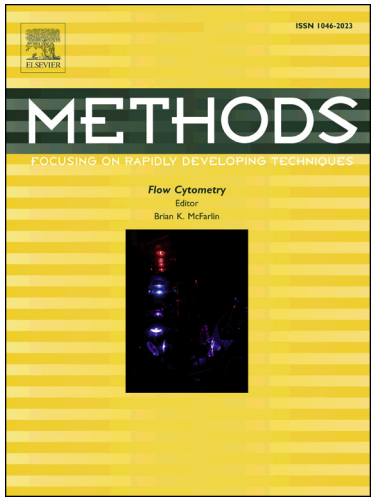


Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition

Andreas Schlundt, Jan-Niklas Tants, Michael Sattler

PII: S1046-2023(17)30120-2
DOI: http://dx.doi.org/10.1016/j.ymeth.2017.03.015
Reference: YMETH 4167

To appear in: *Methods*

Received Date: 5 December 2016
Revised Date: 19 February 2017
Accepted Date: 13 March 2017

Please cite this article as: A. Schlundt, J-N. Tants, M. Sattler, Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition, *Methods* (2017), doi: http://dx.doi.org/10.1016/j.ymeth.2017.03.015

Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition

Andreas Schlundt[1,2], Jan-Niklas Tants[1,2], Michael Sattler[1,2,*]

[1]*Center for Integrated Protein Science Munich (CiPSM) at the Department of Chemistry, Technische Universität München, 85748 Garching, Germany*
[2]*Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany*

* To whom correspondence should be addressed. Tel: +49 (89) 289 13418; Fax: +49 (89) 289 13869; Email: sattler@helmholtz-muenchen.de.

Abstract

Recent advances in RNA sequencing technologies have greatly expanded our knowledge of the RNA landscape in cells, often with spatiotemporal resolution. These techniques identified many new (often non-coding) RNA molecules. Large-scale studies have also discovered novel RNA binding proteins (RBPs), which exhibit single or multiple RNA binding domains (RBDs) for recognition of specific sequence or structured motifs in RNA. Starting from these large-scale approaches it is crucial to unravel the molecular principles of protein-RNA recognition in ribonucleoprotein complexes (RNPs) to understand the underlying mechanisms of gene regulation. Structural biology and biophysical studies at highest possible resolution are key to elucidate molecular mechanisms of RNA recognition by RBPs and how conformational dynamics, weak interactions and cooperative binding contribute to the formation of specific, context-dependent RNPs. While large compact RNPs can be well studied by X-ray crystallography and cryo-EM, analysis of dynamics and weak interaction necessitates the use of solution methods to capture these properties.

Here, we illustrate methods to study the structure and conformational dynamics of protein-RNA complexes in solution starting from the identification of interaction partners in a given RNP. Biophysical and biochemical techniques support the characterization of a protein-RNA complex and identify regions relevant in structural analysis. Nuclear magnetic resonance (NMR) is a powerful tool to gain information on folding, stability and dynamics of RNAs and characterize RNPs in solution. It provides crucial information that is complementary to the static pictures derived from other techniques. NMR can be readily combined with other solution techniques, such as small angle X-ray and/or neutron scattering (SAXS/SANS), electron paramagnetic resonance (EPR), and Förster resonance energy transfer (FRET), which provide information about overall shapes, internal domain arrangements and dynamics. Principles of protein-RNA recognition and current approaches are reviewed and illustrated with recent studies.

Keywords: RNA-protein complex, Integrated structural biology, Nuclear magnetic resonance, Multi-domain proteins, Molecular dynamics, Small angle scattering

1. Introduction

Ribonucleic acids (RNAs) cover a broad range of essential functions in cells, many of them still unknown. While their role in the classical picture of gene expression (i.e. ribosomal RNAs, transfer RNAs), has been uncovered long ago, a growing number of non-coding RNAs has been discovered in the past 25 years, including microRNAs (miRNAs) [1], small interfering RNAs (siRNAs), long non-coding RNAs (lncRNAs) [2], small nuclear and nucleolar RNAs (sn(o)RNAs) [3, 4], circular RNAs and others [5]. Many of these RNAs have been found to play critical roles in modulating gene expression, transcript stability, and RNA processing. Both coding and non-coding RNAs can exhibit secondary and tertiary structures that may potentially confer enzymatic activity. Compared to DNA the RNA world provides a large diversity of conformations and tertiary folds involving stem-loops, pseudoknots, internal bulges, and tertiary interactions that benefit from the presence of multiple functional groups with potential to form hydrogen bonds. The diversity of RNA structural features is also exploited by RNA binding proteins (RBPs) for the formation of RNPs.

RNA binding domains, such as RNA recognition motifs (RRMs), double-stranded RNA binding domains (dsRBDs), K homology (KH) domains or zinc fingers [6-11], are highly abundant and are found in many RBPs [12-15]. These RBPs are involved in the regulation of RNA processing, turnover and modification, including pre-mRNA maturation [16], splicing [17], quality control, degradation [18], transport/localization [19] and translational regulation [20]. Recent genome-wide studies have revealed the presence of previously unknown RBPs that interact with mRNAs [21, 22] through one or more RNA-binding domains (RBDs). Most of the identified RBPs comprise well-known RBDs (RRM, KH, dsRBD, zinc fingers). However, novel RBDs were discovered that were previously not known to bind RNA [21, 22]. For instance, the WD40 domain, a well-known protein–protein interaction module, has been experimentally confirmed to specifically bind RNA [23, 24]. We and others have recently identified the ROQ domain as a novel RBD in the Roquin protein, which plays an important role in post-transcriptional gene regulation in T-cells [25-30]. Structural analysis of these proteins and their RNPs is important to unravel the molecular mechanisms underlying the protein-RNA recognition in these RNPs (**Figure 1A, B**). *Cis* regulatory elements in the mRNA are recognized by *trans* acting proteins (**Figure 1A**) and regulate mRNA processing and stability. Insights into RNA-protein interactions can be the key to understanding why certain RNA elements are functional [31]. On the other hand, unraveling key features required for recognition of target RNAs by an RBP will facilitate the identification of novel RNA targets. Structural analysis therefore is critical to be able to predict protein-RNA interactions [32]. Here, we provide an overview of possible methods to elucidate molecular mechanisms of RNA-protein interactions starting from the initial identification of an RBP and cognate RNA element to a full high-resolution picture of the RNP structure and dynamics.

2. How to discover protein-RNA complexes?

Thanks to recent technical advances protein-RNA interactions can be studied at a genome-wide scale using various methods [33, 34]. Classical approaches to identify protein-RNA complexes involve immunoprecipitation (IP) of the protein to detect the RNA (RNA IP, RIP) via Northern Blots, real-time PCR (which also allows quantification of RNA), microarrays (broad-scale analysis on a chip) or sequencing. RIP has recently been exploited at a quasi-genome-wide

scale to analyze specificity and subcellular localization of hundreds of RBPs [31]. It is possible to pull-down RNAs on large scale, e.g. mRNAs through their poly-A tail or by attaching a specific tag to a particular RNA followed by pulldown (streptavidin-biotin). While the first approach is well suited to isolate endogenous complexes from whole-cell extracts the pulldown of RNA requires isolated RNA to be labeled or tagged. Protein detection is usually achieved through Western blotting. Crosslinking of RNA and interacting proteins prior to co-precipitation is widely used to identify both the RNA and the protein site of interactions within cells. As directed immuno-precipitation is difficult to achieve for a specific RNA (and its bound proteins) methods like CLIP (Cross-Linking with IP) are normally used to identify regions of RNA that are crosslinked to a precipitated RBP (see next paragraph). CLIP will also yield valuable information about the exact regions of a protein that are involved in particular interactions with RNA (**Figure 2**).

Having identified components of an RNP a comprehensive biochemical, biophysical, structural and functional characterization involves the combination of various methods in an integrated manner (**Figure 1C**). These include biochemical and biophysical methods, structural biology techniques up to functional studies *in cells* and *in vivo*.

2.1 Identifying RNA binding regions in RBPs

RNA recognition usually involves specific RNA binding domains (RBDs). The majority of RNA-protein interactions is mediated by highly abundant and well-characterized RRMs, KH-domains, dsRBDs and zinc finger domains [6-11, 35]. The presence of an RBD within a full-length RBP can be identified in various ways (**Figure 2A**). The increasing availability of deposited structural data of RBDs enables the prediction of homologous domains based on the primary sequence. Alternatively, prediction of secondary structure can help to define domain boundaries. Protein expression trials testing various constructs can then identify a soluble globular domain. For the mapping of a novel uncharacterized RNA binding domain [21] limited proteolysis is a valuable tool to identify a protease-resistant globular domain (see reference [36] for further reading). This approach has been successfully used to identify the ROQ domain in the RNA binding protein Roquin, which could not be predicted by other means [25] (see Chapter 3.3 for further details). Notably, an increasing number of RBPs are described where RNA binding involves intrinsically disordered regions (IDRs) in these proteins [37]. The identification and characterization of an RNA binding IDR is challenging as it is non-globular, dynamic and largely unstructured. Here, NMR spectroscopy is a unique tool to map binding interfaces experimentally (see below).

2.2 Identifying protein binding regions of RNAs

Defining a relevant, regulatory region and binding site for an RBP in a large RNA is challenging and requires an assay to monitor functional activity related to the RNA. An efficient approach involves a cellular reporter assay. For example, to identify a regulatory region in the 3' UTR of an mRNA, the 3' UTR or smaller regions are fused to the CDS of an easy-to quantify gene product, e.g. luciferase. This set-up enables testing of different UTR constructs through comparison of mRNA and subsequently protein levels (**Figure 2B**). With this so-called deletion mapping [38, 39] approximate regions of regulatory importance can be defined and then investigated with additional experiments.

Having identified a functionally relevant region in a larger RNA the secondary structure can be mapped using chemical and enzymatic probing methods (see 2.4). SHAPE (Selective 2'-

Hydroxyl acylation Analyzed by Primer Extension) provides a measure of the accessibility of nucleotides based on their chemical reactivity and is a good read-out for RNA secondary structure [40, 41] (**Figure 2B**). This information can be combined with NMR data which also report on RNA base pairing and dynamics [42-45] (see below). The utility of SHAPE and other methods to identify protein-binding sites in RNAs is summarized in Chapter 2.4.

For large-scale analysis of protein-RNA interactions CLIP has proven valuable as it can detect dynamic, low-affine, and transient interactions on a genome-wide scale by crosslinking *in vivo*. It enables the precise identification of RNA binding sites for a given RBP at nucleotide resolution (iCLIP) [46] (**Figure 2C**). CLIP is now routinely combined with high-throughput sequencing (HITS-CLIP [47]) or applied with photo-reactive nucleoside analogs (PAR-CLIP [48]) on a genome-wide scale [46, 49-51]. It is often combined with RNA Bind-n-Seq [52] to characterize individual RBP-bound sequences relative to all input sequences. CLIP can be considered as the most efficient method for the de-novo identification of RNA binding sites. More detailed discussions of CLIP can be found in recent reviews [47, 51, 53]. However, CLIP experiments are usually preformed using full-length RBPs and thus the contributions and specific molecular mechanisms involving the presence of multiple RBDs are not assessed and need to be dissected by biophysical and structural biology techniques (see below).

SELEX (Systematic Evolution of Ligands by EXponential enrichment) has facilitated the identification of DNA- or RNA target sequences of a protein from a systematic screen against a library of randomized sequences [54] (**Figure 2C**). While SELEX is well suited to identify high-affine targets the approach might not provide a complete picture of possible RNA ligands of an RBP, and thus miss interactions that are functionally relevant *in vivo* [55, 56]. RNAcompete is an *in vitro* method to provide an overview with relative affinities of hundreds of short RNA sequences for a given protein [57]. While the two methods reveal well-defined target sequences of an RBP, possible secondary structure in the target RNA is not well detected and requires optimized methods [58]. A combination of RNA fold prediction [59, 60] and experimental validation (for example using NMR methods, see below) is usually required. Prediction of RNA folding is nowadays facilitated computationally and eventually implements temperature, base-pairing restraints and buffer conditions [61, 62].

Biotechnological progress and the computational power to sort and cluster millions of sequences has led to the identification of target sequences/consensus motifs for many RBPs important for, RNA processing and metabolism or the microRNA targetome [63-66]. Deep sequencing is used to identify spatially and temporally confined protein-RNA interactions, e.g. at the ribosome (translatome profiling [67]) or of the nascent transcriptome [68]. Hence, we are now able to partially describe the RNA landscapes in cells with spatiotemporal resolution [69-71]. Combining this with a genome-wide description of RBPs has recently culminated in the quantitative analysis of proteins and RNAs in single cells [72, 73]. Single nucleotide information is also gained by using posttranscriptional modifications of bases (e.g. bisulfite sequencing [74]), which now drives deep-sequencing approaches into the epigenetic field. Data obtained from broad analyses of RNA-protein interactions are used to browse databases predicting binding interfaces, RNA fold, enabling docking and other features based on homology and empiric data [75].

2.3 Molecular principles of protein RNA interactions

Understanding structural features of RBP-RNA recognition beyond their simple binding is important to unravel molecular mechanisms that control many aspects of gene regulation. Over the past two decades our picture of rather static RNA-protein complexes has changed into a scenery of highly dynamic RNPs that constantly remodel their shapes depending on cellular demands. The spliceosome is one striking example of a highly complex protein-RNA machine where large scale changes in composition and structure play crucial roles to perform the catalytic reactions for intron excision [76, 77]. Large catalytic RNP nanomachines like the spliceosome are often targeted by a tightly regulated and dynamic initial assembly of RBPs at the mRNA. This involves the binding of multidomain RBPs to *cis* regulatory RNA motifs in the mRNA (**Figure 1A**), where multiple RBDs can engage several sequence motifs in the RNA. Moreover, multiple RBPs can cooperate to bind the RNA and the dynamic and cooperative binding of multidomain RBPs has been discovered as key feature for the formation of such regulatory RNPs (**Figure 3**), thus necessitating appropriate techniques to study their structural and dynamic features. Most RBPs recognize linear RNA sequence motifs by combinations of common RBDs, i.e. RRM or KH domains. The recognition of structured RNAs is, in general, less well studied, perhaps with the exception of double-stranded RNA helices [35, 78, 79] and RNA hairpin motifs [80-85].

2.3.1 Single-stranded RNA recognition by multidomain RBPs

RNA-protein interactions span several orders of magnitude [86, 87] in affinity and the underlying molecular and structural features are highly versatile. Often an RBP uses a tandem of RBDs to bind to single-stranded RNA (ssRNA) recognizing two different sites to increase affinity and specificity (**Figure 3A**). As such, in a first step the RNA binding of individual domains can be characterized. An additional layer of regulation is the potential presence of closed inactive arrangements of multiple RBDs in the absence of RNA. Conformational rearrangements are then required to enable high-affinity RNA binding providing an additional level of regulation. Tandem RBDs are often connected by flexible linkers [88-91] and can tumble independently in the absence of RNA or sample a range of “inactive” conformations to minimize the formation of non-specific RNPs. This has been observed for the essential splicing factor U2AF65 [90, 92-94] (Chapter 3.2). Alternatively, a prearranged conformation that resembles the RNA-bound structure may facilitate RNA binding but then inhibit splicing of an exon that is looped out, as is the case for the alternative splicing factor PTB [95, 96]. These examples illustrate how different domain arrangements and dynamics are important features for the formation of a regulatory RNP. As a consequence, such dynamic features must be considered when analyzing mechanisms of RNP formation. A simple static picture of the bound state of an RNP will not explain the full biological activity and regulation of RNPs.

2.3.2 Cooperative RNA binding

An important principle of RNA recognition involves the formation of specific RNPs by the cooperative binding of multiple RBDs either within a multidomain RBP or by combining different RBPs [87, 97] (**Figure 3B**). The combinatorial assembly of different RBDs enables the formation of various distinct RNPs and thus can be used to control a wide range of processes with a limited number of RBDs [12, 89, 98]. The simultaneous recognition of a single linear RNA sequence by two RBPs can be further modulated by protein/protein interactions and control the assembly of higher order complexes. Another example for the combination of multiple RBDs involves the combination of sequence-specific and non-sequence specific RNA-binding

domains. Non-specific RNA binding may allow sliding of the RBP on a long RNA sequence to search for high affinity sites that are then recognized sequence-specifically by a second RBD, thereby increasing the efficiency of searching a target *cis* regulatory element. This has been proposed for mRNA binding by Roquin, where shape-specific recognition of a *cis* regulatory RNA stem-loop by the ROQ domain is enhanced by the presence of a zinc finger (ZnF) that is expected to mediate non-specific RNA binding [99]. A guiding function of the zinc finger is suggested by its non-specific RNA-binding properties, which may thus facilitate sliding along single-stranded RNA.

