Published in final edited form as:

Biopolymers. 2013 December ; 99(12): 1114–1123. doi:10.1002/bip.22299.

A Personal Perspective on Chemistry-Driven RNA Research

Ronald Micura, Christoph Kreutz, and Kathrin Breuker

Institute of Organic Chemistry, Center for Molecular Biosciences (CMBI), Center for Chemistry and Biomedicine (CCB), University of Innsbruck, Innrain 80-82, Innsbruck 6020, Austria

Abstract

In this mini review, we discuss how our understanding of ribonucleic acid (RNA) properties becomes significantly deepened when a broad range of modern chemical and biophysical methods is applied. We span our perspective from RNA solid-phase synthesis and site-specific labeling to single-molecule fluorescence-resonance-energy-transfer imaging and NMR spectroscopy approaches to explore the dynamics of RNA over a broad timescale. We then move on to Fourier-transform-ion-cyclotron-resonance mass spectrometry (FT-ICR-MS) as a powerful technique for RNA sequencing and modification analysis. The novel methodological developments are discussed for selected biological systems that include the thiamine-pyrophosphate riboswitch, HIV and ribosomal A-site RNA, and transfer RNA.

Keywords

labeling; modification; structure; dynamics; chemical synthesis

INTRODUCTION

The growing number of newly discovered noncoding RNAs raises the demand to explore and comprehend their distinct roles in cells.^{1–4} From a chemist's point of view, a major challenge is to propose and interrogate the molecular mechanisms behind RNA function, ideally within a structural framework and by taking the complex dynamic features of RNA into account. In recent years, advanced chemical and biophysical approaches have emerged to tackle RNA research related questions. They include novel tools in solid-phase RNA synthesis and labeling^{5,6} to generate RNA derivatives, for example, single-molecule studies, powerful relaxation based or real-time NMR methods to investigate RNA-ligand interactions,^{7–10} and sophisticated mass spectrometric methods for RNA top-down sequencing and analysis of RNA modifications.¹¹ Breaking new ground in methodological developments for RNA characterization and manipulation is a prerequisite to create a solid foundation to propose and evaluate mechanistic models of RNA function, and hence, to further elucidate RNA's role in living systems. Under these aspects, and from a personal

© 2013 Wiley Periodicals, Inc.

Correspondence to: R. Micura; ronald.micura@uibk.ac.at (or) C. Kreutz; christoph.kreutz@uibk.ac.at (or) K. Breuker; kathrin.breuker@uibk.ac.at.

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

standpoint, we discuss three directions of ongoing research in our laboratories. Thereby, one unifying feature is the involvement of chemically synthesized and site-specifically modified RNA. The synthetic strategies to achieve such RNAs require automated RNA solid-phase synthesis (recently reviewed in Refs. 5 and 6). This undertaking also includes the chemical synthesis of building blocks as nucleoside phosphoramidites that carry the desired modification or a tether moiety to which the actual label of interest is attached at a later stage. The solid-phase approach is currently practicable for RNA constructs comprising up to 50–60 nucleotides, as the quality and yield of the biopolymer usually decreases significantly from this length on. However, in many cases the chemically synthesized RNA fragments can be enzymatically ligated by using T4 RNA ligase or T4 DNA ligase and a splint oligonucleotide to obtain the biologically relevant target RNA.¹² The specific synthetic strategies for modified RNA are addressed in the individual sections in context to biological research questions.

RNA Folding and Ligand Recognition Exemplified by the Thiamine Pyrophosphate (TPP) Riboswitch

One of the most intriguing class of RNA with respect to structure, function, interaction with a small molecule, and the underlying folding and ligand-binding dynamics refers to so-called mRNA riboswitches. Their fundamental role, mostly in bacterial gene regulation, is based on the specific, concentration-dependent interaction of a low-molecular weight ligand with the RNA. Ligand binding to the mRNA aptamer domain turns genes “ON” or “OFF” by changing the secondary structure of a downstream expression platform.

A prominent representative of these noncoding RNAs is the TPP riboswitch.^{13–17} Soon after its discovery, crystal structures revealed that a complex three-way junctional fold spans the ligand TPP from one double helical “sensor arm” to the other while the third helix represents the so-called “regulatory” helix that transduces TPP binding to a downstream structural signal for the gene regulation apparatus (Figure 1A).^{18–20} The X-ray structure “snapshot” of the ligand-bound RNA aptamer unveiled that TPP is completely encapsulated by the RNA. How the ligand can enter the binding pocket, or makes initial contacts with the scaffold, remained unexplained. Therefore, a number of studies using diverse chemical, biochemical, and biophysical methods have been applied to explore the RNA ligand recognition process from a dynamics point of view. Enzymatic and chemical probing,²¹ mutational studies,²² as well as isothermal titration calorimetry²³ highlighted that alternative folds play a critical role on the cotranscriptional folding pathway that engages defined polymerase pause sites.²⁴ Furthermore, selective-2'-hydroxyl-acylation-analyzed-by-primer-extension experiments suggested that the well-defined tertiary interaction between the two sensor arms (P3/L5) visible in the X-ray structure is not present in the free RNA state, even at high concentrations of Mg^{2+} ions.^{25,26} These observations were further supported by small angle X-ray scattering measurements that suggested structural compaction in the presence of Mg^{2+} and a further significant compaction when TPP ligand became available.^{27–29} Moreover, an optical tweezer single-molecule study in which force was applied to the 5' and 3' ends of the aptamer's regulatory helix shed light on the energy landscape for TPP riboswitch folding.³⁰ These data indicated that in the absence of TPP, all helical secondary structure elements are folded but no interdomain tertiary contacts are present.

