Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

Dawn of the *in vivo* RNA structurome and interactome

Chun Kit Kwok[1,2,3]

[1]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.; [2]Cancer Research UK, Cambridge Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, U.K.; and [3]Department of Biology and Chemistry, City University of Hong Kong, Kowloon Tong, Hong Kong SAR, China

Correspondence: Chun Kit Kwok (ckk26@cam.ac.uk or ckkwok42@cityu.edu.hk)

RNA is one of the most fascinating biomolecules in living systems given its structural versatility to fold into elaborate architectures for important biological functions such as gene regulation, catalysis, and information storage. Knowledge of RNA structures and interactions can provide deep insights into their functional roles *in vivo*. For decades, RNA structural studies have been conducted on a transcript-by-transcript basis. The advent of next-generation sequencing (NGS) has enabled the development of transcriptome-wide structural probing methods to profile the global landscape of RNA structures and interactions, also known as the RNA structurome and interactome, which transformed our understanding of the RNA structure–function relationship on a transcriptomic scale. In this review, molecular tools and NGS methods used for RNA structure probing are presented, novel insights uncovered by RNA structurome and interactome studies are highlighted, and perspectives on current challenges and potential future directions are discussed. A more complete understanding of the RNA structures and interactions *in vivo* will help illuminate the novel roles of RNA in gene regulation, development, and diseases.

RNA and its structure

RNA molecules and their structures have important roles in diverse cellular processes such as transcription, splicing, polyadenylation, translation, localization, stability, and decay [1]. RNA can perform catalytic and regulatory functions, either alone or in combination with others (e.g. metabolites, proteins, RNAs, and DNAs) [2–4]. For example, ribozymes are RNA enzymes that can catalyze specific biochemical reactions (e.g. cleavage and ligation); riboswitches are regulatory RNAs that can sense the identity and concentration of metabolites or metal ions, and alter their RNA conformations to affect gene expression (e.g. transcription and translation) [5].

The linear combination of RNA bases constitutes the primary structure of RNA (or RNA sequence), and its single-stranded (ss) nature enables the formation of diverse secondary and tertiary structures [6], such as hairpins, kink-turns, G-quadruplexes, and pseudoknots, for its proper function in cells. There is growing evidence, suggesting that RNA structures are influenced by a variety of factors, including but not limited to physiochemical environment (e.g. temperature, metal ions, and metabolites), sequence context (e.g. mutations and native RNA modifications), RNA–protein interactions (e.g. annealers, helicases, and chaperones), and RNA–RNA intramolecular and intermolecular contacts (Figure 1) [7–12]. As such, knowledge about their *in vivo* RNA structures and interactions help to delineate their biological roles, allow us to better understand the RNA-mediated biochemical mechanisms, and facilitate the development of chemical or molecular tools for diagnostic, therapeutic, and gene manipulation purposes.

Received: 22 March 2016
Revised: 19 June 2016
Accepted: 4 July 2016

Version of Record published: 19 October 2016

Phylogenetic comparative sequence and co-variation analysis in RNA

RNA secondary structure can be derived based on phylogenetic comparative sequence and co-variation analysis [13]. The rationale of this approach is based on the notion that RNA structure

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

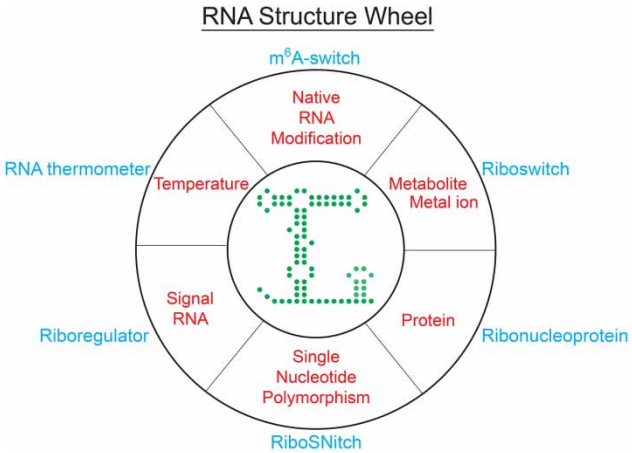


Figure 1. RNA structure wheel.