2.3.3 Conformational selection and induced fit

An important mechanistic aspect of RNA recognition involves conformational changes that are coupled with the formation of a protein-RNA complex (**Figure 3C**). As such it will be difficult to predict interaction surfaces even with prior knowledge of the apo protein and RNA structures. Formation of the protein-RNA complex can involve an "induced fit" mechanism or conformational selection from a dynamic ensemble of structures that includes a pre-existing bound conformation (typically only as a minor populated species) [42, 92, 100-104]. While an induced fit presupposes structural adaptation of at least one interaction partner, conformational selection correlates with the sampling of multiple conformers that exist in a dynamic equilibrium. The "correct" conformer resembling the bound conformation is then captured by binding of the ligand. Both mechanisms require solution-based methods to allow for the detection and characterization of dynamics and conformational ensembles.

2.3.4 Sliding of RBPs on double-stranded RNA

Binding to dsRNA often involves sliding of the RBP along the double-stranded RNA helix. For human TRBP sliding has been shown to be an intrinsic ATP-independent feature [105] (**Figure 3D**). The scanning of a dsRNA conformations enables dsRBDs to contribute to miRNA biogenesis. In general, dsRBPs balance dynamics and specificity for dsRNA binding and may vary from completely independent (dynamic) to specific (static) binding [106]. Usually dsRBDs bind non-sequence-specifically to dsRNA helices [107], while some binding specificity is thought to be linked to the presence of non-A-form helical RNA [78], e.g. in stem-loop RNA structures [81, 108]. For protein-DNA complexes intra- and intermolecular jumping of the protein has been described as additional mechanism [109].

We will give examples for the above mentioned principles and illustrate how they can be unraveled using complementary biophysical and structural biology techniques in section 3 of this review

2.4 Biochemical, biophysical methods and structural biology techniques to study protein-RNA interactions

Standard structural biology techniques, i.e. crystallography, NMR and electron microscopy provide high-resolution structural details. Biochemical and biophysical experiments provide information about binding affinity, stoichiometry, kinetic and thermodynamic parameters for protein-RNA interactions (**Table 1**). In the following we briefly summarize the most common approaches.

2.4.1 Biochemical assays

Classical biochemical assays for RNP analysis *in vitro* often rely on the detection of RNA. Electromobility shift (EMSA, separation of free and bound RNA during gel electrophoresis) and filter binding assays (retention of RNA by proteins that are immobilized on a membrane) determine the binding affinity of RNPs and can indicate the stoichiometry of RNP assemblies [110]. Native EMSAs can also indicate kinking of a dsRNA helix in the free and protein bound form [111, 112] as caused by a significantly impaired migration behavior. Note that EMSAs can also be used for the detection of protein (PEMSA) in a protein-nucleic acid complex while the general setup is comparable [113].

Chemical and biochemical footprinting assays can determine protein binding sites in RNA [114, 115] as they will be protected from enzymatic or chemical degradation after binding to a cognate RBP. Similarly, cleavage assays [116] such as oligonucleotide-targeted RNase H protection assays [117] can indicate a potential binding site and protein activity around a specific nucleotide sequence.

Depending on the sequence- or shape-specificity of an RBP for target RNAs different methods may reveal the exact binding site. A particular useful approach involves chemical probing by SHAPE (see also section 2.2). SHAPE can be performed *in vitro* and *in vivo* [118] and is combined with evolutionary constraints for the structure prediction of large RNAs [40, 119, 120], their analysis, validation and modelling [121-123]. Therefore, these SHAPE-derived models cannot be considered fully experimental high resolution structures. Nonetheless, by comparing SHAPE data in the free and protein bound states the effects of protein binding onto the RNA can be analyzed even in the absence of an experimental structure. It is, however, important to distinguish between direct effects that change the accessibility of the binding surface and indirect effects that may also reflect a (large-scale) conformational change or dynamics in the RNA after protein binding. Thus, additional information is required, preferably from high-resolution structural analysis of the free and complexed states.

Site-specific RNP crosslinking e.g. via thiouridine provides unique information about protein-RNA interfaces and entire complex architectures [124, 125]. Systematic integration of restraint information from protein-RNA crosslinks is now being used in an automated way [126] and can define and validate RNP models in the future.

2.4.2 Biophysical assays

Biophysical experiments provide quantitative kinetic and thermodynamic data, binding stoichiometry and affinity for the interaction. This information is also helpful in order to facilitate structure determination (see more details on key parameters of the methods in **Table 1**). Isothermal titration calorimetry (ITC), fluorescence anisotropy, thermophoresis and surface plasmon resonance (SPR) can be used to compare binding affinities of wildtype and mutant binding partners to validate a given protein-RNA interface. Stoichiometries obtained from static light scattering (SLS) or analytical ultracentrifugation (AUC) characterize a RNP in solution. Differential scanning calorimetry (DSC), Thermofluor and Circular dichroism (CD) spectroscopy are usually applied to detect changes in folding states and stability and may also be used to detect e.g. bending of nucleic acid helices as was shown for DNA [127, 128].

Dynamic light scattering (DLS) yields information about sample homogeneity and complex size distributions and thus complements SLS data. Low-resolution structural information of protein-RNA interfaces can be obtained from mass spectrometry after crosslinking [124, 125]. Förster

resonance energy transfer (FRET) and electron paramagnetic resonance (EPR) spectroscopy require the incorporation of pairwise fluorescent or spin labels, respectively, at specific sites. They then yield valuable long-range distance restraints and can give information about dynamic transitions and structural changes of the RNA or the protein [93, 96]. Also, FRET is an ideal method to detect transition states during RNA folding or probe dynamics of complex systems as it was shown for protein sliding along dsRNA [105].

2.4.3 Structural biology

Structural biology provides high-resolution structural information for biological macromolecules and their complexes. Different techniques have unique benefits but also come with specific drawbacks. Thus, depending on the particular question asked and properties of the system to be studied an integrated structural approach, where complementary techniques are combined, is usually the most promising way to understand structural mechanisms. A complete understanding of a particular protein-RNA interaction will only be obtained from a high-resolution structure. X-ray crystallography and cryo-EM can provide highly resolved structures, but they lack information about the dynamics and flexibility of a complex. In fact, many RNP complexes do not occur as static structures but their functional activity requires conformational flexibility that enables rapid conformational changes in response to a cellular signal. Solution NMR spectroscopy is powerful in analyzing dynamic and transient protein and RNA folds and RNP complexes [129-133] at high resolution in a native-like environment. The three major structural biology techniques (crystallography, NMR, cryo-EM) are complemented by other methods to study the structure and dynamics of RNPs.

X-ray crystallography and cryo-EM

X-ray crystallography is the major technique to obtain high-resolution structural information. The majority of structures deposited in the Protein Data Bank (http://www.wwpdb.org/) have been determined by X-ray crystallography, which can provide high resolution structures of small proteins up to large protein complexes (ribosomes). However, the requirement of obtaining well-diffracting single crystals is a challenge and difficult for large complexes. Also crystallization often involves trimming of the protein or complex to remove flexible or disordered regions, which, however, can be essential for biological function and regulation for example by posttranscriptional modifications. Also, the crystalline state traps low energy conformations, sometimes inducing artificial states and cannot provide information about conformational dynamics.

Recently, technical and methodological advances have put cryo-electron microscopy at the forefront of structural studies of high molecular weight complexes. With this, cryo-EM can provide structural information that approaches the resolution of crystal structures [134]. Furthermore, there is basically no size limitation, and single-particle cryo-EM is an ideal tool to study large complexes or machineries, including huge regulatory RNP complexes [135-137]. Regions of cryo-EM structures that are not well resolved can be complemented with high resolution data from crystallography or NMR. While only small amounts of sample are required a major challenge is the extensive sample optimization and preparation. Moreover, extracting dynamical information is not straightforward and flexible regulatory regions will not be visible.

SAXS/SANS

Small angle X-ray/neutron scattering (SAXS/SANS) can provide useful information on the overall shape of a particle in solution and do not require sample crystallization. SAXS is easy to

implement and yields overall shapes at a resolution above 10 Å. The concentration-dependent molecular weight and assembly in solution can be studied. SANS additionally can exploit the different scattering contrast of protonated or deuterated protein and RNA by matching D_2O contents in the solvent. This allows distinguishing domains and subunits in a complex: By selective deuteration of individual protein subunits their relative position within the overall shape can be studied [138]. SAXS and SANS thus well complement each other and the two methods can yield helpful restraints by restricting the conformational space in systems to be modelled. In general, both methods are ideally combined with high-resolution information from X-ray crystallography and NMR [139-146].

EPR

Electron paramagnetic resonance (EPR) detects interactions between unpaired electrons and can provide long-range distance information. EPR requires the incorporation of paramagnetic metals or spin labels to attach the unpaired electron spins to the biomolecule studied. Distance restraints that can be combined with other structural data to be included in structure calculations. EPR is especially useful to complement short distance information obtained from NMR (e.g. PRE, NOE data) since it detects distances up to 80Å [147, 148]. These long-range information make EPR a viable tool to determine positions and orientations of single molecules, e.g. larger RNAs, within a complex (comparable to SANS).

NMR spectroscopy

A major strength of NMR is the capability to detect and analyze dynamic systems. Biomolecular NMR application require a combination of different experiments applied to isotope labelled samples. Recent developments in selective isotope labelling and NMR pulse schemes have pushed the molecular weight boundaries that usually limit NMR [149-152]. Beyond structure determination of proteins, RNAs and their complexes, solution NMR is a powerful method for the investigation of molecular interactions, dynamics and folding of components. Different NMR experiments are available to obtain information of RNA structure and folding [153, 154]. NMR provides versatile and efficient tools to assess and confirm the secondary structure of structured RNAs based on the detection of imino protons in base pairs (**Figure 4A**). Imino-proton signals indicate double-stranded regions and are thus the primary indicator of folded, base-paired RNA [155, 156]. As the NMR resonance frequencies (chemical shifts) of these protons are normally well-separated from other protons their identification and assignment is straightforward. The 2D imino-NOESY spectrum takes advantage of the fact that neighboring base pairs have close spatial proximity of imino protons and will thus enable a sequential walk through the consecutive stretches of base-pairs (“connectivities”). An observable imino signal in NMR spectra of RNA indicates the presence of secondary structure. Interestingly, imino proton experiments can thus complement the information obtained from SHAPE (see Chapter 2.2). The RNA imino resonances are also useful probes to define the binding site with a cognate RBP (**Figure 4B**) as this will lead to local chemical shift perturbations and/or line broadening of NMR signals. However dynamic binding – e.g. by sliding – can cause RNA signal broadening beyond detection and thus complicates the analysis [7].

It is important to note that imino proton-detected NMR experiments do not require labelling with NMR active isotopes (^{13}C, ^{15}N). On the other hand, to map the RNA-binding site within a protein usually ^{1}H, ^{15}N correlation experiments are used, which require ^{15}N isotope-labelling of the protein. These fingerprint spectra of the protein can be easily recorded and enable tracking of the RNA binding event at residue resolution (**Figure 4C**). In ^{1}H,^{15}N correlation experiments of a

protein each amide group provides a resonance signal that can be assigned through three-dimensional heteronuclear experiments [157]. Once the backbone chemical shift assignment is accomplished various types of information can be assessed for a protein. During a step-wise NMR titration of RNA into the protein sample residues that are affected by the RNA (e.g. through a new chemical environment) will undergo a change in chemical shift which can be quantified and used to determine the binding site and affinity [7, 158]. Similar to RNA imino protons, protein amide resonances are useful to perform a line width analysis that will reveal dynamic binding modes, as was shown for proteins sliding on dsDNA [159].

The analysis of NMR data described above provides structural insight into RNA-protein interactions at residue resolution – even before high-resolution structure determination by NMR or X-ray crystallography. The information obtained is already useful to guide the design of functional studies. For example, the mapping of chemical shift perturbations is a reasonable starting point for defining spots of mutagenesis to validate data *in vitro* or *in vivo* (see Chapter 3.3). On the RNA level, simple one-dimensional proton NMR spectra already provide information about base-pairing, even without assignment of the NMR resonances. NMR provides answers to much more complex questions in the characterization of RNPs (dynamics, solution structures) and is now further enhanced by optimized isotope and segmental labelling approaches to overcome spectral overlap for both proteins and the RNA [45, 92, 160-162]. For RNA, the combination of *in vitro* transcription and segmental or site-specific isotope labelling offers adjustable sample conditions for the NMR study of even large RNAs [45, 85, 163, 164] (and reviewed in [165]).

3. Examples: Structure and dynamics of RNPs using integrated methods

A combination of integrated structural biology with biophysical techniques and computational methods is required to study the structure and dynamics of RNPs, as demonstrated by an increasing number of recent studies of protein-RNA complexes [25, 26, 45, 90, 97, 144, 147, 148, 166-175]. In the following we illustrate the approaches and their utility with recent examples highlighting principles of protein-RNA interaction outlined in section 2.

3.1 Cooperative recognition of linear RNA motifs by SXL-Unr

Determining and quantifying cooperativity in RNP complexes is challenging and requires a combination of biophysical methods and integrated structural biology approaches.

The regulation of the male-specific *msl-2* gene is crucial for dosage compensation in Drosophila. Recently, the structural basis for the cooperative binding of two RBPs, namely Sex-lethal (SXL) and Upstream-of-N-ras (UNR), which are recruited to the 3′ untranslated region of the *msl-2* mRNA, has been unraveled [97]. Assembly of this complex leads to inhibition of *msl-2* translation (**Figure 5A**) by preventing the engagement of the small ribosomal subunit [176]. The ternary Sxl–UNR–*msl2* RNA complex features cooperative interactions of three RNA binding domains, i.e. two RRM domains from SXL, and the first cold-shock domain (CSD) of Unr, with a single-stranded *cis* regulatory RNA motif in the *msl-2* RNA.

Hennig et al. have used the tandem RRMs of SXL and the Unr CSD to analyze the ternary complex formation [97]. First, SLS revealed distinct, unambiguous stoichiometries of complexes with different RNA sequences. For structural analysis the predominant 1:1:1 complex was used

(**Figure 5B**). A first picture was created by SANS measurements indicating that the two RBPs sandwich the RNA (**Figure 5C**). SANS makes use of different scattering propensities of protons, deuterons and heavy atoms, i.e. phosphor atoms in RNA: By matching the scattering contrast of different subunits of a protein-RNA complex with the solvent the relative position of protein subunits and RNA components can be determined [138]. When combined with segmental deuterium labelling of multi-domain proteins SANS is a powerful tool to analyze intramolecular arrangements of individual domains [97, 138, 146, 166, 175], while SAXS is only able to provide a general shape. For the Sxl-Unr RNA complex, a low-resolution SAXS/SANS-derived model was used to validate that the conformation observed in the crystal structure reflects the assembly in solution (**Figure 5D, E**). However, SAXS/SANS provides only low resolution information. Therefore, solution NMR was used to confirm the binding interfaces and cooperative interactions in the complex. Chemical shift perturbations revealed that the two proteins bind *msl-2* RNA in a cooperative manner, as the chemical shift trajectories initially induced by RNA binding to the Unr CSD domain follow a different direction upon the addition of the Sxl protein (**Figure 5F**). This is observed for residues in the so-called triple zipper region, where all three components (Sxl, Unr, and RNA) tightly interact, and thus confirms that these interactions are also present in solution. The $^{1}H,^{15}N$ correlation experiments are used to follow ligand titration experiments by NMR and in general provide valuable information on binding surfaces, affinity, stoichiometry and competition/cooperativity [129, 130] (**Figure 4**). The cooperativity of the Sxl-Unr-*msl-2* complex is also documented by ITC data, which revealed that the RNA binding affinity is greatly increased in the ternary complex (15 nM) (**Figure 5G**) compared to the binary RNA interaction between RNA and Sxl alone (200 nM) [97]. ITC provides detailed information about interaction stoichiometries. The results were fully consistent with the stoichiometries obtained from SLS.

Note that, while Sxl is known to bind to uridine-rich RNA elements in the regulation of splicing in the nucleus, the additional binding of the Unr CSD to an extended *cis* regulatory RNA motifs, enables the same protein to participate in the regulation of mRNA translation in the cytoplasm. As exemplified here, cooperativity of multiple domains is expected to play important roles for the RNA recognition and functional activity of multi-domain RBPs that use tandem or higher-order arrangements of RBDs either to increase affinity or specificity [92, 173, 177-180].

3.2 Dynamic shifts of domain arrangements regulate RNA binding by U2AF

Dynamic assembles of multiple RBDs can play critical roles for the regulation of RNA interactions as the will organize hierarchical binding, functional switches and thus provide an additional level of RNP regulation.