Investigations in the Micura group on the TPP riboswitch started with a spectroscopic approach that made use of 2-aminopurine (2AP) fluorescence³¹ and that was termed the 2AP fold approach.³² It was reasoned that aptamers with strategically positioned 2AP-nucleoside replacements should provide information on the time-dependent course of secondary and/or tertiary structure formation in response to ligand addition. By generating a large number of TPP aptamer variants that were labeled at or close to important tertiary contact sites, and by determining the rate constants based on the time-dependent change in 2AP fluorescence that reflects either increase or decrease of base stacking, gave an unexpected picture: the fastest rates were observed within the secondary structure elements (P5-L3 of the sensor arms) that are most distant from the three-way junction, followed by medium rates for the region around the binding pocket; the slowest response accounted for conformational adjustments in the three-way junction and the regulatory helix. Thus, an “outside-in” folding trend was consistent with this observation. Although it would have been tempting to propose preorganization of the P5-L3 interaction, this could not be argued conclusively because 2AP probes report on the local environment only, and conformational adaptations of that kind can be uncoupled from global structural folding events, as for example, the collapse of the two helical sensor arms of the TPP riboswitch.

To further illuminate the global folding process and the intrinsic dynamics of the TPP riboswitch, one of us has recently performed an extensive fluorescence-resonance-energy-transfer (FRET) study at the single molecule level.³³ Two regions of the molecule are of major interest (Figures 1A–1C): First, the regulatory helix P1, which overlaps the expression platform and transduces aptamer-ligand binding into expression platform “signaling,” and second, the sensor arms that have no obvious contacts with each other in the ligand-unbound form but seem to stick together in the bound form. In short, the surprising observation of the smFRET experiments was that an unexpectedly high population of the free RNA aptamer corresponds to already closed conformations of the regulatory helix (Figure 1D, left). Of critical importance, this “seemingly closed” helix is significantly dynamic and opens up on a millisecond timescale, with the intrinsic dynamics keeping the helix amenable for secondary structure rearrangement that is required to adapt the fold responsible for gene ON-regulation.³³ In case of ligand presence, the regulatory helix becomes “frozen” to adapt the fold that is responsible for gene OFF-regulation (Figure 1D, right).

The second surprising behavior was found for the sensor arms. Although the crystal structure previously implied a very rigid interaction P3-L5 between the forearms in the TPP-aptamer complex, in solution, only about 30% of the RNA molecules were observed in closed conformations of P3 and L5 (Figure 1D, right). This region remains indeed very dynamic in the ligand-bound form and likely represents the ligand entry/exit site to the binding pocket. This finding became more understandable when an aptamer construct with a labeling pattern sensitive for the upper arm dynamics of the “sensor arms” revealed that they are rigidly prefolded in the free form (Figure 1D, left) and hence provide the necessary structural scaffold for initial recognition of the small molecule TPP by the RNA.

Synthetic Aspects—To be able to conduct 2AP and smFRET studies of the kind discussed here, a strong background in synthetic organic chemistry is required for the synthesis of appropriately labeled RNA variants. After the first 2AP fold investigations were

published in the year 2007,^{31,34} this undertaking was regarded as “a tour de force.”³⁵ The required RNA constructs are challenging to synthesize and require enzymatic ligation to introduce the label into a 100 nt RNA.³⁶ However, the wealth of kinetic and thermodynamic information attainable by derivatives with site-specific labels provided novel molecular insights into the recognition process of preQ₁-I,³⁷ SAM-II,³⁸ and very recently TPP riboswitches.^{31,33} Moreover, even larger systems are accessible if enzymatic ligations of chemically synthesized RNA strands to RNA transcripts are performed.³⁹⁻⁴¹ Alternative approaches make use of circularly permuted RNA in combination with synthetic RNAs and hence allow for site-specific labeling of large RNAs and RNA machineries, such as self-splicing group II introns⁴² and ribosomal particles.^{43,44} Other upcoming and exciting developments for RNA labeling concern the use of DNA-templated⁴⁵ and deoxyribozyme (DNA enzyme) promoted labeling of internal positions of large RNAs.⁴⁶⁻⁴⁸

For many RNA labeling strategies, the modification incorporated into the nucleic acid provides the functionality, which is required for subsequent conjugation to the actual label of interest. Beside the classical NHS ester chemistry, tuning of the conditions for Click reactions made this chemistry recently more and more popular in context with RNA,⁴⁹ including exchanged alkyne and azide patterns for inverse Click bioconjugation.^{50,51}

A further aspect of synthetic chemistry in context with smFRET investigations refers to the improvement of the photophysical properties of the dyes involved. A very recent work describes the concept of covalent attachment of triplet state quencher units to the dye core (Cy3 and Cy5) which results in significantly reduced blinking and photobleaching and thus enables extended experimental time courses at high time resolution.^{52,53}

Resolving Rare RNA Conformational States Using Recent Advances in NMR Spectroscopy