RNA structures are dynamic and can be influenced by many factors, which can be classified into specific groups such as riboswitches (response to metabolites and metal ions), RNA thermometers (response to temperature), riboSNitches (response to single-nucleotide polymorphism), m^6A-switches (response to native RNA modifications such as m^6A), RNPs (response to proteins), and riboregulators (response to signal RNAs).

defines it function, meaning that throughout the selective pressure in evolution, the RNA of interest, e.g. ribosomal RNAs (rRNAs), preserves its RNA secondary structure to perform its precise function among different species [14]. Sequence variation and co-variation often occur in orthologous genes across different species, which provide useful structural information regarding the RNA of interest. Phylogenetic comparative sequence and co-variation analysis are used to identify nucleotides with similar patterns of sequence variation in an alignment of RNA sequences, e.g. orthologs. Compensatory base pair co-variation, such as from one Watson-Crick to another Watson-Crick pair or from a Watson-Crick to a wobble pair, is a strong indication of conserved RNA helix and secondary structure. Phylogenetic analysis was used to deduce the structural model of tRNAs (transfer RNAs) and rRNAs [15–17], and these proposed models were later shown to be in good agreement with X-ray crystal structure of ribosome [18]. This strategy is useful for structure prediction of well-conserved RNAs among species, e.g. tRNA, rRNA, and some protein-coding RNAs; however, some RNAs such as long non-coding RNAs (lncRNAs) lack the sequence conservation [19], and thus, experimental approaches are needed to identify the RNA structure.

Enzymatic probing of RNA structure

Structure determination techniques such as X-ray crystallography, nuclear magnetic resonance spectroscopy, and modern cryo-electron microscopy allow the experimental characterization of RNA structures and complexes at high resolution [20]. Nonetheless, most RNAs are of low abundance in cells and are often too long and dynamic to be studied by these methods. Over the past few decades, enzymatic and chemical probing have been routinely used to obtain valuable RNA secondary and tertiary structural information [21,22].

Enzymatic probing utilizes ribonucleases (RNases) that selectively cleave either ss RNA regions or double-stranded (ds) RNA regions to examine the nucleotide accessibility and base-pairing propensity of the RNA of interest *in vitro*, and thereby infer its secondary structure. Commonly used RNases include: RNase V1 (dsRNA-specific), RNase S1 (ssRNA-specific), RNase A (ssC/U-specific), and RNase T1 (ssG-specific) (see refs [21,23] for more information). These RNases can robustly detect structural features of RNA and identify the protein footprint of RNA–protein complex under diverse reaction conditions (e.g. different temperatures and pH). They have been applied extensively to the study of rRNAs and other RNAs [24–26], and the results are shown to be consistent with the structure model derived from phylogenetic analysis [24]. RNases are generally large (>10 000 Da) [21] and thus are sensitive to steric hindrance such as the presence of interacting proteins or certain folding of the RNA. When compared with small chemical probes, they are bulky and somewhat less sensitive in detecting small defects in RNA such as bulges and mismatches. In addition, RNases are largely limited to the study of RNA structures *in vitro* as they cannot cross cell membranes. As the structures of RNA

in vivo are often different from *in vitro* [27,28], an alternative approach, such as chemical probing, is utilized to detect the folding of RNA in cellular environment.

Chemical probing of RNA structure

Chemical probing employs small molecular probes (<500 Da) that can detect RNA structural fingerprints and penetrate cell membranes readily [21,22], thus allowing an *in vivo* examination of RNA structure at nucleotide resolution. Unlike RNases, both the steric and electrostatic environment affects the reactivity of chemical probes [29]. Nucleobase-specific probes, such as dimethyl sulfate (DMS), alkylates the unprotected N1 position of adenine (N_1A), unprotected N3 position of cytosine (N_3C), and unprotected N7 position of guanine (N_7G) (Table 1); 1-cyclohexyl-(2-morpholinoethyl)carbodiimide metho-*p*-toluene sulfonate (CMCT) modifies the unprotected N3 position of uracil (N_3U) and the unprotected N1 position of guanine (N_1G) (Table 1); 3-ethoxy-1,1-dihydroxy-2-butanone (Kethoxal) modifies the unprotected N1 and unprotected exocyclic amine positions of guanine in RNA (N_1G and $N_{(2)}$G) (Table 1). These chemicals report the base-pairing status of the associated bases and have been shown to accurately map the secondary structure and protein-binding sites of RNAs such as rRNAs [30,31] and snRNAs [32,33]. In particular, DMS has been applied to probe *in vivo* structures of many RNAs across different organisms [27,33,34].