The U2AF heterodimer comprising the large 65 kDa subunit of U2 auxiliary factor, (U2AF65) and the small 35 kDa subunit (U2AF35) plays an essential role for the recognition of linear, single-stranded *cis* regulatory RNA elements at the 3' splice site. U2AF65 binds to a poly-pyrimidine-tract (Py-tract) and U2AF35 is thought to recognize the YAG motif at the 3' splice site. Structural studies of the U2AF65 tandem RRMs have been performed using crystallography and NMR in presence and absence of RNA [90, 91, 181, 182]. NMR analysis of the RRM1,2 tandem domains revealed that the domain arrangements of the two RRM domains are drastically different in the absence and presence of RNA [90] (**Figure 6A**). Therefore, domain arrangements in U2AF were probed using NMR paramagnetic relaxation enhancement (PRE) data which provide long-range distance restraints [13, 90]. PREs lead to line-broadening of NMR signals depending on proximity (up to 20Å) to the paramagnetic nitroxyl spin label used

[13, 158] and thereby efficiently report on spatial arrangements of domains or subunits in multidomain proteins and complexes. For U2AF65, PRE data revealed a conformational switch of the protein from a closed form (free) to an open conformation (RNA bound) (**Figure 6A**). The equilibrium between the active and inactive conformations is shifted depending on the overall binding affinity (Py-tract "strength") of a given RNA ligand, which was determined by ITC [90]. The population shifts were also detected based on NMR chemical sift perturbation data (**Figure 6B**). These data show that for a weak Py-tract RRM1 has little contribution to RNA binding, while for strong Py-tracts both domains contribute. In fact, the fraction of open conformations in this dynamic equilibrium is a direct measure of the Py-tract strength that quantitatively correlates with the splicing efficiency of the intron harboring the Py-tract [92]. Note, that the dynamic equilibrium of domain arrangements and its modulation by RNA binding was not possible to detect by crystallographic studies.

Interestingly, SAXS analysis of the free and bound forms revealed that the apo U2AF65 RRM1,2 tandem domains do not only adopt a single closed conformation (as was indicated by the PRE data) but sample a range of domain arrangements, including some where the two RRM domains are not interacting [91]. The conformational dynamics of U2AF65 domains has been recently confirmed using FRET experiments [93] (**Figure 6C**). FRET data revealed that the dynamic switching is faster than milliseconds, fully consistent with NMR data. The example of RNA recognition by the U2AF heterodimer exemplifies the role of conformational dynamics in the regulation of RNA binding and functional activity (pre-mRNA splicing). Analysis of these complex mechanisms requires an integrated approach that combines various solution techniques, i.e. NMR, small angle scattering and FRET with crystallographic data and only thereby enables the full qualitative and quantitative description of the internal dynamics.

3.3 Identification and recognition of RNA stem-loop structures

Compared to linear RNA motifs the identification of structured RNA motifs provides an additional layer of complexity. *In vivo*, such stem-loops can be formed from regions that are distal in sequence. Eventually their recognition is triggered by conserved shape and not sequence which complicates a genome-wide search for motifs. In the daily lab routine, work on structured RNAs requires the permanent control of the native RNA fold during experiments.

The multi-domain protein Roquin (**Figure 7**) represses mRNAs of proteins that mediate differentiation of follicular T helper cells [25-27, 183-186] and thereby antagonizes autoimmunity. How the recognition of *cis* regulatory RNA elements in the 3' UTRs of target genes contributes to this function has long been unknown, although a consensus Roquin target RNA stem-loop had been presented in 2013 [187].Only recently, a number of structural studies from our and other labs have revealed the molecular details and specificity of RNA binding by Roquin [25, 27-30]. To do so, we decided to map the ROQ domain by limited proteolysis (**Figure 7A**) [25] as sequence analysis did not reveal a clear RNA binding domain in Roquin. Roquin binds the constitutive decay element, CDE, a previously described RNA stem-loop element, the secondary structure of which was confirmed with NMR spectroscopy (**Figure 7B**) by us and others [25, 188]. We successfully crystallized this novel RBD type alone and in complex with CDE RNA. Although a dimeric complex is observed in the crystal, a 1:1 stoichiometry of different ROQ domain length versions with RNA was determined in solution based on SLS and SAXS data indicating the need for solution-based methods in the validation of native states. For the ROQ-CDE interaction, NMR data showed that complex formation leads to a stabilization of both the ROQ domain (NMR relaxation data) and of the CDE RNA base

pairing [25]. Nevertheless, an NMR study of the free CDE RNA by the Carlomagno group indicated that the overall structure of the characteristic CDE RNA tri-loop is partially pre-formed in the free RNA [188]. These results underline the power of solution NMR to identify native state intra- and intermolecular dynamics, which is not detectable with other methods.

While the CDE motif had been described earlier as a RNA ligand of Roquin [187], SELEX experiments had also identified a novel sequence motif that resembles a U-rich hexa-loop, named “alternative decay element” (ADE, **Figure 7C, D**) [26]. This was consistent with structural analysis of the ROQ-CDE RNA interaction which indicated that also extended loops can be tolerated by the ROQ domain (**Figure 7E**). This suggested that RNAs with a relaxed consensus can be targeted by Roquin. Bioinformatic analysis of enriched motifs in the SELEX data based on deep sequencing data, consideration of sequence conservation and motif-based sequence analysis suggested the existence of an ADE element in the Roquin target *Ox40* mRNA. Quantitative binding affinities of ROQ with various RNA stem-loops were determined using surface plasmon resonance (SPR) experiments. The contribution of individual nucleotides in the RNA stem-loops for Roquin binding was then assessed by replacing individual nucleotides and performing comparative EMSA analysis. These data helped to define specific features of ROQ interactions with CDE tri- or ADE hexa-loops, which was experimentally demonstrated by crystal structures of ROQ with CDE and ADE RNAs (**Figure 7E**). A final validation of these structural and biophysical data was obtained in cellular assays to probe for the additive functional roles of the CDE and ADE in 3’UTRs for gene expression, mRNA decay and the effect of endogenous levels of mutant Roquin for the presence of Ox40 on the cell surface (not shown). Altogether, the combination of biochemical, biophysical and structural methods was important to obtain a comprehensive picture of the mechanisms of Roquin-RNA interactions. Future studies will need to unravel the role of multiple domains in Roquin for the recognition of RNA [27, 99].

3.4 Integrated structural biology of macromolecular RNP complexes

A special challenge to structural biology is given by macromolecular RNP machineries due to their size and complexity. Thus, the integration of multiple strategies that provide structural data is an unavoidable path to follow.

Recent advances in cryo-EM have allowed us to obtain structural data of macromolecular RNP complexes approaching atomic resolution. One such complex is the Tetrahymena telomerase holoenzyme, a 500 kDa ribonucleoprotein that extends the telomere DNA at the 3’ ends of linear chromosomes. An almost complete atomic picture of a functional telomerase has been obtained by fitting high-resolution crystal and NMR structures into a 9 Å Cryo EM map (**Figure 8A**) [135]. This “classical approach” is a prime example of integrating high-resolution structures into low-resolution maps, where either of them contributes the information that is missing from the other approach. Notably, a relevant combination of crystallography and NMR structural data with cryo-EM allowed interpretation of the cyro-EM map.

A combination of solution techniques was recently presented by the Allain lab [148]. By including long-distance restraints measured with EPR (electron paramagnetic resonance) spectroscopy they were able to determine a structural model of the RsmZ/RsmE complex using multiple samples with double spin-labeled RNAs (**Figure 8B**) [85]. This study shows that with prior optimization of the system the combination of local restraints from NMR and long-distance restraints from EPR is a useful method to determine high-MW structures in the native solution

environment. Moreover, EPR (and NMR) also indicated the presence of multiple conformers, which usually prohibits the structure determination in the crystalline or cryogenic states.

An excellent example for the combination of high-resolution structures and their fit into a global shape is given by the Box C/D methylation complex that resembles a 390 kDa RNP holoenzyme. (**Figure 8C**). Here, individual crystal structures of proteins and the RNA structure are used as input for the calculation of the complete complex arrangement. The conformational space is restricted by SAXS/SANS derived shapes [145]. Lapinaite et al. also include restraints from PREs using methyl probes in ILV-labelled samples on an otherwise deuterated background. The final structural model yields an overall precision of less than 5Å. Albeit, not all details of the macromolecular complex can be resolved, the study also provided a nice application of NMR to characterize the methylation reaction.

Finally, H/D exchange data can be provide information about local stability and dynamics. This has been used in an integrative study by Hurley and colleagues [171] who reported the structure of HIV-1 Tat in complex with HIV-1 TAR RNA and proteins of the human superelongation complex which is hijacked by Tat. Besides RNA SHAPE, SAXS and a low-resolution crystal structure H/D exchange and mass spectrometry data were used to to analyze the stabilization of one protein component in dependence of RNA (**Figure 8D**). Altogether the present a very conclusive and cross-validated structure that fulfills previous knowledge of affinities between the components.

4. Outlook

Ongoing technical and computational developments have facilitated the analysis of an increasing number of RNA-protein complexes with molecular details and spatiotemporal information.

4.1 Open questions and challenges

From a biological point of view we still find completely undiscovered RNP interactomes that have escaped systematic approaches so far. For example, evidence is mounting for a mentionable functional role of intrinsically disordered regions (IDRs) in RBPs, highlighted by recent studies that identify novel RNPs involving IDRs [21, 189]. Their importance is illustrated by the essential roles of unstructured regions [190] or complete intrinsically disordered proteins as exemplified by the Arg/Ser-rich protein nSR100/SRRM4 that acts as an essential splicing regulator in neural cell differentiation [191]. Nevertheless, it is difficult to detect such interactions by large-scale approaches and structural analyses of the corresponding protein-RNA interactions are challenging due to their conformational heterogeneity, intrinsic dynamics and flexibility. Interactions, moreover, often occur in a low-affine and transient manner which complicates their detectability. Low-affine interactions usually require a large excess of one of the binding partners for studies of the bound conformation *in vitro*. Adding to that, low-affine or low-specificity interactions of single RBDs are enhanced by cooperative binding of multiple domains to a more extended target RNA. The impact of functionally important flexible linkers will be followed with growing interest and renders NMR a unique tool for structural analysis.

On the other hand, the conformational flexibility and plasticity of RNA is a central factor for the recognition of RNA elements by RBPs, e.g. for mRNA *cis* regulatory element recognition by protein *trans* factors within 3'-UTRs. With this, a focus will have to be set on the spatiotemporal availability and stability of certain elements beyond their qualitative identification. In other words, we will have to understand the role of minor populated states of RNA elements for a fine-tuned regulation of RNP complex formation preferably in the cellular context. Al-Hashimi and coworkers have revealed a unique role of dynamics in RNA by identifying low-populated states in RNA structural elements [133, 192, 193] that contribute to the regulation of gene expression and signaling. Dynamics within RNA can complicate the analysis of particular protein-RNA interactions in cells, as it is difficult to trap a homogenous complex *in vivo* prior to isolation. NMR is a unique method to characterize and quantify dynamic processes and to identify low-populated states in proteins and RNA [131, 194] beyond static pictures that derive from crystallography or cryo-EM.

Mutually conformational changes of RNA elements and RBPs indicate that induced fit or conformational selection mechanisms can promote the formation of protein-RNA complexes [42, 86, 195]. Presumably, we have only trapped a small portion of biologically relevant RNPs that undergo induced fit or conformational selection mechanisms. In many cases, the binding partners are dynamic and exhibit multiple interconverting conformations, which also sample the bound state. This suggests that conformational selection, linked to the presence of conformational dynamics, plays an important role in protein-RNA molecular recognition. Interestingly, protein-guided RNA dynamics can also drive co-folding during the assembly of RNP complexes. This may aid the recruitment of RBPs during early stages of ribosome formation [196] or the formation and recognition of riboswitches [197, 198]. Detecting dynamical features in RNAs, proteins and their RNPs requires solution techniques such as NMR, small angle scattering, or fluorescence spectroscopy, to complement structural analysis using crystallography, NMR and EM. The sampling of possible conformational states in proteins prior to RNA-binding is nowadays often simulated computationally (eventually guided by experimental restraints) [91, 167], and likely, this approach will be exploited more regularly in the future.

4.2 Methodological advances

To understand the big picture of biological process that involve RNP formation cryo-EM has evolved as a technique that can provide near atomic resolution information. Structures of large, rigid high molecular weight assemblies of RNP machines, such as ribosomes [199], the telomerase holoenzyme [135] or RNA polymerase II [200] are best determined using a combination of single particle cryo-EM methods and X-ray crystallography in the future. However, to understand the regulation and dynamics of these assemblies a combination with solution techniques such as NMR, EPR, SAXS/SANS and FRET will be required that can describe the structure and dynamical features of RNPs.

Mass spectrometry related techniques also become an important method to aid in the determination of RNP assemblies. Ongoing technical developments in the field of native mass spectrometry facilitate the analysis of biomolecular (non-covalent) complexes and their interactions with high sensitivity in a short time. Here, usually electro spray ionization (ESI) is combined with time-of-flight (TOF) mass spectrometers and the structural integrity of sensitive biomolecular samples is maintained [201, 202]. An advantage is the simultaneous detection of multiple states which thus sheds light on dynamics and transitions. However, this technique

currently faces several limitations (suitable buffers, aggregating samples) but promises improvement and further applications in the future. Mass spectrometry analyzing H/D exchange (HDX or HXMS) makes use of different exchange rates of amide backbone protons depending on solvent accessibility or ligands bound [203]. It thus allows to analyze conformational changes, ligand binding and low resolution structural information when combined with proteolysis. Current efforts focus on the application of HXMS to membrane proteins [204]. Most promising, mass spectrometry can be combined with cross-linking [205, 206]. Thereby the conformations of biomolecular complexes are trapped and can be subsequently analyzed by classical MS. This allows analyzing both *in vivo* and *in vitro* trapped complexes and yields further structural information in a way as crosslinks can be implemented as restraints in the definition of a conformational space. Besides that, *in vivo* assays can be used to identify novel targets [207]. Mass spectrometry thus offers future possibilities to further address thriving questions regarding both dynamics and high molecular weight complexes.

In conclusion, a comprehensive picture of molecular principles of RNPs requires the description of the structure and dynamics of protein-RNA recognition by combining static X-ray and cryo-EM structures with information about conformational dynamics that plays essential roles for the assembly and regulation of RNPs.

Acknowledgements

This work was supported by the *Deutsche Forschungsgemeinschaft* (DFG), grants SFB1035, GRK1721 to M.S. and SCHL2062-1/1 to A.S. We apologize for not being able to refer to many additional studies that could not be discussed here.

References

[1] V.N. Kim, J. Han, M.C. Siomi, Nature reviews. Molecular cell biology, 10 (2009) 126-139.
[2] T. Derrien, R. Johnson, G. Bussotti, A. Tanzer, S. Djebali, H. Tilgner, G. Guernec, D. Martin, A. Merkel, D.G. Knowles, J. Lagarde, L. Veeravalli, X. Ruan, Y. Ruan, T. Lassmann, P. Carninci, J.B. Brown, L. Lipovich, J.M. Gonzalez, M. Thomas, C.A. Davis, R. Shiekhattar, T.R. Gingeras, T.J. Hubbard, C. Notredame, J. Harrow, R. Guigo, Genome research, 22 (2012) 1775-1789.
[3] T. Kiss, Journal of cell science, 117 (2004) 5949-5951.
[4] S. Mroczek, A. Dziembowski, Wiley interdisciplinary reviews. RNA, 4 (2013) 581-592.
[5] T.R. Cech, J.A. Steitz, Cell, 157 (2014) 77-94.
[6] A. Clery, M. Blatter, F.H. Allain, Current opinion in structural biology, 18 (2008) 290-298.
[7] G.M. Daubner, A. Clery, F.H. Allain, Current opinion in structural biology, 23 (2013) 100-108.
[8] G. Nicastro, M.F. Garcia-Mayoral, D. Hollingworth, G. Kelly, S.R. Martin, P. Briata, R. Gherzi, A. Ramos, Nature structural & molecular biology, 19 (2012) 1282-1286.
[9] Z. Liu, I. Luyten, M.J. Bottomley, A.C. Messias, S. Houngninou-Molango, R. Sprangers, K. Zanier, A. Kramer, M. Sattler, Science, 294 (2001) 1098-1102.
[10] P.H. Backe, A.C. Messias, R.B. Ravelli, M. Sattler, S. Cusack, Structure, 13 (2005) 1055-1067.