The development of relaxation dispersion NMR methods to monitor dynamic phenomena occurring on a microsecond to millisecond time scale is one of the most recent developments in biomolecular NMR spectroscopy.⁵⁴⁻⁵⁶ The data from these experiments revealed that the function of proteins is often encoded in weakly and transiently populated excited states (ES). For example, such high-energy conformations were found to be essential for catalysis and ligand binding in proteins.⁵⁷ For the isolated amide ¹⁵N-¹H spin pair in proteins the methods can be applied in a straight-forward manner as interfering scalar couplings with other NMR active heteronuclei are absent. Advances in the isotope labeling protocols (e.g., IleLeuVal-¹³C-methyl group labeling) recently allowed ¹³C-based relaxation dispersion experiments.^{58,59} Together with the measurement of residual dipolar couplings solution NMR spectroscopy can nowadays be used to elucidate the structural features of the protein's ES with atomistic resolution giving unrivaled insights into the fundamental mechanisms of biological function.^{60,61}

Whereas the isotope labeling protocols and NMR methodologies for proteins are more or less well established, there is still a lack of comparable methods for nucleic acids.⁶² But for all that, the remaining issues to apply relaxation dispersion methods to nucleic acid systems were addressed and recently published works suggest that the high energy ESs found in DNA and RNA are intricately involved in their respective biological function.⁶³⁻⁶⁵ In the

following, these recent developments are highlighted, which have led to unique insights into functional dynamics of nucleic acids.

An appropriate proton-heteronucleus (^{13}C or ^{15}N) spin pair is fundamentally required for relaxation dispersion experiments. In proteins, the amide protons are suitable readout nuclei of ^{15}N relaxation data as they very slowly exchange with bulk water protons. In nucleic acid systems, the so-called imino protons ($\text{N}^1\text{-H}$ in guanosine and $\text{N}^3\text{-H}$ in uridine) also represent suitable spin pairs, but the accelerated exchange rate with bulk water protons can lead to strongly broadened or even unobservable signals. Thus, the ^{15}N relaxation dispersion experiments are restricted to study dynamics in well-defined double helical regions, where the imino protons are protected from chemical exchange with water. Using ^{15}N - and ^{13}C -relaxation dispersion methods, an unprecedented complexity of structural flexibility in canonical duplex DNA was found.^{64,65} Owing to the fast chemical exchange of single stranded imino protons with bulk water one has to resort to ^{13}C - ^1H spin systems if solvent accessible regions of nucleic acids are probed for micro- to millisecond dynamics. Currently, only uniformly labeled ^{13}C -, ^{15}N -ribonucleotide triphosphates (rNTPs) are commercially available that are suitable for assignment purposes and subsequent solution structure determination protocols.⁶² However, the heteronuclear scalar couplings in these compounds impair with a straightforward application of relaxation dispersion experiments. This problem was addressed recently either by advancing the isotope labeling protocols or by novel NMR spectroscopic methods.^{63,66,67} Al-Hashimi and coworkers designed an elegant NMR spectroscopic work-around by using a selective carbon $\text{R}_{1\rho}$ relaxation dispersion experiment.⁶⁷ The off-resonance ^{13}C - $\text{R}_{1\rho}$ NMR experiment uses Hartmann-Hahn polarization transfers to selectively excite specific spins and collect data in a 1-dimensional manner either making it suitable to apply to unlabeled nucleic acids or uniformly labeled RNA/DNA samples. Using the selective excitation method, experimental artifacts originating from scalar couplings are eliminated. The methodology was recently used to visualize transient low-populated structures in biologically relevant RNAs.⁶⁸ The HIV TAR apical loop was found to fluctuate between a ground state (87%) and a transiently formed suboptimal conformation (13%). From the ^{13}C nucleobase and sugar chemical shift data of the ES obtained from a two state fit of the $\text{R}_{1\rho}$ relaxation curves a loop conformation resembling a UUCG tetraloop with a flipped-out guanosine residue was postulated (Figure 2A). The ES structure was further corroborated by a “mutate-and-chemical-shift-fingerprint” strategy. In this process, the conformational ensemble is biased either towards or off the ES by nucleotide mutations. The ES structure is not able to bind the viral transactivator proteins Tat and human cyclin T1, which are involved in the transcription activation of the HIV-1 genome. Thus, the ES structure of the TAR apical loop could be involved in the down-regulation of the transactivation of the HIV genome or in the release mechanism of TAT and cyclin T1. Along these lines, the biological functions encoded in the excited structures of the ribosomal A-site RNA and the HIV-1 stem loop were investigated. For the A-site RNA two adenosine (A1492 and A1493) residues were found to form non-canonical base pairs in the ES structure making them unavailable to decode the mRNA. Furthermore, the ES conformation affects residues (C7, G94, and U95) involved in protein recognition and formation of the B2a intersubunit crossbridge (Figure 2B). Even more interesting, the HIV-1 stem loop 1 RNA shows two processes, which represent excursions to two mutually

exclusive ESs. In the ES structures, the bulge residues (A27, G28, and G29) migrate either up- (ES1, 9%) or downward (ES2, 2%). The ES structures and the ground state conformation ensemble define a moving zipper in which the bulge residues displace base pairs in the upper and lower helical region. This base pair zipping is presumed to be involved in homoduplex formation of the SL1 (Figure 2C).