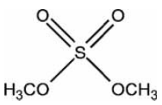
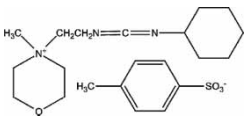
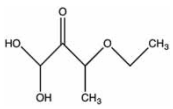
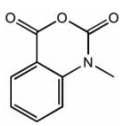
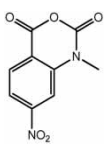
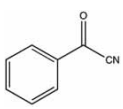
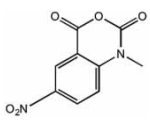
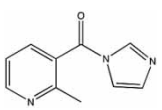
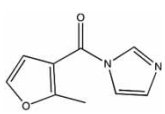
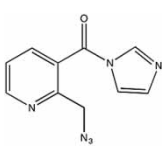
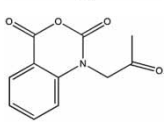
Ribose-specific probes, such as selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) [35], acylate the flexible C2′-hydroxyl group of the ribose (C_2′-OH) and report the local structural environment at each nucleotide (Table 1). Generally, flexible 2′-OH has higher propensity to adopt specific local conformations for acylation to occur [36], thus yielding higher SHAPE reactivity, e.g. unpaired nucleotides. One key advantage of SHAPE reagents over nucleobase-specific probes is that it targets the ribose of all four bases, producing more structural information in a single experiment. SHAPE reagents also react readily with water at different rates and therefore each reagent has a reaction half-life in aqueous solution [37]. Commonly used SHAPE reagents include *N*-methylisatoic anhydride [NMIA, $t_{1/2}$ hydrolysis (37°C) = 260 s], 1-methyl-7-nitroisatoic anhydride [1M7, $t_{1/2}$ hydrolysis (37°C) = 14 s], benzoyl cyanide [BzCN, $t_{1/2}$ hydrolysis (37°C) = 0.25 s], and 2-methylnicotinic acid imidazolide [NAI, $t_{1/2}$ hydrolysis (37°C) = 33 min] (Table 1). Depending on the biological question and application, caution is needed when choosing the appropriate SHAPE reagents. For example, fast-reacting SHAPE reagents, such as BzCN and 1M7, are desirable when studying time-resolved RNA-folding dynamics and ribonucleoprotein (RNP) assembly [38,39]; otherwise, a higher dose of a slow-reacting reagent and an external quench reagent is needed [27]. *In vivo* compatible SHAPE reagents, such as NAI and 1M7, should be used when probing RNA structures in cells [28,40].

Another important class of ribose-specific probe is the hydroxyl radical, which targets the solvent accessible C4′ hydrogen group of the ribose (C_4′-H) and reports the tertiary fold of the RNA of interest at single-nucleotide resolution [41,42] (Table 1). Hydroxyl radicals can be produced using iron(II) EDTA and hydrogen peroxide (Fenton's reagent) [43]. Alternatively, hydroxyl radicals can be generated in live cells using X-rays from a synchrotron source [44,45]. A high flux (>10^{16} photons/s) beam with tens of milliseconds exposure time can produce sufficient hydroxyl radicals to probe the RNA structure *in vivo*. The short exposure time required makes this chemical probe ideal to study the rapid kinetics of RNA assembly and RNA conformational dynamics [44,46]. Other chemical probes are also available, such as lead(II) ion [47,48], which cleave ss regions, loops, and bulges of RNA. Besides modifying RNAs, most chemical probes also react with other biomolecules, such as proteins and DNAs, which may affect the stability of protein–RNA complexes and DNA–RNA interactions; thus, proper control experiments and/or orthogonal experimental approaches are needed to verify the RNA structural findings.

The RNA structural information from enzymatic and chemical probing approaches are detected via direct RNA scission or indirect reverse transcription, and the probing results obtained are often incorporated into thermodynamic-based RNA structure prediction programs, such as RNAstructure [49], Vienna [50], and Mfold [51], as experimental constraints (usually via a scale of probe reactivity values) to improve the overall RNA structure prediction. In addition, phylogenetic analysis can be performed to improve and validate the proposed structure model if orthologous RNA sequences are available from other species. Different computational approaches have been developed to analyze and incorporate the probing data for RNA structure prediction, and detailed discussion on this topic can be found in recent outstanding reviews [52–55]. Many computational programs, such as SeqFold [56], SAVoR [57], StructureFold [58], ShapeMapper/SuperFold [59], and RNA structure framework [60], were also developed to handle large structure probing dataset from RNA structurome studies.