[11] L.R. Saunders, G.N. Barber, FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 17 (2003) 961-983.
[12] J. Hennig, M. Sattler, BioEssays : news and reviews in molecular, cellular and developmental biology, 37 (2015) 899-908.
[13] J. Hennig, L.R. Warner, B. Simon, A. Geerlof, C.D. Mackereth, M. Sattler, in: A.W. Sarah, H.T.A. Frédéric (Eds.) Methods in enzymology, Academic Press2015, pp. 333-362.
[14] T. Ban, J.K. Zhu, K. Melcher, H.E. Xu, Cellular and molecular life sciences : CMLS, 72 (2015) 1045-1058.
[15] A. Re, T. Joshi, E. Kulberkyte, Q. Morris, C.T. Workman, Methods in molecular biology, 1097 (2014) 491-521.
[16] Y. Shi, J.L. Manley, Genes & development, 29 (2015) 889-897.
[17] X.D. Fu, M. Ares, Jr., Nature reviews. Genetics, 15 (2014) 689-701.
[18] D.R. Schoenberg, L.E. Maquat, Nature reviews. Genetics, 13 (2012) 246-259.
[19] A.R. Buxbaum, G. Haimovich, R.H. Singer, Nature reviews. Molecular cell biology, 16 (2015) 95-109.
[20] J. Schott, G. Stoecklin, Wiley interdisciplinary reviews. RNA, 1 (2010) 432-456.
[21] A. Castello, B. Fischer, K. Eichelbaum, R. Horos, B.M. Beckmann, C. Strein, N.E. Davey, D.T. Humphreys, T. Preiss, L.M. Steinmetz, J. Krijgsveld, M.W. Hentze, Cell, 149 (2012) 1393-1406.
[22] A.G. Baltz, M. Munschauer, B. Schwanhausser, A. Vasile, Y. Murakawa, M. Schueler, N. Youngs, D. Penfold-Brown, K. Drew, M. Milek, E. Wyler, R. Bonneau, M. Selbach, C. Dieterich, M. Landthaler, Molecular cell, 46 (2012) 674-690.
[23] S.C. Kwon, H. Yi, K. Eichelbaum, S. Fohr, B. Fischer, K.T. You, A. Castello, J. Krijgsveld, M.W. Hentze, V.N. Kim, Nature structural & molecular biology, 20 (2013) 1122-1130.
[24] I. Loedige, M. Stotz, S. Qamar, K. Kramer, J. Hennig, T. Schubert, P. Loffler, G. Langst, R. Merkl, H. Urlaub, G. Meister, Genes & development, 28 (2014) 749-764.
[25] A. Schlundt, G.A. Heinz, R. Janowski, A. Geerlof, R. Stehle, V. Heissmeyer, D. Niessing, M. Sattler, Nature structural & molecular biology, 21 (2014) 671-678.
[26] R. Janowski, G.A. Heinz, A. Schlundt, N. Wommelsdorf, S. Brenner, A.R. Gruber, M. Blank, T. Buch, R. Buhmann, M. Zavolan, D. Niessing, V. Heissmeyer, M. Sattler, Nature communications, 7 (2016) 11032.
[27] A. Schlundt, D. Niessing, V. Heissmeyer, M. Sattler, Wiley interdisciplinary reviews. RNA, (2016).
[28] D. Tan, M. Zhou, M. Kiledjian, L. Tong, Nature structural & molecular biology, 21 (2014) 679-685.
[29] M. Srivastava, G. Duan, N.J. Kershaw, V. Athanasopoulos, J.H. Yeo, T. Ose, D. Hu, S.H. Brown, S. Jergic, H.R. Patel, A. Pratama, S. Richards, A. Verma, E.Y. Jones, V. Heissmeyer, T. Preiss, N.E. Dixon, M.M. Chong, J.J. Babon, C.G. Vinuesa, Nature communications, 6 (2015) 6253.
[30] A. Schuetz, Y. Murakawa, E. Rosenbaum, M. Landthaler, U. Heinemann, Nature communications, 5 (2014) 5701.
[31] B. Sundararaman, L. Zhan, S.M. Blue, R. Stanton, K. Elkins, S. Olson, X. Wei, E.L. Van Nostrand, G.A. Pratt, S.C. Huelga, B.M. Smalec, X. Wang, E.L. Hong, J.M. Davidson, E. Lecuyer, B.R. Graveley, G.W. Yeo, Molecular cell, 61 (2016) 903-913.
[32] Z.P. Liu, L. Chen, Current topics in medicinal chemistry, 16 (2016) 604-615.
[33] J. Konig, K. Zarnack, N.M. Luscombe, J. Ule, Nature reviews. Genetics, 13 (2012) 77-83.
[34] Z.T. Campbell, M. Wickens, Trends in biochemical sciences, 40 (2015) 157-164.
[35] S. Gerstberger, M. Hafner, T. Tuschl, Nature reviews. Genetics, 15 (2014) 829-845.
[36] L. Columbus, Current opinion in structural biology, 32 (2015) 131-138.
[37] S. Calabretta, S. Richard, Trends in biochemical sciences, 40 (2015) 662-672.
[38] R. Shen, W.A. Miller, Journal of virology, 78 (2004) 4655-4664.

[39] V.S. Sriskanda, G. Pruss, X. Ge, V.B. Vance, Journal of virology, 70 (1996) 5266-5271.
[40] J.T. Low, K.M. Weeks, Methods, 52 (2010) 150-158.
[41] K.A. Wilkinson, E.J. Merino, K.M. Weeks, Nature protocols, 1 (2006) 1610-1616.
[42] F.T. Edelmann, A. Schlundt, R.G. Heym, A. Jenner, A. Niedner-Boblenz, M.I. Syed, J.C. Paillart, R. Stehle, R. Janowski, M. Sattler, R.P. Jansen, D. Niessing, Nat. Struct. Mol. Biol., in press (2016).
[43] R.P. Barnwal, E. Loh, K.S. Godin, J. Yip, H. Lavender, C.M. Tang, G. Varani, Nucleic acids research, 44 (2016) 9426-9437.
[44] Y. Chen, L. Zubovic, F. Yang, K. Godin, T. Pavelitz, J. Castellanos, P. Macchi, G. Varani, Nucleic acids research, 44 (2016) 4381-4395.
[45] S.C. Keane, X. Heng, K. Lu, S. Kharytonchyk, V. Ramakrishnan, G. Carter, S. Barton, A. Hosic, A. Florwick, J. Santos, N.C. Bolden, S. McCowin, D.A. Case, B.A. Johnson, M. Salemi, A. Telesnitsky, M.F. Summers, Science, 348 (2015) 917-921.
[46] Y. Sugimoto, J. Konig, S. Hussain, B. Zupan, T. Curk, M. Frye, J. Ule, Genome biology, 13 (2012) R67.
[47] R.B. Darnell, Wiley interdisciplinary reviews. RNA, 1 (2010) 266-286.
[48] M. Hafner, M. Landthaler, L. Burger, M. Khorshid, J. Hausser, P. Berninger, A. Rothballer, M. Ascano, A.C. Jungkamp, M. Munschauer, A. Ulrich, G.S. Wardle, S. Dewell, M. Zavolan, T. Tuschl, Journal of visualized experiments : JoVE, (2010).
[49] J. Ule, K.B. Jensen, M. Ruggiu, A. Mele, A. Ule, R.B. Darnell, Science, 302 (2003) 1212-1215.
[50] C. Zhang, R.B. Darnell, Nature biotechnology, 29 (2011) 607-614.
[51] T. Wang, G. Xiao, Y. Chu, M.Q. Zhang, D.R. Corey, Y. Xie, Nucleic acids research, 43 (2015) 5263-5274.
[52] N. Lambert, A. Robertson, M. Jangi, S. McGeary, P.A. Sharp, C.B. Burge, Molecular cell, 54 (2014) 887-900.
[53] I. Huppertz, J. Attig, A. D'Ambrogio, L.E. Easton, C.R. Sibley, Y. Sugimoto, M. Tajnik, J. Konig, J. Ule, Methods, 65 (2014) 274-287.
[54] L. Gold, Journal of molecular evolution, 81 (2015) 140-143.
[55] C. Tuerk, L. Gold, Science, 249 (1990) 505-510.
[56] A.D. Ellington, J.W. Szostak, Nature, 346 (1990) 818-822.
[57] D. Ray, H. Kazan, E.T. Chan, L. Pena Castillo, S. Chaudhry, S. Talukder, B.J. Blencowe, Q. Morris, T.R. Hughes, Nature biotechnology, 27 (2009) 667-670.
[58] Y. Sugimoto, A. Vigilante, E. Darbo, A. Zirra, C. Militti, A. D'Ambrogio, N.M. Luscombe, J. Ule, Nature, 519 (2015) 491-494.
[59] R. Backofen, Methods in molecular biology, 1097 (2014) 417-435.
[60] X. Li, H. Kazan, H.D. Lipshitz, Q.D. Morris, Wiley interdisciplinary reviews. RNA, 5 (2014) 111-130.
[61] M. Zuker, Nucleic acids research, 31 (2003) 3406-3415.
[62] I.L. Hofacker, Methods in molecular biology, 1097 (2014) 71-84.
[63] D.D. Licatalosi, A. Mele, J.J. Fak, J. Ule, M. Kayikci, S.W. Chi, T.A. Clark, A.C. Schweitzer, J.E. Blume, X. Wang, J.C. Darnell, R.B. Darnell, Nature, 456 (2008) 464-469.
[64] G.W. Yeo, N.G. Coufal, T.Y. Liang, G.E. Peng, X.D. Fu, F.H. Gage, Nature structural & molecular biology, 16 (2009) 130-137.
[65] J. Konig, K. Zarnack, G. Rot, T. Curk, M. Kayikci, B. Zupan, D.J. Turner, N.M. Luscombe, J. Ule, Nature structural & molecular biology, 17 (2010) 909-915.
[66] P.M. Clark, P. Loher, K. Quann, J. Brody, E.R. Londin, I. Rigoutsos, Scientific reports, 4 (2014) 5947.
[67] G.A. Brar, J.S. Weissman, Nature reviews. Molecular cell biology, 16 (2015) 651-664.
[68] L.S. Churchman, J.S. Weissman, Nature, 469 (2011) 368-373.

[69] J. Ma, Z. Liu, N. Michelotti, S. Pitchiaya, R. Veerapaneni, J.R. Androsavich, N.G. Walter, W. Yang, Nature communications, 4 (2013) 2414.
[70] C.M. Connelly, R. Uprety, J. Hemphill, A. Deiters, Molecular bioSystems, 8 (2012) 2987-2993.
[71] Y. Ding, Y. Tang, C.K. Kwok, Y. Zhang, P.C. Bevilacqua, S.M. Assmann, Nature, 505 (2014) 696-700.
[72] C. Albayrak, C.A. Jordi, C. Zechner, J. Lin, C.A. Bichsel, M. Khammash, S. Tay, Molecular cell, 61 (2016) 914-924.
[73] B. Wu, A.R. Buxbaum, Z.B. Katz, Y.J. Yoon, R.H. Singer, Cell, 162 (2015) 211-220.
[74] C.X. Song, C. Yi, C. He, Nature biotechnology, 30 (2012) 1107-1116.
[75] V.J. Henry, A.E. Bandrowski, A.S. Pepin, B.J. Gonzalez, A. Desfeux, Database : the journal of biological databases and curation, 2014 (2014).
[76] M.M. Golas, B. Sander, S. Bessonov, M. Grote, E. Wolf, B. Kastner, H. Stark, R. Luhrmann, Molecular cell, 40 (2010) 927-938.
[77] P. Wysoczanski, C. Schneider, S. Xiang, F. Munari, S. Trowitzsch, M.C. Wahl, R. Luhrmann, S. Becker, M. Zweckstetter, Nature structural & molecular biology, 21 (2014) 911-918.
[78] R. Stefl, F.C. Oberstrass, J.L. Hood, M. Jourdan, M. Zimmermann, L. Skrisovska, C. Maris, L. Peng, C. Hofr, R.B. Emeson, F.H. Allain, Cell, 143 (2010) 225-237.
[79] J.M. Ryter, S.C. Schultz, The EMBO journal, 17 (1998) 7505-7513.
[80] R. Thapar, A.P. Denmon, E.P. Nikonowicz, Wiley interdisciplinary reviews. RNA, 5 (2014) 49-67.
[81] Z. Wang, E. Hartman, K. Roy, G. Chanfreau, J. Feigon, Structure, 19 (2011) 999-1010.
[82] F.C. Oberstrass, A. Lee, R. Stefl, M. Janis, G. Chanfreau, F.H. Allain, Nature structural & molecular biology, 13 (2006) 160-167.
[83] P.E. Johnson, L.W. Donaldson, Nature structural & molecular biology, 13 (2006) 177-178.
[84] C. Johansson, L.D. Finger, L. Trantirek, T.D. Mueller, S. Kim, I.A. Laird-Offringa, J. Feigon, Journal of molecular biology, 337 (2004) 799-816.
[85] O. Duss, E. Michel, M. Yulikov, M. Schubert, G. Jeschke, F.H. Allain, Nature, 509 (2014) 588-592.
[86] E. Jankowsky, M.E. Harris, Nature reviews. Molecular cell biology, 16 (2015) 533-544.
[87] S. Helder, A.J. Blythe, C.S. Bond, J.P. Mackay, Current opinion in structural biology, 38 (2016) 83-91.
[88] Y. Shamoo, N. Abdul-Manan, K.R. Williams, Nucleic Acids Res., 23 (1995) 725-728.
[89] B.M. Lunde, C. Moore, G. Varani, Nature reviews. Molecular cell biology, 8 (2007) 479-490.
[90] C.D. Mackereth, T. Madl, S. Bonnal, B. Simon, K. Zanier, A. Gasch, V. Rybin, J. Valcarcel, M. Sattler, Nature, 475 (2011) 408-411.
[91] J.R. Huang, L.R. Warner, C. Sanchez, F. Gabel, T. Madl, C.D. Mackereth, M. Sattler, M. Blackledge, Journal of the American Chemical Society, 136 (2014) 7068-7076.
[92] C.D. Mackereth, M. Sattler, Current opinion in structural biology, 22 (2012) 287-296.
[93] L. Voith von Voithenberg, C. Sanchez-Rico, H.S. Kang, T. Madl, K. Zanier, A. Barth, L.R. Warner, M. Sattler, D.C. Lamb, Proceedings of the National Academy of Sciences of the United States of America, 113 (2016) E7169-E7175.
[94] J.L. Jenkins, A.A. Agrawal, A. Gupta, M.R. Green, C.L. Kielkopf, Nucleic acids research, 41 (2013) 3859-3873.
[95] F.C. Oberstrass, S.D. Auweter, M. Erat, Y. Hargous, A. Henning, P. Wenter, L. Reymond, B. Amir-Ahmady, S. Pitsch, D.L. Black, F.H. Allain, Science, 309 (2005) 2054-2057.
[96] R. Lamichhane, G.M. Daubner, J. Thomas-Crusells, S.D. Auweter, C. Manatschal, K.S. Austin, O. Valniuk, F.H. Allain, D. Rueda, Proceedings of the National Academy of Sciences of the United States of America, 107 (2010) 4105-4110.