The above described approach using selective ^{13}C -excitation, however, comes to limitations if carbon signals are located in crowded or overlapping chemical shift regions (i.e., in larger RNAs comprising more than 35 nucleotides like riboswitch aptamer domains) and thus, the selective excitation conditions are not fulfilled anymore. In this case, the site-specific incorporation of ^{13}C -isotopes into ribonucleotides (and deoxyribonucleotides) represents an inevitable synthetic challenge to apply relaxation dispersion experiments to DNA and RNA.

Synthetic Aspects—Several methods to produce sufficient amounts of RNA for NMR spectroscopic studies exist. The most widely used method relies on enzymatic synthesis by T7 RNA polymerase using a DNA template and rNTPs with ^{13}C - and/or ^{15}N -labels.⁶⁹⁻⁷¹ The method can be used to introduce stable isotope label patterns on a nucleotide level by mixing unlabeled and the respective labeled rNTP. That is why several research groups are strongly focused on elucidating the bacterial biosynthetic pathway of rNTPs and on the identification of the metabolites used therein to produce tailor made rNTPs that can be used in relaxation dispersion experiments.^{63,66,72,73}

Alternatively, chemical RNA synthesis on polymer solid support is capable to produce high quality oligonucleotides at amounts sufficient for solution NMR spectroscopy.⁷⁴ The major advantage of the approach lies in the high flexibility on positioning the isotope label and on choosing the number of labels to be incorporated into the target RNA. This, however, comes at the costs of the labor-intensive chemical synthesis of appropriate isotope ^{13}C -modified phosphoramidites prior to sequence assembly.

Recently, one of us published a reliable and robust protocol for the chemical synthesis of RNAs comprising 6- ^{13}C -modified pyrimidine building blocks (Figure 3A).⁷⁵ The 6- ^{13}C -uridine and -cytidine phosphoramidites were incorporated into the 27 nt HIV-1 TAR RNA. Using this labeling pattern dynamics occurring on a microsecond to millisecond time scale were probed using Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion (Figure 3B). One residue with a nonflat relaxation dispersion profile was identified, namely uridine 23 (U23, Figure 3B). For this residue, data were acquired at two magnetic field strengths and fitting the data to a two-site exchange process in the intermediate NMR chemical shift time scale according to the Richards Carver equation resulted in an ES population of 1.0% and an exchange rate k_{ex} of 2200 s^{-1} . The two state equilibrium fits to the earlier proposed conformational selection mechanism as U23 undergoes a rigorous structural rearrangement on the addition of the ligand argininamide. In the ligand bound state, a base triple between U38-A27-U23 is formed resulting in a significant chemical shift change for the 6- ^{13}C atom of U23 ($\delta_{\text{free-ARG}} = 3.3\text{ ppm}$), which is in reasonable good agreement with the relaxation dispersion data for U23 ($\delta_{\text{RD}} = 2.4\text{ ppm}$). Based on this chemical shift signature, a transient formation of the base triple between U38, A27, and U23 in the ES was postulated (Figure 3C).

The advances in the RNA ^{13}C -labeling protocols accompanied with the development of improved NMR methods and hardware will make the use of CPMG- and $R_{1\rho}$ -relaxation dispersion methods amenable to study microsecond to millisecond dynamics of RNA at atomic resolution. This will give interesting insights into the functional dynamics of recently discovered RNAs, like riboswitch aptamer domains and ribozymes, and will deepen the understanding of the mechanisms underlying their function.

Top-Down Mass Spectrometry (MS) of RNA

The vast majority of MS studies identifies and/or characterizes the proteins rather than the RNA of biological samples.^{76,77} However, detailed characterization of RNA primary structure by MS is highly desirable whenever posttranscriptional modifications may play a role. Moreover, accurate sequence confirmation for synthetic or extracted RNA is required for a solid interpretation of data from experiments for probing of structural or dynamic features of RNA, such as X-ray crystallography, FRET, and NMR spectroscopy. This can be achieved by high-resolution top-down MS,^{78,79} which is an emerging and powerful method for RNA sequence and modification analysis.⁸⁰ Top-down MS of biomolecules generally involves separate mass measurements of the intact sample and its fragment ions from biomolecular dissociation, as illustrated for RNA in Figure 4. The mass spectrum of the intact RNA (Figure 4A) immediately reveals information about sample heterogeneity, including accurate (to within 1 ppm) mass values and relative abundances of all species observed. Exact mass measurements are highly valuable for comparison with predicted mass values, and can be used for determination of the extent of isotope label incorporation. In further experimental stages, each species of interest is isolated (Figure 4B) and dissociated (Figure 4C) in the mass spectrometer; the resulting fragments from RNA backbone cleavage provide sequence information and localize native or synthetic modifications.

The general concept of top-down MS dates back to the early days of MS,⁸¹ but techniques suitable for desorption/ionization of biomolecules and dissociation of their $(M + nH)^{n+}$ or $(M - nH)^{n-}$ ions have not been available until recently. A major breakthrough for top-down MS of proteins was the invention of electron capture dissociation⁸² and its descendant electron transfer dissociation,⁸³ both of which are, however, impractical for ribonucleic acids (RNAs) as these ionize more readily as negative ions whose cross sections for electron capture or transfer is marginal. Top-down MS of RNA has instead gained momentum with the even more recent extension of collisionally activated dissociation (CAD)^{80,84} and electron detachment dissociation (EDD)⁸⁰ to larger (up to 76 nt, ~25 kDa) and highly modified transfer RNA.