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

Table 1 Chemical probes commonly used for RNA structure probing

Chemical	Chemical structure	Specificity	Commercially available?*	References
DMS		N_1A N_3C N_7G	Yes. CAS No. 77-78-1	[27,33,34]
CMCT		N_3U N_1G	Yes. CAS No. 2491-17-0	[30,119]
Kethoxal		N_1G, $N_{(2)}G^{\dagger}$	Yes. CAS No. 27762-78-3	[30,120]
Hydroxyl radical	•OH	C_4'-H	No. See [42,45] for synthesis	[41,44]
NMIA		C_2'-OH	Yes. CAS No. 10328-92-4	[35,121]
1M7		C_2'-OH	Yes. CAS No. 73043-80-8	[122,123]
BzCN		C_2'-OH	Yes. CAS No. 613-90-1	[38,124]
1M6		C_2'-OH	Yes. CAS No. 4693-01-0	[125,126]
NAI		C_2'-OH	Yes. CAS No. 1055970-47-2	[27,40]
FAI		C_2'-OH	Yes. CAS No. 1415238-77-5	[40]
NAI-N3		C_2'-OH	Yes. CAS No. 1612756-29-2	[70]
NPIA		C_2'-OH	Yes. CAS No. 57384-79-9	[127]

DMS, dimethyl sulfate; CMCT, 1-cyclohexyl-(2-morpholinoethyl)carbodiimide metho-*p*-toluene sulfonate; Kethoxal, 3-ethoxy-1,1-dihydroxy-2-butanone; NMIA, *N*-methylisatoic anhydride; 1M7, 1-methyl-7-nitroisatoic anhydride; BzCN, benzoyl cyanide; 1M6, 1-methyl-6-nitroisatoic anhydride; NAI, 2-methylnicotinic acid imidazolide; FAI, 2-methyl-3-furoic acid imidazolide; NAI-N3, 2-methylnicotinic acid imidazolide-azide; NPIA, *N*-propanone isatoic anhydride.*SciFinder was used for the searching. Chemical Abstracts Service numbers (CAS No.) are provided.
†Exocylic amine.

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

Insights from RNA structurome studies

The advent of NGS has enabled the analysis of RNA structure probing data in a high-throughput fashion, in contrast with the lower-throughput electrophoresis-based methods either in a slab gel or a capillary format. The first *in vitro* transcriptome-wide enzymatic probing-sequencing method was the parallel analysis of RNA structures (PARS), initially applied in yeast [61]. Other enzymatic probing-sequencing methods are listed in Table 2 and their comparison has been reviewed elsewhere [7]. PARS uses RNase V1 and RNase S1 to probe for dsRNA and ssRNA regions, respectively, in separate reactions, and the structural information are read out by NGS. Wan et al. [62] later applied PARS to examine the RNA thermostability at a range of temperatures and interestingly found that yeast mRNAs were generally less thermostable (lower melting temperature, T_m) than non-coding RNAs (ncRNAs), such as rRNA and tRNAs. The ability to determine T_m per nucleotide using PARS offers an attractive approach to discover RNA thermometers in a transcriptome-wide fashion in any organism. Recently, Righetti et al. [63] applied PARS to obtain the *in vitro* RNA structurome of *Yersinia pseudotuberculosis*, a foodborne pathogen, at three physiologically relevant temperatures (25, 37, and 42°C), and identified 16 RNA thermometers that are able to regulate gene expression in a temperature-dependent fashion.

PARS has also been applied to study the effect of single-nucleotide polymorphism (SNP) on RNA structures, and thousands of riboSNitches (SNP-mediated RNA structure switch) were reported in an *in vitro* study using healthy human parent–offspring trio samples [64]. These riboSNitches were found to be depleted in 3′-untranslated regions (3′UTRs), miRNA-binding sites, and protein-binding sites [64]. This work provides a useful dataset for evaluation and the development of computational riboSNitch prediction programs [65], and promising candidates for *in vivo* verification and further characterization to establish a link between SNP, RNA structure, and gene regulation. Overall, enzymatic probing-sequencing methods (Table 2) uncovered unique insights that could not be obtained by low-throughput single-transcript studies, and set the stage for the exploration of the *in vivo* RNA structurome using chemical probing-sequencing methods, as discussed in Table 2.