[97] J. Hennig, C. Militti, G.M. Popowicz, I. Wang, M. Sonntag, A. Geerlof, F. Gabel, F. Gebauer, M. Sattler, Nature, 515 (2014) 287-290.
[98] J. Hennig, F. Gebauer, M. Sattler, Cell cycle, 13 (2014) 3619-3620.
[99] Y. Murakawa, M. Hinz, J. Mothes, A. Schuetz, M. Uhl, E. Wyler, T. Yasuda, G. Mastrobuoni, C.C. Friedel, L. Dolken, S. Kempa, M. Schmidt-Supprian, N. Bluthgen, R. Backofen, U. Heinemann, J. Wolf, C. Scheidereit, M. Landthaler, Nature communications, 6 (2015) 7367.
[100] D.D. Boehr, R. Nussinov, P.E. Wright, Nature chemical biology, 5 (2009) 789-796.
[101] A.D. Vogt, E. Di Cera, Biochemistry, 52 (2013) 5723-5729.
[102] T. Bohnuud, D. Kozakov, S. Vajda, PLoS computational biology, 10 (2014) e1003872.
[103] E. Kligun, Y. Mandel-Gutfreund, RNA biology, 10 (2013) 982-989.
[104] E. Kligun, Y. Mandel-Gutfreund, RNA biology, 12 (2015) 720-727.
[105] H.R. Koh, M.A. Kidwell, K. Ragunathan, J.A. Doudna, S. Myong, Proceedings of the National Academy of Sciences of the United States of America, 110 (2013) 151-156.
[106] X. Wang, L. Vukovic, H.R. Koh, K. Schulten, S. Myong, Nucleic acids research, 43 (2015) 7566-7576.
[107] G. Masliah, P. Barraud, F.H. Allain, Cellular and molecular life sciences : CMLS, 70 (2013) 1875-1895.
[108] H. Wu, A. Henras, G. Chanfreau, J. Feigon, Proceedings of the National Academy of Sciences of the United States of America, 101 (2004) 8307-8312.
[109] J. Iwahara, M. Zweckstetter, G.M. Clore, Proceedings of the National Academy of Sciences of the United States of America, 103 (2006) 15062-15067.
[110] L.M. Hellman, M.G. Fried, Nature protocols, 2 (2007) 1849-1861.
[111] D.M. Lilley, A. Bhattacharyya, S. McAteer, Biotechnology & genetic engineering reviews, 10 (1992) 379-401.
[112] X. Zheng, P.C. Bevilacqua, Proceedings of the National Academy of Sciences of the United States of America, 97 (2000) 14162-14167.
[113] E. Nkansah, R. Shah, G.W. Collie, G.N. Parkinson, J. Palmer, K.M. Rahman, T.T. Bui, A.F. Drake, J. Husby, S. Neidle, G. Zinzalla, D.E. Thurston, A.F. Wilderspin, FEBS letters, 587 (2013) 833-839.
[114] P. Tijerina, S. Mohr, R. Russell, Nature protocols, 2 (2007) 2608-2623.
[115] Y. Peng, T.J. Soper, S.A. Woodson, Methods in molecular biology, 905 (2012) 213-224.
[116] N.K. Sinha, K.D. Trettin, P.J. Aruscavage, B.L. Bass, Molecular cell, 58 (2015) 406-417.
[117] A. Gunzl, Z. Palfi, A. Bindereif, Methods, 26 (2002) 162-169.
[118] B. Lee, R.A. Flynn, A. Kadina, J.K. Guo, E.T. Kool, H.Y. Chang, Rna, (2016).
[119] Z.Z. Xu, D.H. Mathews, Methods in molecular biology, 1490 (2016) 163-176.
[120] K.E. Watters, J.B. Lucks, Methods in molecular biology, 1490 (2016) 135-162.
[121] M.J. Smola, G.M. Rice, S. Busan, N.A. Siegfried, K.M. Weeks, Nature protocols, 10 (2015) 1643-1669.
[122] C.A. Lavender, R. Lorenz, G. Zhang, R. Tamayo, I.L. Hofacker, K.M. Weeks, PLoS computational biology, 11 (2015) e1004126.
[123] P.J. Homan, A. Tandon, G.M. Rice, F. Ding, N.V. Dokholyan, K.M. Weeks, Biochemistry, 53 (2014) 6825-6833.
[124] K. Kramer, T. Sachsenberg, B.M. Beckmann, S. Qamar, K.L. Boon, M.W. Hentze, O. Kohlbacher, H. Urlaub, Nature methods, 11 (2014) 1064-1070.
[125] S. Qamar, K. Kramer, H. Urlaub, Methods in enzymology, 558 (2015) 417-463.
[126] J. Veit, T. Sachsenberg, A. Chernev, F. Aicheler, H. Urlaub, O. Kohlbacher, Journal of proteome research, 15 (2016) 3441-3448.
[127] Y.C. Lou, I. Wang, M. Rajasekaran, Y.F. Kao, M.R. Ho, S.T. Hsu, S.H. Chou, S.H. Wu, C. Chen, Nucleic acids research, 42 (2014) 4080-4093.

[128] I. Papapanagiotou, S.D. Streeter, P.D. Cary, G.G. Kneale, Nucleic acids research, 35 (2007) 2643-2650.
[129] M. Bieri, A.H. Kwan, M. Mobli, G.F. King, J.P. Mackay, P.R. Gooley, The FEBS journal, 278 (2011) 704-715.
[130] A.H. Kwan, M. Mobli, P.R. Gooley, G.F. King, J.P. Mackay, The FEBS journal, 278 (2011) 687-703.
[131] J.R. Bothe, E.N. Nikolova, C.D. Eichhorn, J. Chugh, A.L. Hansen, H.M. Al-Hashimi, Nature methods, 8 (2011) 919-931.
[132] H.M. Al-Hashimi, Journal of magnetic resonance, 237 (2013) 191-204.
[133] Y. Xue, D. Kellogg, I.J. Kimsey, B. Sathyamoorthy, Z.W. Stein, M. McBrairty, H.M. Al-Hashimi, Methods in enzymology, 558 (2015) 39-73.
[134] A. Merk, A. Bartesaghi, S. Banerjee, V. Falconieri, P. Rao, M.I. Davis, R. Pragani, M.B. Boxer, L.A. Earl, J.L. Milne, S. Subramaniam, Cell, 165 (2016) 1698-1707.
[135] J. Jiang, E.J. Miracco, K. Hong, B. Eckert, H. Chan, D.D. Cash, B. Min, Z.H. Zhou, K. Collins, J. Feigon, Nature, 496 (2013) 187-192.
[136] T.H. Nguyen, W.P. Galej, X.C. Bai, C.G. Savva, A.J. Newman, S.H. Scheres, K. Nagai, Nature, 523 (2015) 47-52.
[137] T.H. Nguyen, W.P. Galej, X.C. Bai, C. Oubridge, A.J. Newman, S.H. Scheres, K. Nagai, Nature, 530 (2016) 298-302.
[138] F. Gabel, Methods in enzymology, 558 (2015) 391-415.
[139] A. Grishaev, J. Wu, J. Trewhella, A. Bax, Journal of the American Chemical Society, 127 (2005) 16621-16628.
[140] C.D. Putnam, M. Hammel, G.L. Hura, J.A. Tainer, Quarterly reviews of biophysics, 40 (2007) 191-285.
[141] F. Gabel, B. Simon, M. Nilges, M. Petoukhov, D. Svergun, M. Sattler, Journal of biomolecular NMR, 41 (2008) 199-208.
[142] T. Madl, F. Gabel, M. Sattler, Journal of structural biology, 173 (2011) 472-482.
[143] Y. Takayama, C.D. Schwieters, A. Grishaev, R. Ghirlando, G.M. Clore, Journal of the American Chemical Society, 133 (2011) 424-427.
[144] J. Hennig, I. Wang, M. Sonntag, F. Gabel, M. Sattler, Journal of biomolecular NMR, 56 (2013) 17-30.
[145] A. Lapinaite, B. Simon, L. Skjaerven, M. Rakwalska-Bange, F. Gabel, T. Carlomagno, Nature, 502 (2013) 519-523.
[146] J. Hennig, M. Sattler, Protein science : a publication of the Protein Society, 23 (2014) 669-682.
[147] O. Duss, M. Yulikov, G. Jeschke, F.H. Allain, Nature communications, 5 (2014) 3669.
[148] O. Duss, M. Yulikov, F.H. Allain, G. Jeschke, Methods in enzymology, 558 (2015) 279-331.
[149] R. Sprangers, A. Velyvis, L.E. Kay, Nature methods, 4 (2007) 697-703.
[150] T.L. Religa, A.M. Ruschak, R. Rosenzweig, L.E. Kay, Journal of the American Chemical Society, 133 (2011) 9063-9068.
[151] R. Rosenzweig, S. Moradi, A. Zarrine-Afsar, J.R. Glover, L.E. Kay, Science, 339 (2013) 1080-1083.
[152] R. Rosenzweig, L.E. Kay, Annual review of biochemistry, 83 (2014) 291-315.
[153] C. Helmling, S. Keyhani, F. Sochor, B. Furtig, M. Hengesbach, H. Schwalbe, Journal of biomolecular NMR, 63 (2015) 67-76.
[154] M.P. Latham, D.J. Brown, S.A. McCallum, A. Pardi, Chembiochem : a European journal of chemical biology, 6 (2005) 1492-1505.
[155] A. Dallmann, B. Simon, M.M. Duszczyk, H. Kooshapur, A. Pardi, W. Bermel, M. Sattler, Angewandte Chemie, 52 (2013) 10487-10490.
[156] A. Dallmann, M. Sattler, Current protocols in nucleic acid chemistry, 59 (2014) 7 22 21-19.

[157] M. Sattler, J. Schleucher, C. Griesinger, Progress in nuclear magnetic resonance spectroscopy, 34 (1999) 93-158.
[158] C. Gobl, T. Madl, B. Simon, M. Sattler, Prog Nucl Magn Reson Spectrosc, 80 (2014) 26-63.
[159] K. Loth, M. Gnida, J. Romanuka, R. Kaptein, R. Boelens, Journal of biomolecular NMR, 56 (2013) 41-49.
[160] S. Nozinovic, C. Richter, J. Rinnenthal, B. Furtig, E. Duchardt-Ferner, J.E. Weigand, H. Schwalbe, Journal of the American Chemical Society, 132 (2010) 10318-10329.
[161] S. Saxena, J. Stanek, M. Cevec, J. Plavec, W. Kozminski, Journal of biomolecular NMR, 60 (2014) 91-98.
[162] Y.L. Ahmed, R. Ficner, RNA biology, 11 (2014) 427-432.
[163] I. Kim, P.J. Lukavsky, J.D. Puglisi, Journal of the American Chemical Society, 124 (2002) 9338-9339.
[164] I. Kawahara, K. Haruta, Y. Ashihara, D. Yamanaka, M. Kuriyama, N. Toki, Y. Kondo, K. Teruya, J. Ishikawa, H. Furuta, Y. Ikawa, C. Kojima, Y. Tanaka, Nucleic acids research, 40 (2012) e7.
[165] O. Duss, P.J. Lukavsky, F.H. Allain, Advances in experimental medicine and biology, 992 (2012) 121-144.
[166] L. Freiburger, M. Sonntag, J. Hennig, J. Li, P. Zou, M. Sattler, Journal of biomolecular NMR, 63 (2015) 1-8.
[167] M. Krepl, A. Clery, M. Blatter, F.H. Allain, J. Sponer, Nucleic acids research, 44 (2016) 6452-6470.
[168] T. Carlomagno, Journal of magnetic resonance, 241 (2014) 126-136.
[169] D.K. Yadav, P.J. Lukavsky, Prog Nucl Magn Reson Spectrosc, 97 (2016) 57-81.
[170] J. Weber, H. Bao, C. Hartlmuller, Z. Wang, A. Windhager, R. Janowski, T. Madl, P. Jin, D. Niessing, eLife, 5 (2016).
[171] U. Schulze-Gahmen, I. Echeverria, G. Stjepanovic, Y. Bai, H. Lu, D. Schneidman-Duhovny, J.A. Doudna, Q. Zhou, A. Sali, J.H. Hurley, eLife, 5 (2016).
[172] A. Mourao, S. Bonnal, K. Soni, L. Warner, R. Bordonne, J. Valcarcel, M. Sattler, eLife, 5 (2016).
[173] Z. Cienikova, S. Jayne, F.F. Damberger, F.H. Allain, C. Maris, Rna, 21 (2015) 1931-1942.
[174] L. Vukovic, C. Chipot, D.L. Makino, E. Conti, K. Schulten, Journal of the American Chemical Society, 138 (2016) 4069-4078.
[175] A. Lapinaite, Simon, B., Skjaerven, L., Rakwalska-Bange, M., Gabel, F., Carlomagno, T., Nature, doi:10.1038/nature12581 (2013).
[176] A. Graindorge, C. Militti, F. Gebauer, Wiley interdisciplinary reviews. RNA, 2 (2011) 534-545.
[177] T. Afroz, Z. Cienikova, A. Clery, F.H. Allain, Methods in enzymology, 558 (2015) 235-278.
[178] I. Wang, J. Hennig, P.K. Jagtap, M. Sonntag, J. Valcarcel, M. Sattler, Nucleic acids research, 42 (2014) 5949-5966.
[179] H. Wang, F. Zeng, Q. Liu, H. Liu, Z. Liu, L. Niu, M. Teng, X. Li, Acta crystallographica. Section D, Biological crystallography, 69 (2013) 373-380.
[180] X. Wang, T.M. Tanaka Hall, Nature structural biology, 8 (2001) 141-145.
[181] E.A. Sickmier, K.E. Frato, H. Shen, S.R. Paranawithana, M.R. Green, C.L. Kielkopf, Molecular cell, 23 (2006) 49-59.
[182] A.A. Agrawal, E. Salsi, R. Chatrikhi, S. Henderson, J.L. Jenkins, M.R. Green, D.N. Ermolenko, C.L. Kielkopf, Nature communications, 7 (2016) 10950.
[183] E. Glasmacher, K.P. Hoefig, K.U. Vogel, N. Rath, L. Du, C. Wolf, E. Kremmer, X. Wang, V. Heissmeyer, Nat Immunol, 11 (2010) 725-733.
[184] V. Heissmeyer, K.U. Vogel, Immunol Rev, 253 (2013) 273-289.

[185] K.U. Vogel, S.L. Edelmann, K.M. Jeltsch, A. Bertossi, K. Heger, G.A. Heinz, J. Zoller, S.C. Warth, K.P. Hoefig, C. Lohs, F. Neff, E. Kremmer, J. Schick, D. Repsilber, A. Geerlof, H. Blum, W. Wurst, M. Heikenwalder, M. Schmidt-Supprian, V. Heissmeyer, Immunity, 38 (2013) 655-668.
[186] K.M. Jeltsch, D. Hu, S. Brenner, J. Zoller, G.A. Heinz, D. Nagel, K.U. Vogel, N. Rehage, S.C. Warth, S.L. Edelmann, R. Gloury, N. Martin, C. Lohs, M. Lech, J.E. Stehklein, A. Geerlof, E. Kremmer, A. Weber, H.J. Anders, I. Schmitz, M. Schmidt-Supprian, M. Fu, H. Holtmann, D. Krappmann, J. Ruland, A. Kallies, M. Heikenwalder, V. Heissmeyer, Nat Immunol, 15 (2014) 1079-1089.
[187] K. Leppek, J. Schott, S. Reitter, F. Poetz, M.C. Hammond, G. Stoecklin, Cell, 153 (2013) 869-881.
[188] L. Codutti, K. Leppek, J. Zalesak, V. Windeisen, P. Masiewicz, G. Stoecklin, T. Carlomagno, Structure, (2015).
[189] S. Basu, R.P. Bahadur, Cellular and molecular life sciences : CMLS, (2016).
[190] T. Bradley, M.E. Cook, M. Blanchette, Rna, 21 (2015) 75-92.
[191] M. Quesnel-Vallieres, M. Irimia, S.P. Cordes, B.J. Blencowe, Genes & development, 29 (2015) 746-759.
[192] J. Lee, E.A. Dethoff, H.M. Al-Hashimi, Proceedings of the National Academy of Sciences of the United States of America, 111 (2014) 9485-9490.
[193] I.J. Kimsey, K. Petzold, B. Sathyamoorthy, Z.W. Stein, H.M. Al-Hashimi, Nature, 519 (2015) 315-320.
[194] P. Vallurupalli, G. Bouvignies, L.E. Kay, Journal of the American Chemical Society, 134 (2012) 8148-8161.
[195] J.R. Williamson, Nature structural biology, 7 (2000) 834-837.
[196] H. Kim, S.C. Abeysirigunawarden, K. Chen, M. Mayerle, K. Ragunathan, Z. Luthey-Schulten, T. Ha, S.A. Woodson, Nature, 506 (2014) 334-338.
[197] J.C. Grigg, A. Ke, Methods in enzymology, 558 (2015) 213-232.
[198] A. Reining, S. Nozinovic, K. Schlepckow, F. Buhr, B. Furtig, H. Schwalbe, Nature, 499 (2013) 355-359.
[199] A. Brown, S. Shao, J. Murray, R.S. Hegde, V. Ramakrishnan, Nature, 524 (2015) 493-496.
[200] C. Bernecky, F. Herzog, W. Baumeister, J.M. Plitzko, P. Cramer, Nature, 529 (2016) 551-554.
[201] M. Zhou, C.M. Jones, V.H. Wysocki, Analytical chemistry, 85 (2013) 8262-8267.
[202] H. Zhang, W. Cui, M.L. Gross, R.E. Blankenship, FEBS letters, 587 (2013) 1012-1020.
[203] X. Yan, C.S. Maier, Methods in molecular biology, 492 (2009) 255-271.
[204] L. Konermann, J. Pan, Y.H. Liu, Chemical Society reviews, 40 (2011) 1224-1234.
[205] J. Rappsilber, Journal of structural biology, 173 (2011) 530-540.
[206] A.N. Holding, Methods, 89 (2015) 54-63.
[207] J. Dejardin, R.E. Kingston, Cell, 136 (2009) 175-186.
[208] M.I. Recht, S.P. Ryder, J.R. Williamson, Methods in molecular biology, 488 (2008) 117-127.
[209] Y. Tanabe, S. Nagatoishi, K. Tsumoto, Molecular bioSystems, 11 (2015) 2441-2448.
[210] I. Kim, H. Kwak, H.K. Lee, S. Hyun, S. Jeong, Nucleic acids research, 40 (2012) 6863-6872.
[211] J.M. Pagano, C.C. Clingman, S.P. Ryder, Rna, 17 (2011) 14-20.
[212] I. Wong, T.M. Lohman, Proceedings of the National Academy of Sciences of the United States of America, 90 (1993) 5428-5432.
[213] P.G. Stockley, Methods in molecular biology, 543 (2009) 1-14.
[214] D.C. Rio, Cold Spring Harbor protocols, 2012 (2012) 1078-1081.
[215] M. Jerabek-Willemsen, C.J. Wienken, D. Braun, P. Baaske, S. Duhr, Assay and drug development technologies, 9 (2011) 342-353.