The unique and rather intricate dissociation chemistry involved in EDD of RNA produces non-complementary, even-electron *d* and *w* fragment ions (Figure 5) from radical precursor ions, $(M - nH)^{(n-1)-\bullet}$, by elimination of an uncharged, radical nucleoside unit.⁸⁵ CAD of RNA, on the other hand, produces complementary *c* and *y* fragment ions, provided that $(M - nH)^{n-}$ ion charge is sufficiently small (~0.2 charges/nt).⁸⁵⁻⁸⁷ What at first glance appears to be unimpressive fragment ion nomenclature associated with unimolecular dissociation chemistry, which is of interest to only a few devoted mass spectrometrists is actually key to RNA characterization by top-down MS. As illustrated in Figure 5, the mass values of *c* and

d ions (which include the original 5' terminus) differ by 18.01 Da (H₂O), whereas those of *w* and *y* ions (which include the original 3' terminus) differ by 79.97 Da (HPO₃).⁷⁸ The mass values in an EDD spectrum of RNA that are 18.01 Da higher than mass values in a CAD spectrum of the same RNA can confidently be assigned to *d* and *c* ions, respectively. Likewise, mass values in the EDD spectrum that are 79.97 Da higher than mass values in the CAD spectrum can be assigned to *w* and *y* ions, respectively. Thus the analysis of only two spectra, one from EDD and one from CAD, generates a total of four RNA fragment “mass ladders” with known sequence direction, even for completely unknown RNA with as yet unknown modifications.⁸⁰

While residues with the same elemental composition and therefore mass, such as uridine and pseudouridine, cannot be distinguished from each other with this new approach, additional experiments based on MS or other techniques can be employed for further elucidation.⁸⁸⁻⁹¹ For example, the Breuker group has recently proposed a label-free method for the identification, localization, and relative quantification (accurate to within 10%) of pseudouridine which exploits the strikingly different dissociation behavior of the constitutional isomers uridine (N-glycosidic bond) and pseudouridine (C-glycosidic bond) in CAD of highly charged (~0.5 charges/nt) (M – nH)^{n–} ions of RNA.⁸⁸ Other MS methods for the detection of pseudouridine generally involve chemical labeling reactions, as recently reviewed by Durairaj and Limbach⁸⁹ and Helm and coworkers.⁹¹ However, none of the current MS-based methods for the detection of pseudouridylation sites is based on top-down MS.

The top-down MS approach for sequencing of RNA described above currently necessitates special instrumentation, namely a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer,⁹² which not only provides the required mass resolving power and accuracy but is also operated under sufficiently low pressure for EDD experiments. Although not yet demonstrated for larger RNA, negative ion electron transfer dissociation^{93,94} could potentially be used for RNA dissociation instead of EDD with other types of MS instruments.

For sequencing of unmodified RNA, MS cannot yet compete with any of the “next generation sequencing methods”⁹⁵ established during the past few years, but it has already demonstrated outstanding potential for detailed characterization of modified RNA^{80,84,96,97} by offering full sequence information, including the localization and identification of mass-altering modifications in top-down experiments. Further advancement of top-down MS could even make possible the (relative) quantification of modifications in partially modified RNA, and the analysis of possible “combinatorial codes”⁹⁸ in RNA modifications.

Synthetic Aspects—The development of top-down MS for RNA characterization generally requires high-quality samples in sufficient amount. These can be provided by solid-phase synthesis, which offers high flexibility with respect to sequence variability and positioning of modifications that, in turn, allows for method development and validation using sufficiently large sample sets. Under this aspect, MS is not primarily used for validation of synthetic RNA, rather the synthetic RNA is used to develop and validate new

methodology that can then be used for analysis of RNA from various sources, including natural ones.

CONCLUSIONS

With this personal reflection, we intend to point out how our understanding of RNA properties becomes deepened when a broad range of chemical and biophysical methods is applied. Specifically, solid-phase synthesis and site-specific labeling of the RNA of interest is frequently required to use the full potential of modern spectroscopic methods, such as single-molecule FRET imaging. Synthetic, labeled RNA is also needed to develop and implement new, advanced NMR spectroscopic approaches to explore the dynamics of RNA over a broad timescale. More recently, FT-ICR-MS is emerging as a high-precision tool for RNA sequencing and modification analysis. The novel methodological developments discussed above will play an important role to explore and comprehend the diverse functions of the continuously growing number of newly discovered, biologically relevant RNAs.

Acknowledgments

Contract grant sponsor: Austrian Science Fund (FWF)