The first few *in vivo* transcriptome-wide chemical probing-sequencing methods reported were structure-sequencing (Structure-seq) (Figure 2) [66], DMS-sequencing (DMS-seq) [67], and modification-sequencing (Mod-seq) (Table 2) [68]. These three methods used DMS to probe the *in vivo* RNA structurome in *Arabidopsis* seedling, yeast, and human cells. The *in vivo* DMS probing data from these studies were shown to map consistently to the phylogenetic or X-ray crystal structures of rRNAs [66–68]. Comparison between these methods has been discussed extensively by Kwok et al. [9] and Aviran et al. [69].

Remarkably, Structure-seq and DMS-seq both revealed that many mRNAs were less structured and/or more dynamic *in vivo* versus *in vitro*/*in silico* [66,67]. Interestingly, this suggests that the cellular environment (e.g. proteins and physiological stress) significantly contributes to the RNA structure that predominates *in vivo*. The finding that mRNAs are less structured *in vivo* was also observed later in mouse cells by *in vivo* click selective 2′-hydroxyl acylation and profiling experiment (icSHAPE)(Table 2) [70], a transcriptome-wide method that uses a novel SHAPE reagent called NAI-N3 (Table 1) to probe RNA structures (Figure 2). The mechanistic basis of this observation is largely unknown, although it was speculated to be partly due to RNA helicases' unwinding function *in vivo* [67]. Future experiments may shed light into how RNA helicases co-ordinate with other proteins (e.g. annealers and chaperones) and/or other cellular factors to control *in vivo* mRNA structure unfolding and remodeling, and how these events may couple to translation, mRNA decay, mRNA localization, and other biological processes.

Using icSHAPE, Spitale et al. [70] identified that upon native RNA methylation at N_6A position (m^6A), the common sequence motif GGm^6ACU is less structured compared with unmodified m^6A sites, suggesting that the presence of m^6A weakens the stem region of RNA *in vivo*. To support this, icSHAPE in Mettl3 (a key m^6A methyltransferase) knockout cells was performed and the results showed that the region at, or around m^6A, sites becomes more structured [70]. This interesting finding was corroborated by detailed biophysical and biochemical characterizations [71–73]. It will be exciting to see more examples of functional m^6A structural switches *in vivo*, and to explore other native RNA modification structural switches given the abundance and diversity of RNA modified bases being discovered in various organisms.

By comparing the *in vivo* and *in vitro* icSHAPE reactivities, a parameter called *in vivo–in vitro* difference (VTD) was introduced to identify regions where cellular factors may have caused RNA structural rearrangements [70]. Notably, VTD analysis has enabled the identification of distinct differential icSHAPE pattern in Rbfox and HuR protein-binding sites in RNA. This result suggests that in addition to consensus motif, RNA

