hairpin family. *RNA Biol. 9*, 1305−1310.
(74) Chiang, C., Chen, G. W., and Shih, S. R. (2008) Mutations at alternative 5′ splice sites of M1 mRNA negatively affect influenza A virus viability and growth rate. *J. Virol. 82*, 10873−10886.
(75) Cox, R. J., Brokstad, K. A., and Ogra, P. L. (2004) Influenza virus: immunity and vaccination strategies. Comparison of the immune response to inactivated and live, attenuated influenza vaccines. *Scand. J. Immunol. 59*, 1−15.
(76) Lilley, D. M. J. (2000) Structures of helical junctions in nucleic acids. *Q. Rev. Biophys. 33*, 109−159.
(77) Lipfert, J., Ouellet, J., Norman, D. G., Doniach, S., and Lilley, D. M. J. (2008) The complete VS ribozyme in solution studied by small-angle X-ray scattering. *Structure 16*, 1357−1367.
(78) Kim, I., Lukavsky, P. J., and Puglisi, J. D. (2002) NMR study of 100 kDa HCV IRES RNA using segmental isotope labeling. *J. Am. Chem. Soc. 124*, 9338−9339.
(79) Lukavsky, P. J., Kim, I., Otto, G. A., and Puglisi, J. D. (2003) Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol. 10*, 1033−1038.
(80) Ouellet, J., Melcher, S., Iqbal, A., Ding, Y., and Lilley, D. M. J. (2010) Structure of the three-way helical junction of the hepatitis C virus IRES element. *RNA 16*, 1597−1609.
(81) Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature 372*, 68−74.
(82) Scott, W. G., Finch, J. T., and Klug, A. (1995) The crystal structure of an aII-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell 81*, 991−1002.
(83) Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science 273*, 1678−1685.
(84) Donghi, D., Pechlaner, M., Finazzo, C., Knobloch, B., and Sigel, R. K. O. (2013) The structural stabilization of the κ three-way junction by Mg(II) represents the first step in the folding of a group II intron. *Nucleic Acids Res. 41*, 2489−2504.
(85) Lescoute, A., and Westhof, E. (2006) Topology of three-way junctions in folded RNAs. *RNA 12*, 83−93.
(86) Almakarem, A. S. A., Petrov, A. I., Stombaugh, J., Zirbel, C. L., and Leontis, N. B. (2012) Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res. 40*, 1407−1423.
(87) Brown, J. W. (1999) The ribonuclease P database. *Nucleic Acids Res. 27*, 314−314.
(88) Klein, D., Schmeing, T., Moore, P., and Steitz, T. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J. 20*, 4214−4221.
(89) Krasilnikov, A. S., Yang, X., Pan, T., and Mondragón, A. (2003) Crystal structure of the specificity domain of ribonuclease P. *Nature 421*, 760−764.
(90) Lee, J. C., Cannone, J. J., and Gutell, R. R. (2003) The lonepair triloop: A new motif in RNA structure. *J. Mol. Biol. 325*, 65−83.
(91) Rinnenthal, J., Buck, J., Ferner, J., Wacker, A., Furtig, B., and Schwalbe, H. (2011) Mapping the landscape of RNA dynamics with NMR spectroscopy. *Acc. Chem. Res. 44*, 1292−1301.
(92) Reining, A., Nozinovic, S., Schlepckow, K., Buhr, F., Furtig, B., and Schwalbe, H. (2013) Three-state mechanism couples ligand and temperature sensing in riboswitches. *Nature 499*, 355−359.
(93) Santner, T., Rieder, U., Kreutz, C., and Micura, R. (2012) Pseudoknot preorganization of the PreQ1 class I riboswitch. *J. Am. Chem. Soc. 134*, 11928−11931.
(94) Wenter, P., Bodenhausen, G., Dittmer, J., and Pitsch, S. (2006) Kinetics of RNA refolding in dynamic equilibrium by ^{1}H-detected ^{15}N exchange NMR spectroscopy. *J. Am. Chem. Soc. 128*, 7579−7587.