Contract grant numbers: I317; P21641; I1040; I844; Y372

REFERENCES

1. Aalto AP, Pasquinelli AE. *Curr Opin Cell Biol.* 2012; 24:333–340. [PubMed: 22464106]
2. Hüttenhofer A, Schattner P, Polacek N. *Trends Genet.* 2005; 21:289–297. [PubMed: 15851066]
3. Serganov A, Patel DJ. *Curr Opin Struct Biol.* 2012; 22:279–286. [PubMed: 22579413]
4. Breaker RR. *Mol Cell.* 2011; 43:867–879. [PubMed: 21925376]
5. Wachowius F, Höbartner C. *ChemBioChem.* 2010; 11:469–480. [PubMed: 20135663]
6. Caruthers MH. *Biochem Soc Trans.* 2011; 39:575–580. [PubMed: 21428942]
7. Bardaro MF Jr, Varani G. *Wiley Interdiscip Rev RNA.* 2012; 3:122–132. [PubMed: 22180312]
8. Bothe JR, Nikolova EN, Eichhorn CD, Chugh J, Hansen AL, Al-Hashimi HM. *Nat Methods.* 2011; 8:919–931. [PubMed: 22036746]
9. Wacker A, Buck J, Richter C, Schwalbe H, Wöhnert J. *RNA Biol.* 2012; 9:672–680. [PubMed: 22647526]
10. Buck J, Fürtig B, Noeske J, Wöhnert J, Schwalbe H. *Proc Natl Acad Sci U S A.* 2007; 104:15699–15704. [PubMed: 17895388]
11. Giessing AM, Kirpekar F. *J Proteomics.* 2012; 75:3434–3449. [PubMed: 22348820]
12. Stark MR, Pleiss JA, Deras M, Scaringe SA, Rader SD. *RNA.* 2006; 12:2014–2019. [PubMed: 16983143]
13. Winkler W, Nahvi A, Breaker RR. *Nature.* 2002; 419:952–956. [PubMed: 12410317]
14. Sudarsan N, Barrick JE, Breaker RR. *RNA.* 2003; 9:644–647. [PubMed: 12756322]
15. Cheah MT, Wachter A, Sudarsan N, Breaker RR. *Nature.* 2007; 447:497–500. [PubMed: 17468745]
16. Sudarsan N, Hammond MC, Block KF, Welz R, Barrick JE, Roth A, Breaker RR. *Science.* 2006; 314:300–304. [PubMed: 17038623]
17. Welz R, Breaker RR. *RNA.* 2007; 13:573–582. [PubMed: 17307816]
18. Serganov A, Polonskaia A, Phan AT, Breaker RR, Patel DJ. *Nature.* 2006; 441:1167–1171. [PubMed: 16728979]
19. Thore S, Leibundgut M, Ban N. *Science.* 2006; 312:1208–1211. [PubMed: 16675665]

20. Edwards TE, Ferré-D'Amaré AR. *Structure*. 2006; 14:1459–1468. [PubMed: 16962976]
21. Rentmeister A, Mayer G, Kuhn N, Famulok M. *Nucl Acids Res*. 2007; 35:3713–3722. [PubMed: 17517779]
22. Kulshina N, Edwards TE, Ferré-D'Amaré AR. *RNA*. 2009; 16:186–196. [PubMed: 19948769]
23. Burnouf D, Ennifar E, Guedich S, Puffer B, Hoffmann G, Bec G, Disdier F, Baltzinger M, Dumas P. *J Am Chem Soc*. 2012; 134:559–565. [PubMed: 22126339]
24. Wong TN, Pan T. *Methods Enzymol*. 2009; 468:167–193. [PubMed: 20946770]
25. Steen K-A, Malhotra A, Weeks KM. *J Am Chem Soc*. 2010; 132:9940–9943. [PubMed: 20597503]
26. Steen K-A, Rice GM, Weeks KM. *J Am Chem Soc*. 2012; 134:13160–13163. [PubMed: 22852530]
27. Ali M, Lipfert J, Seifert S, Herschlag D, Doniach S. *J Mol Biol*. 2010; 396:153–165. [PubMed: 19925806]
28. Baird NJ, Kulshina N, Ferré-D'Amaré AR. *RNA Biology*. 2010; 7:328–332. [PubMed: 20458165]
29. Baird NJ, Ferré-D'Amaré AR. *RNA*. 2010; 16:598–609. [PubMed: 20106958]
30. Anthony PC, Perez CF, García-García C, Block SM. *Proc Natl Acad Sci USA*. 2012; 109:1485–1489. [PubMed: 22219369]
31. Lang K, Rieder R, Micura R. *Nucl Acids Res*. 2007; 35:5370–5378. [PubMed: 17693433]
32. Haller A, Soulière MF, Micura R. *Acc Chem Res*. 2011; 44:1339–1348. [PubMed: 21678902]
33. Haller A, Altman RB, Soulière MF, Blanchard SC, Micura R. *Proc Natl Acad Sci USA*. 2013; 110:4188–4193. [PubMed: 23440214]
34. Rieder R, Lang K, Graber D, Micura R. *ChemBioChem*. 2007; 8:896–902. [PubMed: 17440909]
35. Hall KB. *Methods Enzymol*. 2009; 469:269–285. [PubMed: 20946794]
36. Lang K, Micura R. *Nat Protoc*. 2008; 3:1457–1466. [PubMed: 18772873]
37. Santner T, Rieder U, Kreutz C, Micura R. *J Am Chem Soc*. 2012; 134:11928–11931. [PubMed: 22775200]
38. Haller A, Rieder U, Aigner M, Blanchard SC, Micura R. *Nat Chem Biol*. 2011; 7:393–400. [PubMed: 21532598]
39. Heppell B, Lafontaine DA. *Biochemistry*. 2008; 47:1490–1499. [PubMed: 18205390]
40. Duss O, Maris C, Schroetter, von C, Allain FHT. *Nucl Acids Res*. 2010; 38:e188. [PubMed: 20798173]
41. Duss O, Lukavsky PJ, Allain FH. *Adv Exp Med Biol*. 2012; 992:121–144. [PubMed: 23076582]
42. Steiner M, Rueda D, Sigel RKO. *Angew Chem Int Ed*. 2009; 48:9739–9743. [PubMed: 19924747]
43. Clementi N, Chirkova A, Puffer B, Micura R, Polacek N. *Nat Chem Biol*. 2010; 6:344–351. [PubMed: 20348921]
44. Erlacher MD, Chirkova A, Voegelé P, Polacek N. *Nat Protoc*. 2011; 6:580–592. [PubMed: 21527916]
45. Jahn K, Olsen EM, Nielsen MM, Tørring T, MohammadZadegan R, Andersen ES, Gothelf KV, Kjems J. *Bioconj Chem*. 2011; 22:95–100. [PubMed: 21138255]
46. Baum DA, Silverman SK. *Angew Chem Int Ed*. 2007; 46:3502–3504. [PubMed: 17394278]
47. Silverman SK, Baum DA. *Methods Enzymol*. 2009; 469:95–117. [PubMed: 20946786]
48. Büttner L, Seikowski J, Wawrzyniak K, Ochmann A, Höbartner C. *Bioorg Med Chem*. in press.
49. Paredes E, Das SR. *ChemBioChem*. 2011; 12:125–131. [PubMed: 21132831]
50. Gerowska M, Hall L, Richardson J, Shelbourne M, Brown T. *Tetrahedron*. 2012; 68:857–864.
51. Fauster K, Hartl M, Santner T, Aigner M, Kreutz C, Bister K, Ennifar E, Micura R. *ACS Chem Biol*. 2012; 7:581–589. [PubMed: 22273279]
52. Altman RB, Terry DS, Zhou Z, Zheng Q, Geggier P, Kolster RA, Zhao Y, Javitch JA, Warren JD, Blanchard SC. *Nat Methods*. 2011; 9:68–71. [PubMed: 22081126]
53. Zheng Q, Jockusch S, Zhou Z, Altman RB, Warren JD, Turro NJ, Blanchard SC. *J Phys Chem Lett*. 2012; 3:2200–2203. [PubMed: 22984636]
54. Mittermaier AK, Kay LE. *Trends Biochem Sci*. 2009; 34:601–611. [PubMed: 19846313]