[216] S.A. Seidel, P.M. Dijkman, W.A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J.S. Joseph, P. Srinivasan, P. Baaske, A. Simeonov, I. Katritch, F.A. Melo, J.E. Ladbury, G. Schreiber, A. Watts, D. Braun, S. Duhr, Methods, 59 (2013) 301-315.
[217] C.J. Wong, K. Launer-Felty, J.L. Cole, Methods in enzymology, 488 (2011) 59-79.
[218] D. Matulis, J.K. Kranz, F.R. Salemme, M.J. Todd, Biochemistry, 44 (2005) 5258-5266.
[219] K.R. Rupesh, A. Smith, P.E. Boehmer, Biochemical and biophysical research communications, 454 (2014) 604-608.
[220] M. Lorenz, Rna, 15 (2009) 97-103.
[221] T.R. Patel, G. Chojnowski, Astha, A. Koul, S.A. McKenna, J.M. Bujnicki, Methods, (2016).
[222] A. Bhardwaj, M.P. Myers, E. Buratti, F.E. Baralle, Nucleic acids research, 41 (2013) 5062-5074.

Figure legends

Fig. 1. Workflow for the identification and characterization of protein-RNA interactions. A) The coding sequence (CDS) in a pre-mRNA transcript is usually embedded in a short 5'-UTR and a more complex 3'-UTR. Linear or structured *cis* regulatory elements are recognized by *trans*-acting RNA binding proteins. **B)** Identification of RNA targets and their cognate RNA binding proteins (see Fig. 2) is a prerequisite for structural and mechanistic studies of an RNP. **C)** Biochemical, biophysical and structural biology techniques provide complementary information to characterize molecular features of an RNP *in vitro*. Information from high-resolution structural analysis of the RNP is confirmed by functional studies using *in vivo* assays.

Fig. 2. Identification of minimal RNA and protein binding regions. A) Computational, bioinformatic and experimental methods can be used to define the boundaries of an RNA binding domain in a full-length RBP. Note, that the given sources do not well include RNA-binding IDRs of a protein as those will normally escape experimental approaches and computational predictions of a domain. **B)** Identification of RNA target motifs that are recognized by an RBP. Top, molecular biology allows to identify a functionally important RNA region by deletion mapping. A reporter gene assay is used to obtain a functional readout. The RNA region identified is then subject to further structural and biophysical studies. Bottom, the secondary structure and surface accessibility of folded RNAs can be experimentally analyzed with chemical probing methods. Here, SHAPE identifies regions of non-paired and/or dynamic nucleotides. Combined with computational analysis structures of RNAs can be predicted from both *in vitro* and *in vivo* SHAPE. **C)** Broad-scale investigations of RNP complexes to define interaction sites. Top, CLIP approaches are used to crosslink an RBP to genome-wide RNA targets, followed by sequencing and bioinformatics analysis of the cross-linked RNA sequences to yield consensus binding motifs. Bottom, SELEX is widely applied to systematically enrich and identify high-affinity RNA ligands for an RBP. Bioinformatic analysis identifies linear consensus motifs and predicts the potential presence of structured motifs. The identified motifs can then be mapped on a genome-wide scale.

Fig. 3. Mechanistic features of protein-RNA recognition. A) Examples for the recognition of linear RNA motifs by tandem RBDs are shown. These can act independently or present a pre-formed rigid tandem RBD arrangement, which can cause looping of a single-stranded RNA ligand. **B)** Cooperativity of multiple RBDs enables the formation of high affinity and specific RNPs (see Fig. 4 for an example). This can involve avidity effects due to multivalent interactions. In a multi-domain RBPs one of the RBDs (for example a zinc finger domain) with low RNA binding affinity and specificity can promote sliding along a linear RNA to help identifying a high-affinity RNA motif to be specifically recognized by a second RBD. **C)** Conformational dynamics of RNA and/or protein components of an RNP often contributes to the formation of an RNA-protein complex. The conformational changes can mechanistically involve (mutually) induced fit or conformational selection mechanisms. **D)** RNP dynamics can involve RBD sliding along ssRNA or dsRNA molecules (top), mainly implying a lack of sequence specificity.

Fig. 4. NMR tools to determine RNA folding and RNP complex formation. A) Schematic drawing of a one-dimensional imino proton spectrum. Visible imino signals indicate the presence of hydrogen-bonding and thus base-pairing in a structured RNA. The experiment is a direct readout for RNA secondary structure. **B)** Imino signals (only detectable in base pairs) that are affected by the binding of an RBP experience changes in chemical shifts and/or line-widths

and thus report about the binding site on the RNA. **C)** Schematic drawing of a $^{1}H,^{15}N$ correlation experiment in proteins which provides one correlation cross peak per amide proton and thus amino acid. Titration of RNA into a ^{15}N-isotope-labeled protein sample yields chemical shift perturbations at the site of interaction and thus allows for mapping of the RNA-protein interface. Both experiments yield complementary residue-resolved information about the protein-RNA interface in an RNP.

Fig. 5. Cooperativity during RNA recognition by multiple RBDs. A) The *msl-2* RNA is only present in male animals as it is required for dosage compensation for the lack of one X chromosome. It has been shown to form a trimeric complex with tandem RRMs from SXL and a Unr CSD. **B)** SLS is used to validate the complex MW and homogeneity, and thus stoichiometry of the complex. Here, the predominant 1:1:1 fraction was used for structural studies. **C)** The graph on the left shows SANS data for the ternary SXL-CSD1-RNA complex (adapted from [97]). A SANS derived sphere model shows density for SXL (green) UNR-CSD1 (blue) and RNA (magenta), providing low-resolution information on the relative position of individual subunits. **D)** SAXS data confirm the overall shape of the SANS derived model in C. Shown is the overlay of an experimental curve (blue) with the theoretically back-calculated curve obtained from the crystal structure (red). The χ^2 value reports about an excellent fit. **E)** Crystal structure of the ternary complex reveals the atomic details of the SXL-Unr-*msl2* complex. The triple zipper region is indicated. **F)** $^{1}H,^{15}N$ correlation protein NMR spectra reveal the cooperativity in binding of SXL and Unr to *msl-2* as indicated by the two residues located in the Unr CSD. Colors are: CSD alone, blue; CSD+RNA, red; ternary complex, green. Arrows indicate trajectories of chemical shift changes and show that residues are differently affected by the subsequent titration of SXL RRMs and *msl-2* RNA to the Unr CSD. **G)** ITC measurements dissect binding contributions from the two proteins and confirm the 1:1:1 stoichiometry. Shown is the titration of SXL RRMs to a pre-formed complex of Unr CSD with RNA.

Fig. 6. Dynamic shifts of domain arrangements during linear-motif-recognition by tandem RRMs. A) NMR PRE experiments have been used to reveal different domain arrangements of U2AF65 RRM1-RRM2 alone (left panel, the spin label affects both domains in the closed conformation) or in complex with U9 RNA (right panel, only RRM2 is affected in the open conformation). The green dot represents the spin label with its affected sphere. Distances are translated into a scale between 0 (minimum) and 1 (distal, unaffected). **B)** NMR titrations show that strong (U9) or weak (U4A8U4) RNA ligands will provoke different CSPs illustrating different contributions of RRM1 and RRM2. While U9 RNA clearly involves binding to both RRMs in an open conformation, U4A8U4 RNA shifts the conformational equilibrium between open and closed states only partially as indicated by reduced CSPs observed for RRM1. **C)** In line with NMR data, FRET measurements reveal open, closed and a dynamic intermediate state (with subpopulations of the open and closed conformations) of the U2AF65 RRM1,2 arrangements depending on the presence of no (grey), weak (blue) or strong (red) RNA ligands. The dashed lines indicate reference positions for the fully open or closed states, respectively. The FRET fluorophores (green and red labels) were attached pairwise with one moiety at each RRM. Shorter distances between them will increase the FRET efficiency. The panel is adapted from [93]. Panels A and B are adapted from [90].

Fig. 7. Integrated approach to identify and characterize Roquin-RNA stem-loop interactions. A) Identification of the two protease-protected globular RNA-binding protein fragments from an N-terminal Roquin fragment resembling the core and extended ROQ domains, respectively, by limited proteolysis [25]. The arrow indicates the fragment that has

been used in later NMR and crystallographic studies. **B)** The secondary structure of the CDE stem-loop, a known target of Roquin, is confirmed by imino NMR spectra [25]. Red bars between bases in the scheme represent detectable H-bonds. Visible G and U bases are numbered. Dashed lines in the spectrum indicate the 'imino walk' based on cross peaks between neighboring H-bonds. **C)** SELEX was used to identify a novel ROQ target sequence [26] termed ADE, which resembles a U-rich hexa-loop RNA hairpin. **D)** The secondary structure of the novel ADE stem-loop is confirmed by imino NMR spectra analogously to the CDE in panel B. **E)** Comparison of ROQ-RNA stem-loop co-crystal structures with the novel ADE or the CDE motifs. The closing base-pairs of the RN hairpins are annotated; the unpaired loop bases are shown in comparison.

Fig. 8. Integrated structural biology approaches applied to macromolecular RNP complexes. A) Cryo-EM map of the Tetrahymena telomerase holoenzyme complex (Electron Microscopy Data Bank accession number EMD-6442). The panel is adapted from reference [135] with permission from the original publisher. Individual high-resolution structures obtained from X-ray crystallography and NMR are modelled into the map. With this, a global picture of the complex is created while either method alone does not sufficiently report about the arrangement of proteins and RNA. The structure -albeit still incomplete- allows deriving detailed mechanisms of the enzyme action at the atomic level. **B)** Example for the combined use of local and long-distance restraints in solution-based methods. While local restraints (up to 6 Å) are obtained from NOEs in NMR NOESY spectra (here within protein monomers), EPR yields restraints of up to 100 Å and is thus well-suited for RNAs up to 100 nts in size. In this study different pairwise spin-labelled RNAs have been used to obtain distance restraints within the triple-homodimeric RsmE bound to RsmZ RNA. Both methods together provide a useful mixture of restraints obtained in solution. EPR also allowed to confirm the presence of two conformations, which makes it ideally complementary to NMR. The shown structure refers to PDB entry 2MF1 [147]. **C)** Method integration for the structure determination of the apo-box C/D methylation complex. While SAXS/SANS data were used to restrict the conformational space a de novo complex structure determination was undertaken based on individual protein start structures (L7Ae, fibrillarin and the Nop platform consisting of two copies in light and dark grey, respectively) and a set of distance restraints between them from PRE experiments using state of the art isotope labelling (see arrows for examples). The dimeric RNA arrangement was determined with SANS measurements. Notably, NMR has also been used to visualize the RNA methylation event. The PDB code is 4BY9; the panel has been adapted from [145] with permission from the original publisher. **D)** The combination of SAXS with a low-resolution X-ray structure, H/D-exchange coupled to mass spectrometry and RNA SHAPE has been used to determine the structure of the HIV-1 Tat protein (red) bound to TAR RNA (green) and the human proteins CDK9 (blue), AFF4 (grey) and CycT1 (yellow). The PDB code is 5L1Z [171]. SHAPE yields information about the RNA stability in dependence of bound protein and thus its interaction with the complex. Quantification of H/D exchange data show that binding of RNA leads to the stabilization of adjacent fragments in AFF4 and CycT1.

Figure 1

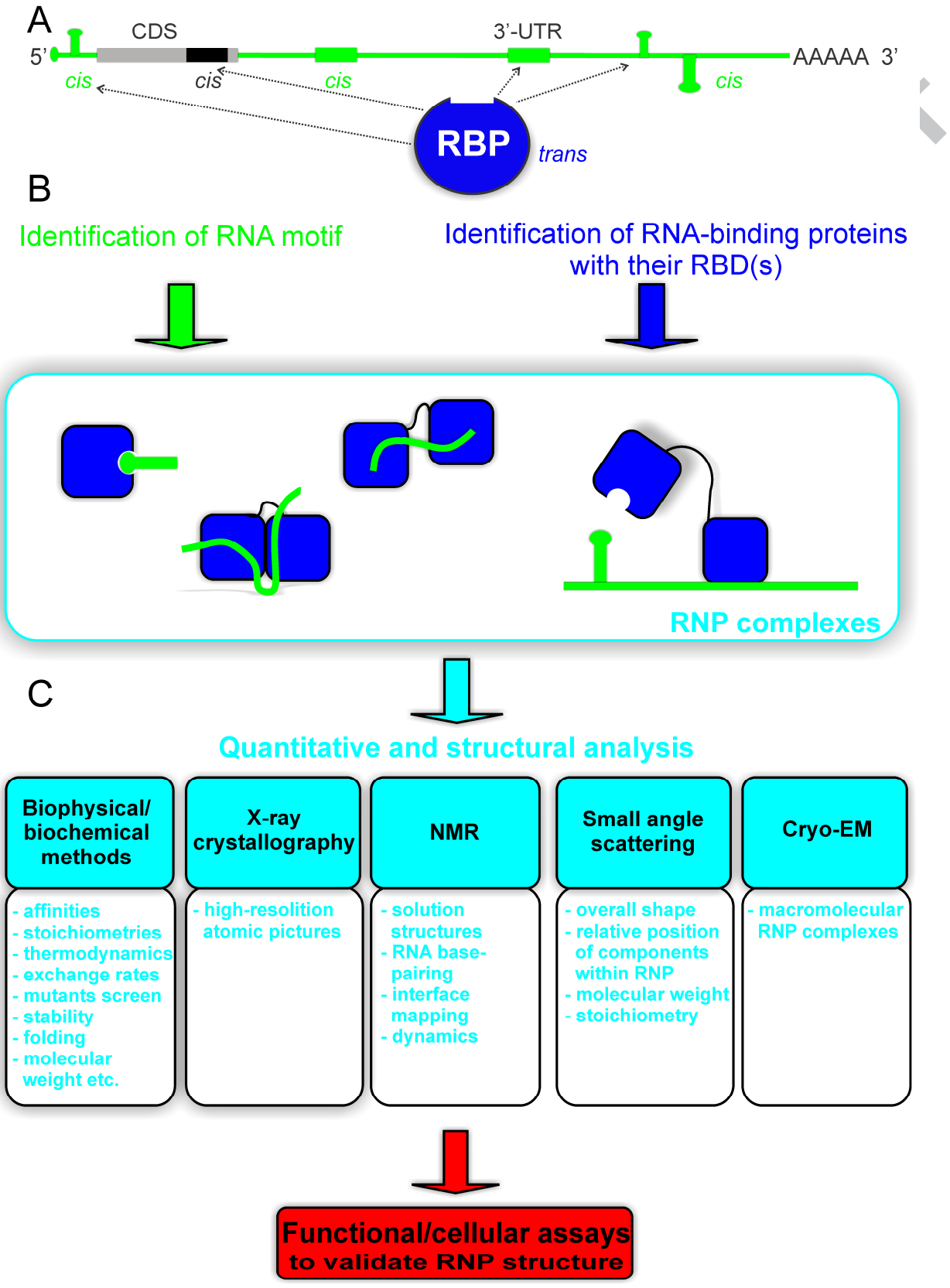

A
CDS
3'-UTR
5'
AAAAA 3'
cis
cis
cis
cis
RBP
trans
B
Identification of RNA motif
Identification of RNA-binding proteins with their RBD(s)
RNP complexes
C
Quantitative and structural analysis
Biophysical/ biochemical methods
- affinities
- stoichiometries
- thermodynamics
- exchange rates
- mutants screen
- stability
- folding
- molecular weight etc.
X-ray crystallography
- high-resolition atomic pictures
NMR
- solution structures
- RNA base-pairing
- interface mapping
- dynamics
Small angle scattering
- overall shape
- relative position of components within RNP
- molecular weight
- stoichiometry
Cryo-EM
- macromolecular RNP complexes
Functional/cellular assays to validate RNP structure