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

Table 2 RNA structure-sequencing methods developed to date

Methods	Application to date	Probes used	References
Enzymatic probing-sequencing			
PARS	*In vitro*	RNase V1 (dsRNA), RNase S1 (ssRNA) or RNase A/T1 (ssU/C/G)	[61–64,128,129]
ds/ssRNA-seq	*In vitro*	RNase V1, RNase I (ssRNA)	[130–132]
Frag-seq	*In vitro*	RNase P1 (ssRNA)	[133]
Chemical probing-sequencing			
SHAPE-seq	*In vitro*	1M7	[134,135]
MAP-seq	*In vitro*	1M7, DMS, CMCT	[136]
Structure-seq	*In vivo*	DMS	[66,117]
DMS-seq	*In vitro*, *in vivo*	DMS	[67]
HRF-seq	*In vitro*	Hydroxyl radical	[87]
Mod-seq	*In vivo*	DMS	[68,137]
SHAPE-MaP	*In vitro*, *in vivo*	1M7, 1M6, NMIA	[59,74,93,138,139]
ChemMod-seq	*In vitro*	DMS, 1M7	[140]
RING-MaP	*In vitro*	DMS	[141]
SHAPE-seq 2.0	*In vitro*	1M7	[142,143]
CIRS-seq	*In vitro*	DMS, CMCT	[144]
icSHAPE	*In vitro*, *in vivo*	NAI-N3	[70,118]
SHAPES	*In vitro*	NPIA	[127]
MOHCA-seq	*In vitro*	Hydroxyl radical	[88]
In cell SHAPE-seq	*In vitro*, *in vivo*	1M7, DMS	[143,145]
Targeted Structure-seq	*In vivo*	DMS	[110]
RNA hybrid ligation-sequencing			
CLASH	*In vivo*	UV cross-linking and proximity ligation	[79,80,89]
HiCLIP	*In vivo*	UV cross-linking and proximity ligation	[81]
RPL	*In vitro*, *in vivo*	Proximity ligation	[91]
PARIS	*In vivo*	AMT cross-linking and promixity ligation	[82]
SPLASH	*In vivo*	Psoralen-Biotin cross-linking and promixity ligation	[83]
LiGR-seq	*In vivo*	AMT cross-linking and promixity ligation	[84]

Working pipeline of the italicized methods is described in Figures 2 and 3.

local structural variations, as identified by *in vivo* and *in vitro* icSHAPE comparison, can provide additional information to predict authentic protein-binding sites in RNA. Independently, another SHAPE-based method called SHAPE-MaP (SHAPE-mutational profiling) (Figure 2 and Table 2) presented a similar approach called ΔSHAPE (*in cellulo ex vivo* difference) to detect protein-binding sites in many abundant ncRNAs, such as 5S rRNA, U1 snRNA, and RNase MRP, in mouse cells [74]. It is uncertain whether the VTD and ΔSHAPE approaches are applicable to non-sequence-specific dsRNA-binding proteins, which may only cause a minimal variation in local RNA structural environment between *in vivo* and *in vitro* conditions, i.e. VTD and ΔSHAPE ~0. Nonetheless, these RNA-centric strategies to identify protein-binding sites in RNA are powerful and are complementary to protein-centric methods such as individual cross-linking immunoprecipitation (iCLIP) [75], enhanced-CLIP [76], infrared-CLIP [77], and cross-linking and analysis of cDNAs (CRAC) [78] that identifies RNA-binding sites of protein.