(95) Kladwang, W., Mann, T. H., Becka, A., Tian, S., Kim, H., Yoon, S., and Das, R. (2014) Standardization of RNA chemical mapping experiments. *Biochemistry 53*, 3063−3065.
(96) Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared A_{anti}·A_{anti} base pairs in a destabilizing 2 × 2 internal loop: The NMR structure of 5′(rGGCAAGCCU)$_2$. *Biochemistry 41*, 14969−14977.
(97) Chen, G., Kennedy, S. D., Qiao, J., Krugh, T. R., and Turner, D. H. (2006) An alternating sheared AA pair and elements of stability for a single sheared purine-purine pair flanked by sheared GA pairs in RNA. *Biochemistry 45*, 6889−6903.

(98) Van Nostrand, K. P., Kennedy, S. D., Turner, D. H., and Mathews, D. H. (2011) Molecular mechanics investigation of an adenine−adenine non-canonical pair conformational change. *J. Chem. Theory Comput. 7*, 3779−3792.
(99) Popenda, L., Bielecki, L., Gdaniec, Z., and Adamiak, R. W. (2009) Structure and dynamics of adenosine bulged RNA duplex reveals formation of the dinucleotide platform in the C:G-A triple. *ARKIVOC 3*, 130−144.
(100) Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J., and Strobel, S. A. (2004) Crystal structure of a self-splicing group I intron with both exons. *Nature 430*, 45−50.
(101) Petry, S., Brodersen, D. E., Murphy, F. V. I. V., Dunham, C. M., Selmer, M., Tarry, M. J., Kelley, A. C., and Ramakrishnan, V. (2005) Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell 123*, 1255−1266.
(102) Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D. (2005) Structures of the bacterial ribosome at 3.5 Å resolution. *Science 310*, 827−834.
(103) Schuwirth, B. S., Day, J. M., Hau, C. W., Janssen, G. R., Dahlberg, A. E., Cate, J. H. D., and Vila-Sanjurjo, A. (2006) Structural analysis of kasugamycin inhibition of translation. *Nat. Struct. Mol. Biol. 13*, 879−886.
(104) Kierzek, E., Christensen, S. M., Eickbush, T. H., Kierzek, R., Turner, D. H., and Moss, W. N. (2009) Secondary structures for 5′ regions of R2 retrotransposon RNAs reveal a novel conserved pseudoknot and regions that evolve under different constraints. *J. Mol. Biol. 390*, 428−442.
(105) Dorsett, Y., and Tuschl, T. (2004) siRNAs: applications in functional genomics and potential as therapeutics. *Nat. Rev. Drug Discovery 3*, 318−329.
(106) Crooke, S. T. (2004) Antisense strategies. *Curr. Mol. Med. 4*, 465−487.
(107) Nimjee, S. M., Rusconi, C. P., and Sullenger, B. A. (2005) Aptamers: an emerging class of therapeutics. *Annu. Rev. Med. 56*, 555−583.
(108) Shen, M., Bellaousov, S., Hiller, M., de La Grange, P., Creamer, T. P., Malina, O., Sperling, R., Mathews, D. H., Stoilov, P., and Stamm, S. (2013) Pyrvinium pamoate changes alternative splicing of the serotonin receptor 2C by influencing its RNA structure. *Nucleic Acids Res. 41*, 3819−3832.
(109) Singh, N. N., Lawler, M. N., Ottesen, E. W., Upreti, D., Kaczynski, J. R., and Singh, R. N. (2013) An intronic structure enabled by a long-distance interaction serves as a novel target for splicing correction in spinal muscular atrophy. *Nucleic Acids Res. 41*, 8144−8165.
(110) Childs-Disney, J., Wu, M., Pushechnikov, A., Aminova, O., and Disney, M. (2007) A small molecule microarray platform to select RNA internal loop−ligand interactions. *ACS Chem. Biol. 2*, 745−754.