Figure 2:

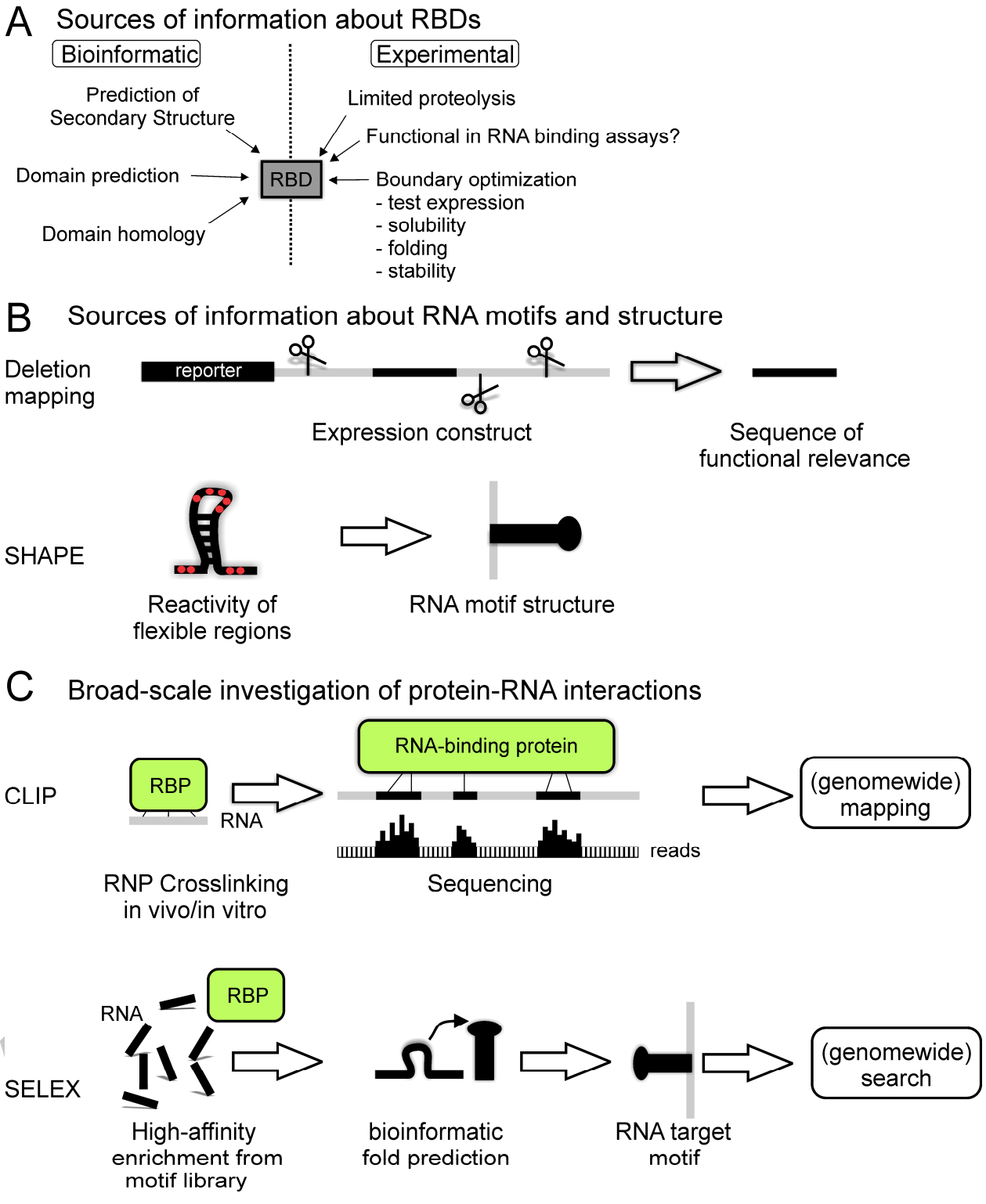

A
Sources of information about RBDs
Bioinformatic
Experimental
Prediction of Secondary Structure
Limited proteolysis
Functional in RNA binding assays?
Domain prediction
RBD
Boundary optimization
- test expression
- solubility
- folding
- stability
Domain homology
B
Sources of information about RNA motifs and structure
Deletion mapping
reporter
Expression construct
Sequence of functional relevance
SHAPE
Reactivity of flexible regions
RNA motif structure
C
Broad-scale investigation of protein-RNA interactions
CLIP
RBP
RNA
RNA-binding protein
reads
(genomewide) mapping
RNP Crosslinking in vivo/in vitro
Sequencing
SELEX
RNA
RBP
High-affinity enrichment from motif library
bioinformatic fold prediction
RNA target motif
(genomewide) search

Figure 3

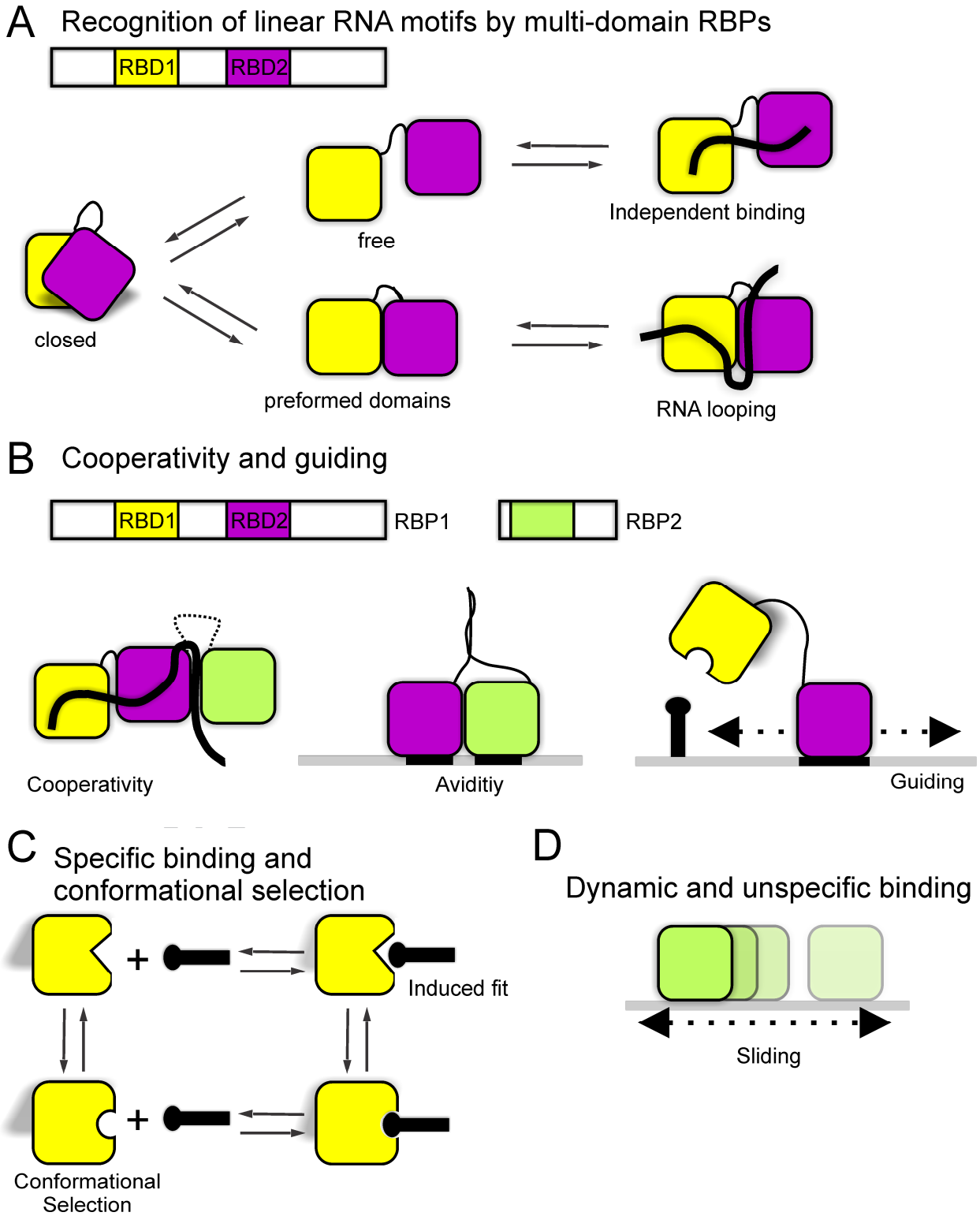
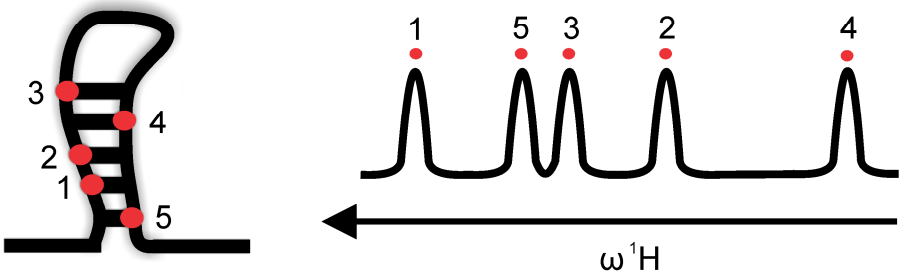
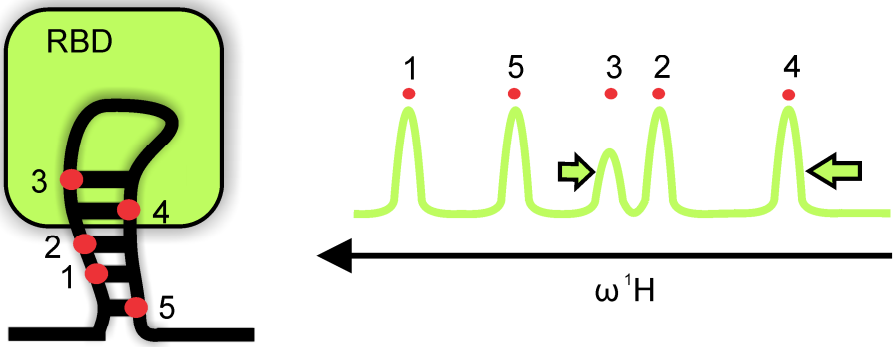

A Recognition of linear RNA motifs by multi-domain RBPs
RBD1
RBD2
free
Independent binding
closed
preformed domains
RNA looping
B Cooperativity and guiding
RBD1
RBD2
RBP1
RBP2
Cooperativity
Aviditiy
Guiding
C Specific binding and conformational selection
Induced fit
Conformational Selection
D Dynamic and unspecific binding
Sliding

Figure 4

A Imino protons as a probe for secondary structure



B Binding site mapping by imino protons



C Binding site mapping by protein amide correlation spectra

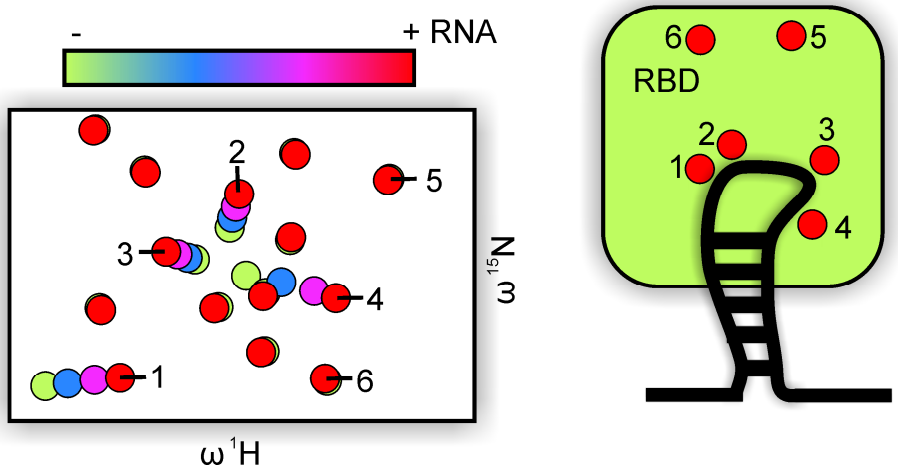
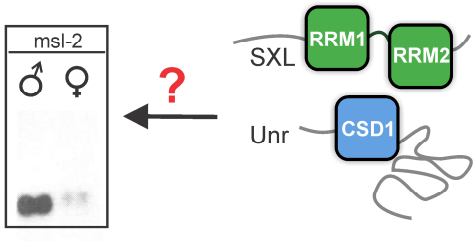
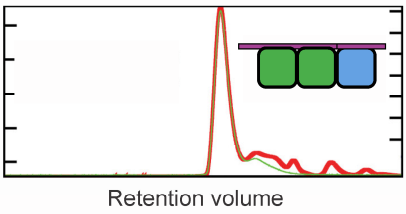


Figure 5

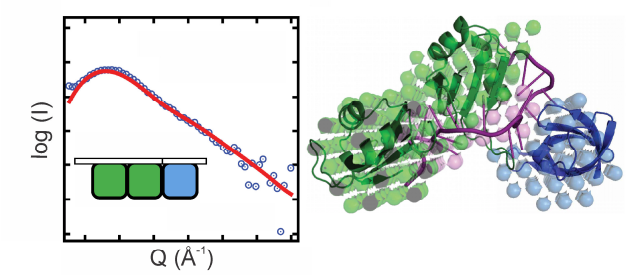
A Target identification by Northern Blotting



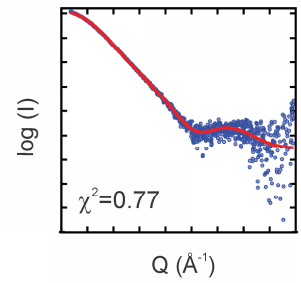
B Static light scattering reveals 1:1:1 stoichiometry



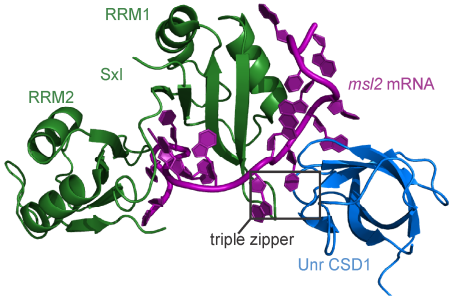
C Domain arrangement by SANS: Sandwich-like binding of RNA by SXL and Unr



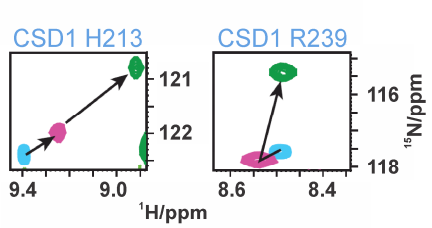
D Overall complex shape by SAXS



E Crystal structure of ternary complex: confirms postulated binding mechanism; reveals key residues



F CSP: different shifts caused by RNA and Sxl reveal cooperativity



G Affinity by ITC: highes affinity observed for simultaenous binding

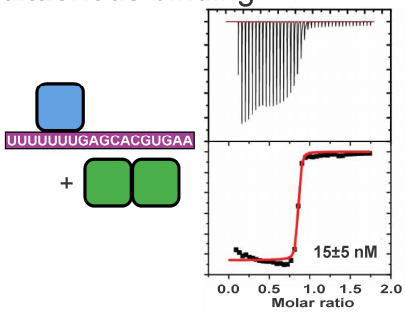
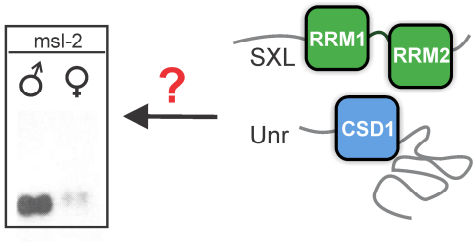
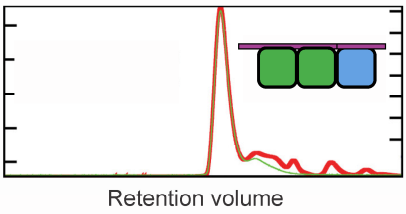


Figure 5

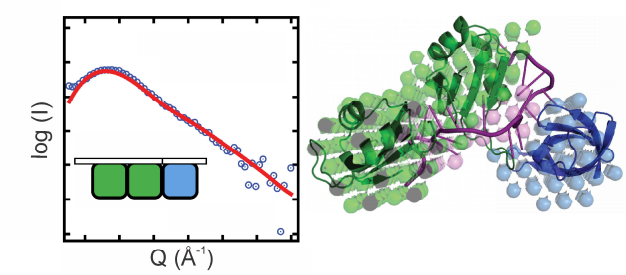
A Target identification by Northern Blotting