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

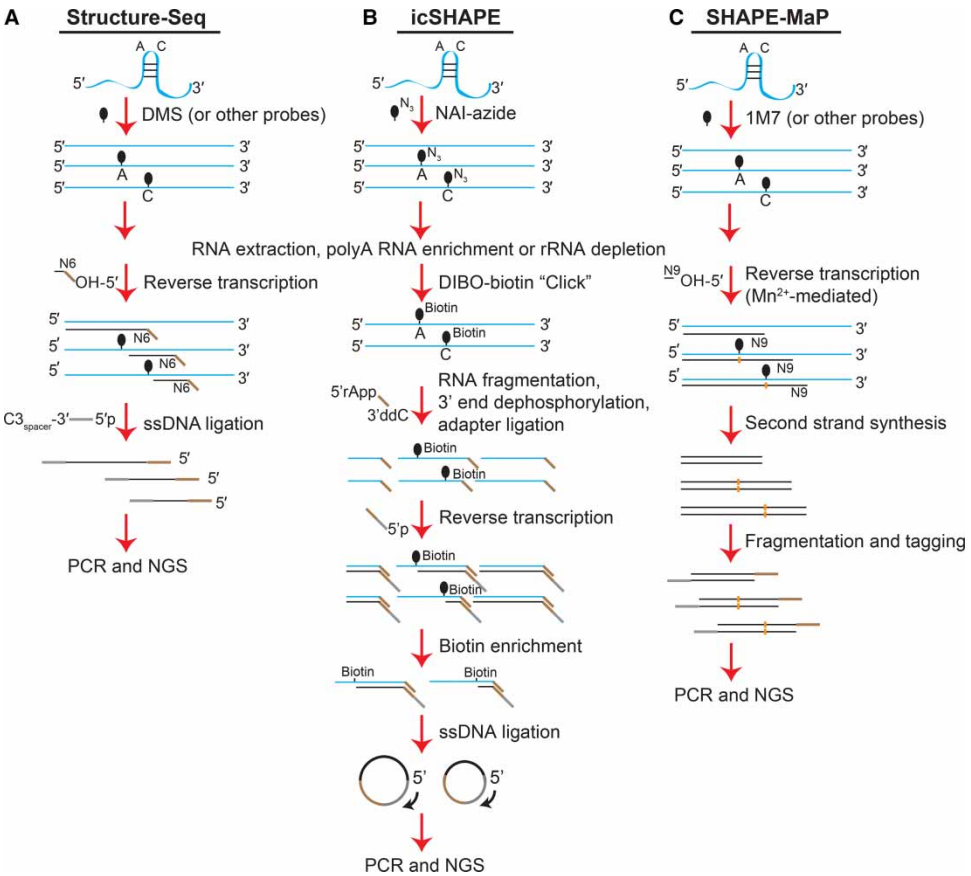


Figure 2. Experimental pipeline for Structure-seq, icSHAPE, and SHAPE-MaP.

The major steps are shown. The chemical modification is marked with black oval, which can be detected by chemical-induced reverse transcriptase halting (for Structure-seq and icSHAPE) or chemical-induced reverse transcriptase read-through (for SHAPE-MaP). (**A**) For Structure-seq [117], RNA (blue) is treated with DMS (or other probes such as NAI) *in vivo*. After RNA extraction and polyA enrichment/rRNA depletion, adapter-fused (brown) random hexamer (N6) reverse transcription (RT) is performed to generate the cDNA (black). Then, the cDNA is ligated to an ss DNA (gray), which contains part of the adapter sequence for PCR purpose. The ligated cDNA is subjected to PCR amplification, and submitted for NGS. (**B**) For icSHAPE [118], RNA is treated with NAI-azide (N3) *in vivo*. After RNA extraction and polyA enrichment/rRNA depletion, click chemistry is performed between RNA-NAI-N_3 and dibenzocyclooctynes (DIBO)-biotin. Next, RNA is subjected to random fragmentation by Zn^{2+}-mediated hydrolysis, followed by 3′-dephosphorylation, and ligation with a pre-adenylated adapter (brown). An adapter-specific RT is then performed to generate cDNA (black). The biotinylated RNA–cDNA hybrid is then enriched using streptavidin magnetic beads. The enriched cDNA is eluted and followed by ssDNA ligation (circularization), and subsequently PCR and NGS. (**C**) For SHAPE-MaP [59], RNA is treated with 1M7 (or other probes such as DMS) *in vivo*. After RNA extraction and polyA enrichment/rRNA depletion, random nonamer (N9) RT is performed to generate first-strand cDNA. Manganese ion (Mn^{2+}) is present in the RT reaction to facilitate reverse transcriptase to read-through the modified site by incorporating noncomplementary bases/mutation (orange). Next, second-strand synthesis, fragmentation and tagging are conducted based on the Nextera library preparation protocol, and followed by PCR and NGS.

Insights from RNA interactome studies

From the analysis of CRAC data, Tollervey and colleagues discovered chimeric RNAs [small nucleolar RNA (snoRNA)–rRNA hybrids], which occurred at low frequency, from the RNA ligation step in the CRAC protocol [79]. Driven by this observation, they optimized the experimental and computational protocol and developed CLASH [79] (UV cross-linking, ligation, and sequencing of hybrids) (Table 2), to enable the direct detection of RNA–RNA hybrids that are bound by protein of interest. CLASH was applied to study intermolecular interactions of snoRNA–rRNA in yeast [79] and miRNA-target RNAs in human cells [80], adding another dimension

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

of RNA structural information. Similarly, Sugimoto *et al.* modified the iCLIP protocol to develop hybrid iCLIP (hiCLIP) (Figure 3 and Table 2) [81], which was applied to detect dsRNA bound by STAU1 in human cells. There are two key differences between CLASH and hiCLIP. First, CLASH utilized tagged protein of interest in the cells for more stringent purification. Second, hiCLIP introduced an adapter between the two RNA fragments to allow more efficient ligation and better identification of the two RNA fragments in analysis (Figure 3A, orange). Using hiCLIP [81], many new intramolecular long-range RNA base-pairings, which were missed or generally not being considered in RNA structure prediction programs, were observed, especially within 3′-UTRs [81]. For example, hiCLIP detected an RNA helix spanning 858 nucleotides in the XBP1 3′-UTR. Further duplex mutant and rescue experiments verified this long-range interaction, and showed that the duplex formation affects the mRNA stability [81].