55. Palmer AG III, Massi F. *Chem Rev.* 2006; 106:1700–1719. [PubMed: 16683750]
56. Palmer AG III, Grey MJ, Wang C. *Methods Enzymol.* 2005; 394:430–465. [PubMed: 15808232]
57. Bouvignies G, Vallurupalli P, Hansen DF, Correia BE, Lange O, Bah A, Vernon RM, Dahlquist FW, Baker D, Kay LE. *Nature.* 2011; 477:111–114. [PubMed: 21857680]
58. Vallurupalli P, Bouvignies G, Kay LE. *J Phys Chem B.* 2011; 115:14891–14900. [PubMed: 22077866]
59. Ruschak AM, Velyvis A, Kay LE. *J Biomol NMR.* 2010; 48:129–135. [PubMed: 20949307]
60. Baldwin AJ, Kay LE. *Nat Chem Biol.* 2009; 5:808–814. [PubMed: 19841630]
61. Korzhnev DM, Kay LE. *Acc Chem Res.* 2008; 41:442–451. [PubMed: 18275162]
62. Lu K, Miyazaki Y, Summers MF. *J Biomol NMR.* 2010; 46:113–125. [PubMed: 19789981]
63. Johnson JE Jr, Hoogstraten CG. *J Am Chem Soc.* 2008; 130:16757–16769. [PubMed: 19049467]
64. Nikolova EN, Kim E, Wise AA, O'Brien PJ, Andricioaei I, Al-Hashimi HM. *Nature.* 2011; 470:498–502. [PubMed: 21270796]
65. Nikolova EN, Gottardo FL, Al-Hashimi HM. *J Am Chem Soc.* 2012; 134:3667–36670. [PubMed: 22309937]
66. Thakur CS, Luo Y, Chen B, Eldho NV, Dayie TK. *J Biomol NMR.* 2012; 52:103–114. [PubMed: 22124680]
67. Hansen AL, Nikolova EN, Casiano-Negroni A, Al-Hashimi HM. *J Am Chem Soc.* 2009; 131:3818–3819. [PubMed: 19243182]
68. Dethoff EA, Petzold K, Chugh J, Casiano-Negroni A, Al-Hashimi HM. *Nature.* 2012; 491:724–728. [PubMed: 23041928]
69. Nikonowicz EP, Sirr A, Legault P, Jucker FM, Baer LM, Pardi A. *Nucleic Acids Res.* 1992; 20:4507–4513. [PubMed: 1383927]
70. Nikonowicz EP, Pardi A. *J Mol Biol.* 1993; 232:1141–1156. [PubMed: 8396648]
71. Batey RT, Inada M, Kujawinski E, Puglisi JD, Williamson JR. *Nucleic Acids Res.* 1992; 20:4515–4523. [PubMed: 1383928]
72. Thakur CS, Brown ME, Sama JN, Jackson ME, Dayie TK. *Appl Microbiol Biotechnol.* 2010; 88:771–779. [PubMed: 20730533]
73. Schultheisz HL, Szymczyna BR, Scott LG, Williamson JR. *J Am Chem Soc.* 2011; 133:297–304. [PubMed: 21166398]
74. Micura R. *Angew Chem Int Ed Engl.* 2002; 41:2265–2269. [PubMed: 12203559]
75. Wunderlich CH, Spitzer R, Santner T, Fauster K, Tollinger M, Kreutz C. *J Am Chem Soc.* 2012; 134:7558–7569. [PubMed: 22489874]
76. Yu YH, Ji H, Doudna JA, Leary JA. *Protein Sci.* 2005; 14:1438–1446. [PubMed: 15883184]
77. Chen SS, Williamson JR. *J Mol Biol.* 2013; 425:767–779. [PubMed: 23228329]
78. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW. *J Am Chem Soc.* 1999; 121:806–812.
79. Cui W, Rohrs HW, Gross ML. *Analyst.* 2011; 136:3854–3864. [PubMed: 21826297]
80. Taucher M, Breuker K. *Angew Chem Int Ed.* 2012; 51:11289–11292. [PubMed: 23042528]
81. McLafferty FW. *Annu Rev Anal Chem.* 2011; 4:1–22.
82. Zubarev RA, Kelleher NL, McLafferty FW. *J Am Chem Soc.* 1998; 120:3265–3266.
83. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. *Proc Natl Acad Sci USA.* 2004; 101:9528–9533. [PubMed: 15210983]
84. Huang TY, Liu JA, McLuckey SA. *J Am Soc Mass Spectrom.* 2010; 21:890–898. [PubMed: 20080046]
85. Taucher M, Breuker K. *J Am Soc Mass Spectrom.* 2010; 21:918–929. [PubMed: 20363646]
86. Huang TY, Kharlamova A, Liu J, McLuckey SA. *J Am Soc Mass Spectrom.* 2008; 19:1832–1840. [PubMed: 18799321]
87. Taucher M, Breuker K. *J Am Soc Mass Spectrom.* 2010; 21:278–285. [PubMed: 19932627]
88. Taucher M, Ganisl B, Breuker K. *Int J Mass Spectrom.* 2011; 304:91–97. [PubMed: 21960742]
89. Durairaj A, Limbach PA. *Anal Chim Acta.* 2008; 623:117–125. [PubMed: 18620915]