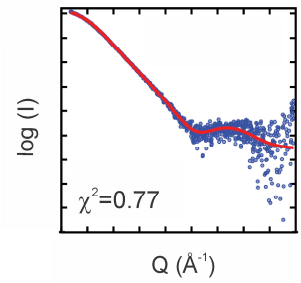
B Static light scattering reveals 1:1:1 stoichiometry



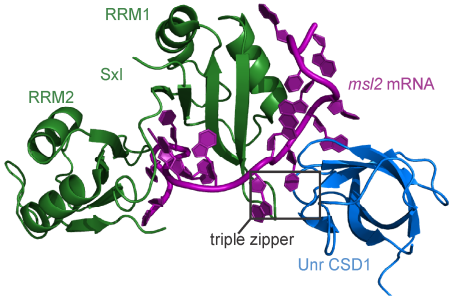
C Domain arrangement by SANS: Sandwich-like binding of RNA by SXL and Unr



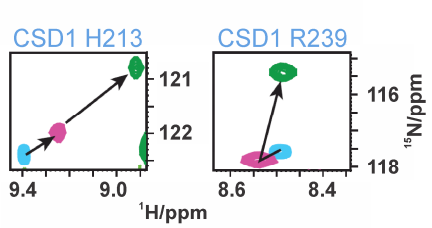
D Overall complex shape by SAXS



E Crystal structure of ternary complex: confirms postulated binding mechanism; reveals key residues



F CSP: different shifts caused by RNA and Sxl reveal cooperativity



G Affinity by ITC: highes affinity observed for simultaenous binding

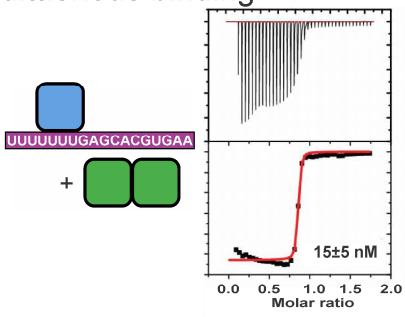
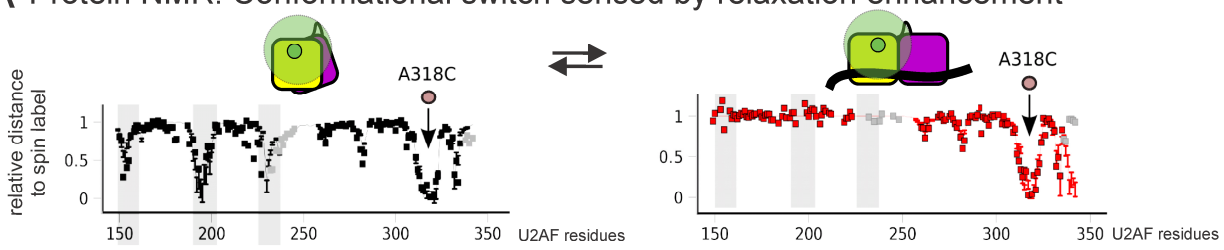
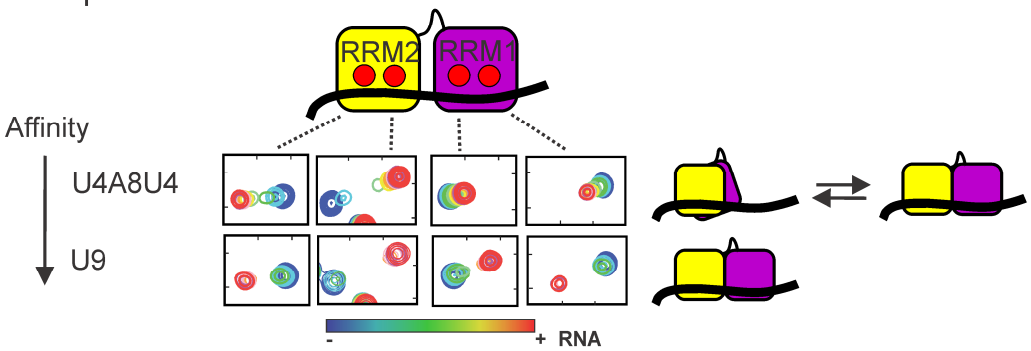


Figure 6

A Protein NMR: Conformational switch sensed by relaxation enhancement



B Protein NMR: Amide chemical shift perturbations as probes to reflect RNA-dependent conformational transitions



C FRET reveals open-closed arrangement of domains in dependence of RNA

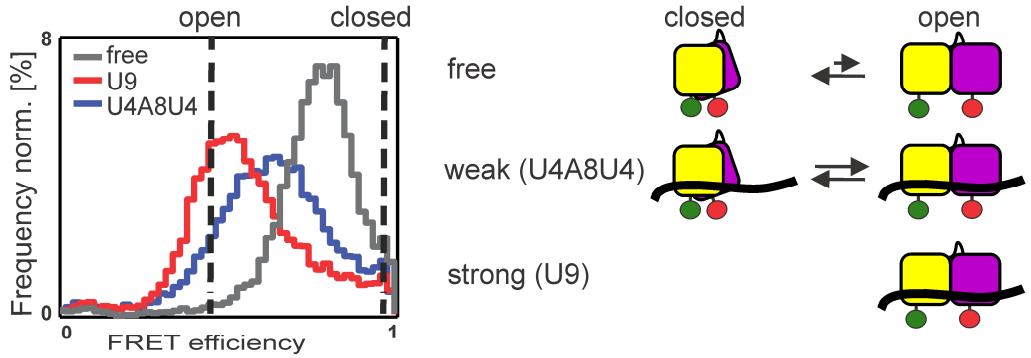


Figure 7

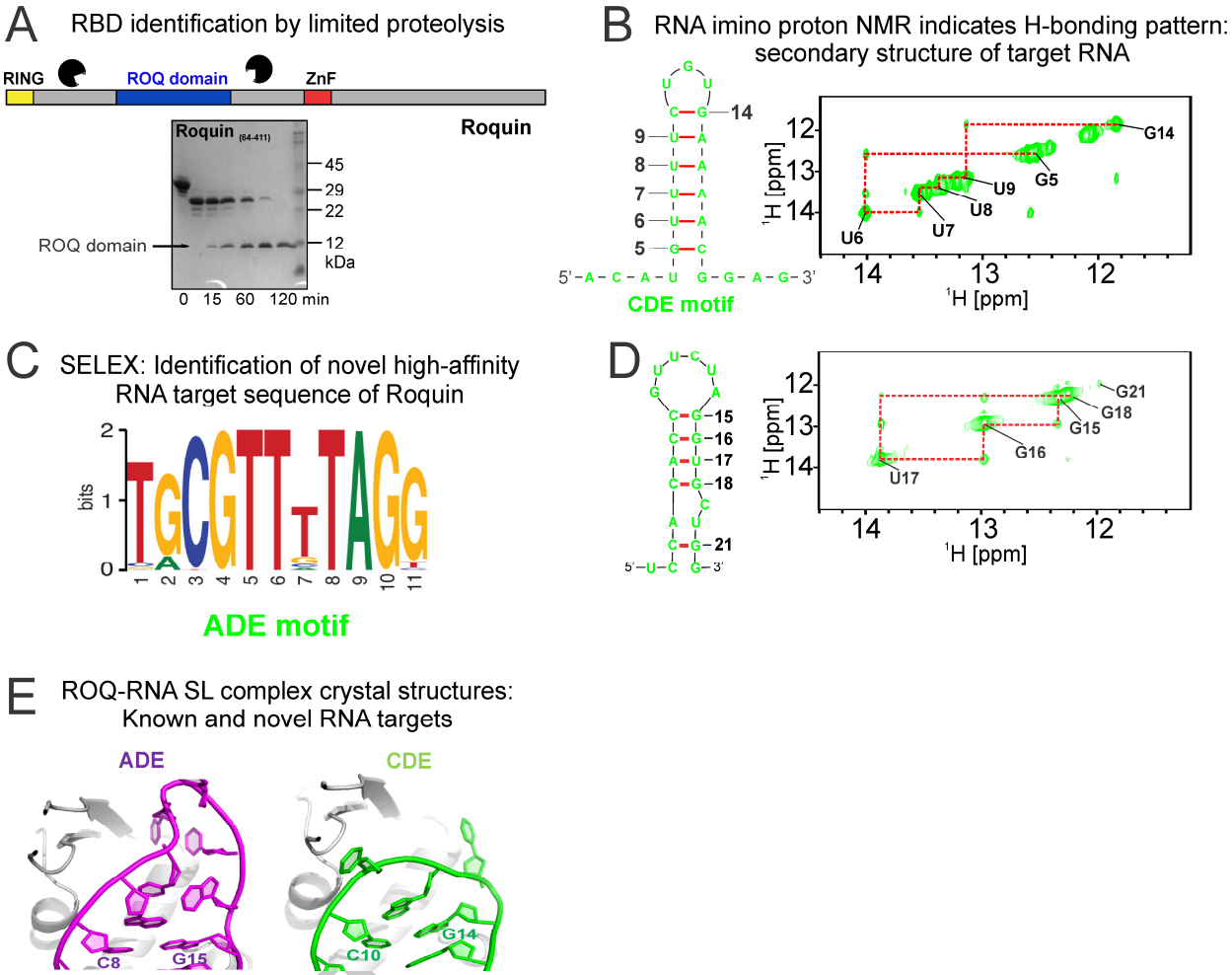


Figure 8

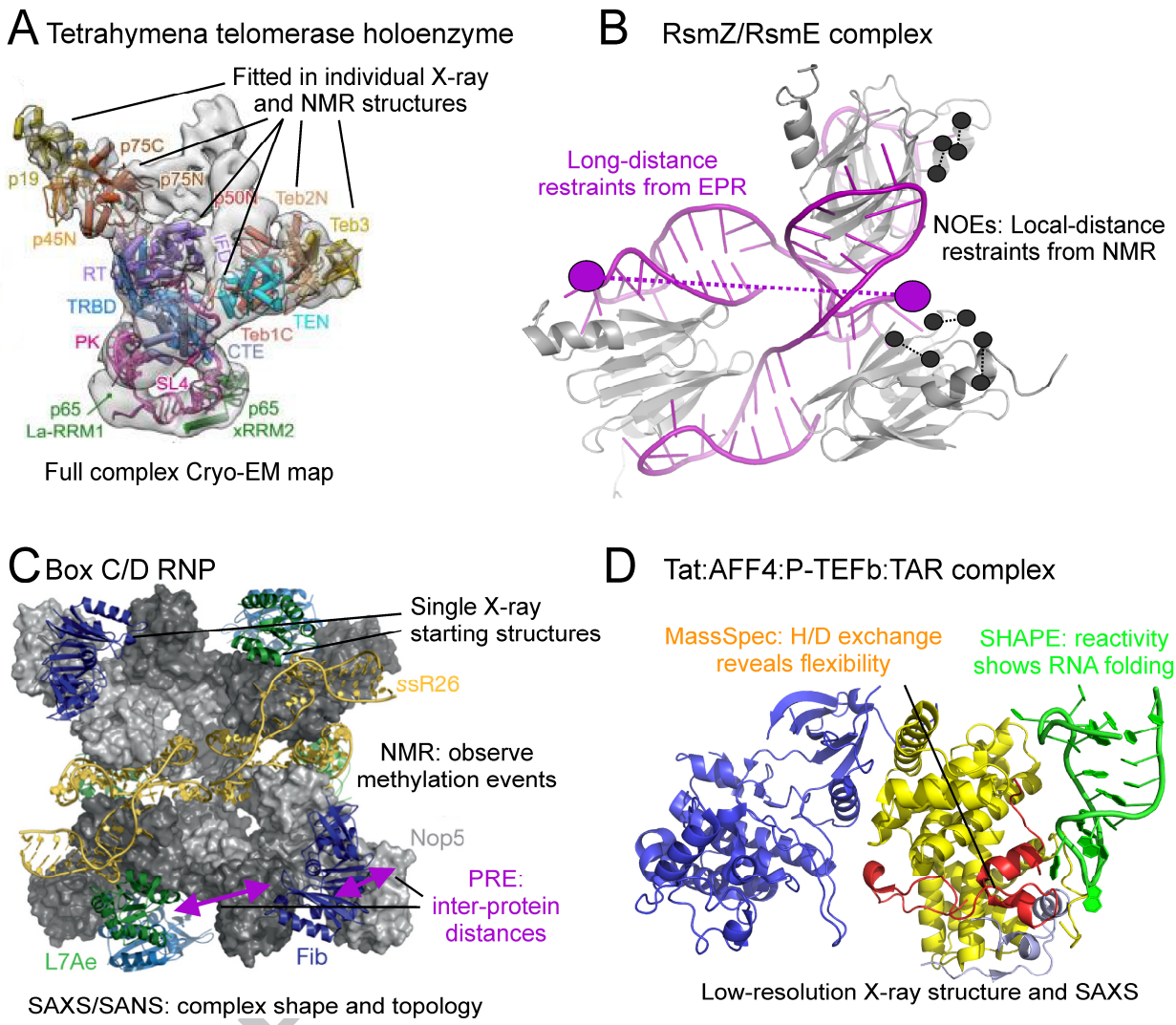


Table 1. Biophysical techniques to probe RNA-protein interactions and/or additional interactions with small molecules *in vitro* (see references for further reading or example studies in brackets)

Method	**Information**	**Observation**	**Pros**	**Cons**
Isothermal titration calorimetry (ITC) [208, 209]	K_D, stoichiometry, thermodynamic parameters	measures heat production complex formation in a titration experiment	direct measurement of reaction heat, multi-parameter fit, multiple binding transitions visible	not suitable for dynamic RNAs, sample stirring, requires relatively large sample amounts, needs change in net enthalpy, low throughput, some buffer limitations
Surface plasmon resonance (SPR) [26, 210]	k_{on}, k_{off}, K_D	detects changes of refractive index during binding	different ligands can be tested simultaneously with one chip, low sample amount needed, all affinity ranges	surface binding, needs immobilization of RNA or protein (can interfere with binding), high-affinity bias, artefacts from mass transport
Fluorescence anisotropy and rotational correlation spectroscopy [211]	K_D	measure fluorescence-changes linked to rotational diffusion	High-throughput possible	needs active groups without overlap in absorption, significant increase in mass is needed, dye might interfere with binding
Electromobility shift assay (EMSA) [42, 112]	K_D, (complex stoichiometry)	measures change in electrophoretic mobility in PAGE	multiple, stepwise binding observable, very low concentrations needed, only relative K_Ds are reliable	usually radioactive or fluorescent labelling needed, not suitable for low affinities, large RNAs difficult to quantify
Filter binding assay (FBA)[212-214]	K_D or relative affinity in multiple setup, eventually k_{on}/k_{off}-rates	detects absorption of formed RNP complex on cellulose filter	suitable for screening multiple RNA ligands, low sample amount	prone to unspecific interactions with filter, radioactive labelling needed
Microscale thermophoresis [215, 216]	K_D, stoichiometry, homogeneity	change in motion of molecules in temperature gradient	possible for very small ligands, native-like, multiple binding sites, cooperativity	needs fluorescent labels for RNP complex measurements
Analytical ultracentrifugation [217]	stoichiometry (K_D), multiple sample states	sedimentation	very accurate, distribution of multiple complex states in equilibrium	needs ultracentrifuge
Thermal shift assay [218, 219]	melting point, complex stability	measures melting points by detection of hydrophobic patches induced by thermal unfolding	precise, cheap, HT possible, suitable for buffer screen	difficult for simultaneous or multiple unfolding events, requires labeling or good sensitivity w/o labeling
Förster resonance energy transfer (FRET) [196, 220]	dynamics, distances, populations	measures FRET efficiency	ideal for domain motions, dynamic systems, to detect multiple conformations	needs fluorescent dye labels, which are bulky and hydrophobic and can interfere
Static light scattering (SLS) [221]	molecular weight, stoichiometry	measures refractive index	easy, precise, usually includes integrated purification	only for medium / high affine complexes

Dynamic light scattering (DLS) [221]	molecular weight, stoichiometry, melting point	uses light scattering of molecules	easy, can detect multiple species	needs very pure samples, otherwise misleading
Circular dichroism (CD) [222]	secondary structure, bending of RNA/DNA helices, melting points	detects structural alignments of chiral features, i.e. secondary structure in proteins	fast and easy, folding transitions can be monitored	only global effect of sample can be measured

- We illustrate methods to study the structure and dynamics of protein-RNA complexes.
- The identification and principles of protein-RNA complex formation are summarized.
- Their various challenges for structural and mechanistic studies are discussed.
- We provide examples for the successful integration of methods to study RNPs.