90. Bruckl T, Globisch D, Wagner M, Muller M, Carell T. *Angew Chem Int Ed*. 2009; 48:7932–7934. [PubMed: 19746495]
91. Kellner S, Burhenne J, Helm M. *RNA Biol*. 2010; 7:237–247. [PubMed: 20224293]
92. Marshall AG, Hendrickson CL, Jackson GS. *Mass Spectrom Rev*. 1998; 17:1–35. [PubMed: 9768511]
93. Coon JJ, Shabanowitz J, Hunt DF, Syka JEP. *J Am Soc Mass Spectrom*. 2005; 16:880–882. [PubMed: 15907703]
94. Huang TY, McLuckey SA. *Int. J. Mass Spectrom*. 2011; 304:140–147.
95. Metzker ML. *Nat Rev*. 2010; 11:31–46. [PubMed: 19997069]
96. Kellersberger KA, Yu E, Kruppa GH, Young MM, Fabris D. *Anal Chem*. 2004; 76:2438–2445. [PubMed: 15117181]
97. Nyakas A, Stucki SR, Schürch S. *J Am Soc Mass Spectrom*. 2011; 22:875–887. [PubMed: 21472522]
98. Lanucara F, Evers CE. *Mass Spectrom Rev*. 2013; 1:27–42. [PubMed: 22718314]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

FIGURE 2.

Exited states of the HIV-1 TAR RNA. A: Ground (GS) and excited (ES) state structures of the TAR RNA. Residues with nonflat dispersion profiles are color-coded (blue: chemical shift changes associated to increased stacking and/or antiglycosidic angles and C3'-endo sugar pucker, red: chemical shift changes associated to decreased stacking and/or *syn* glycosidic angles and non-C3'-endo sugar pucker, gray: residues with flat dispersion profiles). B: Ground (GS) and excited (ES) state structures of the A-site mimic RNA. Color-coding as in A. C: Ground (GS) and excited (ES) state structures of the HIV-1 stem loop 1 RNA. Color-coding as in A. Figure adapted from Ref. 68.

FIGURE 3.

Site-specific ¹³C-RNA labeling for CPMG relaxation dispersion NMR experiments. A: 6-¹³C-modified pyrimidine building blocks. B: CPMG relaxation dispersion profiles of uridine residues of HIV TAR-1 RNA. Residue U23 displays a nonflat dispersion profile. C: Proposed ground and ES structure of the HIV TAR-1 RNA.

FIGURE 4.

Schematic illustration of a hypothetical RNA top-down MS experiment. A: The mass spectrum of the intact RNA reveals, in this example, the presence of two RNA species with relative abundances of 62 and 38% that differ in mass by 28.03 Da, indicating the addition of one C₂H₄ or two CH₂ units. B: Each RNA species is isolated (shown here for RNA species 2) and (C) dissociated in the mass spectrometer; the resulting fragment mass values (only three signals from fragment ions F_n , F_{n+1} , and F_{n+2} are shown, where F stands for c , y , d , or w ions, see Figure 5) provide extensive sequence information for the selected species, including the site and type of modification (methylation of uridine at position $n + 2$ from the 5' or 3' end, depending on the identity of F).

FIGURE 5.

Schematic illustration of c and y , and d and w ions from RNA backbone cleavage by CAD and EDD, respectively, highlighting the mass differences of 18.01 and 79.97 Da between c_n and d_n , and w_{m-n-1} and y_{m-n-1} ions, respectively, in RNA consisting of m nucleotides.