Recently, three psoralen-based transcriptome-wide structure probing methods, including psoralen analysis of RNA interactions and structures (PARIS) [82], sequencing of psoralen cross-linked, ligated, and selected hybrids (SPLASH) [83], and LIGation of interacting RNA followed by high-throughput sequencing (LIGR-seq) (Figure 3 and Table 2) [84], have been developed to reveal RNA interactomes in human cells, mouse cells, and yeast. Psoralen and its derivatives, such as 4′-aminomethyltrioxsalen (AMT) or biotinylated psoralen, intercalate in RNA helices and undergo interstranded cross-link upon UV 365 nm irradiation (Figure 3). Using PARIS (Figure 3) [82], Lu *et al.* reveal prevalent long-range intramolecular RNA interactions, alternative structures, and intermolecular RNA interactions in human and mouse cells. Combining PARIS, icSHAPE, and phylogenetic analysis, they have determined the structure of XIST lncRNA and found that it was organized in four major domains, of which three of them were conserved. They also showed by iCLIP that SPEN, a key epigenetic silencing protein, interacted with the A-repeat of XIST RNA, highlighting the higher-order structure of XIST lncRNA in living cells.

Similarly, SPLASH and LIGR-seq reveal diverse long-range intramolecular and intermolecular RNA interactions in human cells and yeast. Using SPLASH (Figure 3) [83], Aw *et al.* found that efficiently translated mRNAs have long-range interactions that usually connected the beginning and end of transcripts, whereas less efficiently translated mRNAs have short-range interactions that are within the 5′-UTR. This result suggests that stable structures within 5′-UTRs will inhibit translation, whereas end-to-end interactions facilitate translation, possibly via ribosome recycling. Using LiGR-seq [84], Blencowe and colleagues identified previously unknown interactions between snoRNAs and mRNAs. In particular, they showed that the orphan C/D box snoRNA, SNORD83B, regulates the steady-state level of the mRNAs that it interacts with, such as NOP14, RPS5, and SRSF3. Taken together, these novel RNA structurome and interactome methods allows original hypotheses to be tested, and opens new doors for us to explore the wonders of RNA in biological system.

Current challenges and potential future directions

At present, many chemical probing-sequencing methods utilize either nucleobase-specific probes (e.g. DMS or CMCT) or ribose-specific probes (e.g. 1M7 or hydroxyl radical) (Tables 1 and 2). Given that they yield complementary structural information as discussed above, it is desirable to use two or more probes in parallel to obtain a more comprehensive view of the *in vivo* RNA structurome. In particular, hydroxyl radicals can provide tertiary structure information of RNA that can complement the secondary structure information provided by DMS or SHAPE [85,86]. Proof-of-concept *in vitro* studies using hydroxyl radical footprinting-sequencing (HRF-seq) [87] and massive ·OH cleavage analysis-sequencing (MOHCA-seq) [88] (Table 2) showed that the probing results were consistent with the structures of the 16S rRNA, as well as of several ncRNAs. It will be exciting to see *in vivo* transcriptome-wide hydroxyl radical probing [28,29] to be developed in the near future.

One major limitation of current chemical probes (Table 1) is their inability to directly probe dsRNA regions like RNase V1 does, making RNA structure predictions, especially for long RNAs (>200 nt), still challenging. The RNA hybrid ligation-sequencing methods (Table 2), such as CLASH [79,80,89] and HiCLIP [81], can be applied to complement the chemical probing-sequencing methods; however, the low UV cross-linking efficiency [90], as well as inefficient RNA–RNA ligation, poses some technical challenges to obtain decent amount of RNA–RNA contact information (< 2% of the data) [81,89]. In addition, the requirement of protein binding limits their use to reveal RNA–RNA hybrids (either intramolecular or intermolecular) that are not bound by a protein. RNA proximity ligation (RPL) [91] alleviates the requirement of RBP, yet the lack of enrichment step produces high background noise and only yields <0.3% RNA–RNA contact information in the data. PARIS, SPLASH, and LIGR-seq were recently developed to overcome these issues by

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

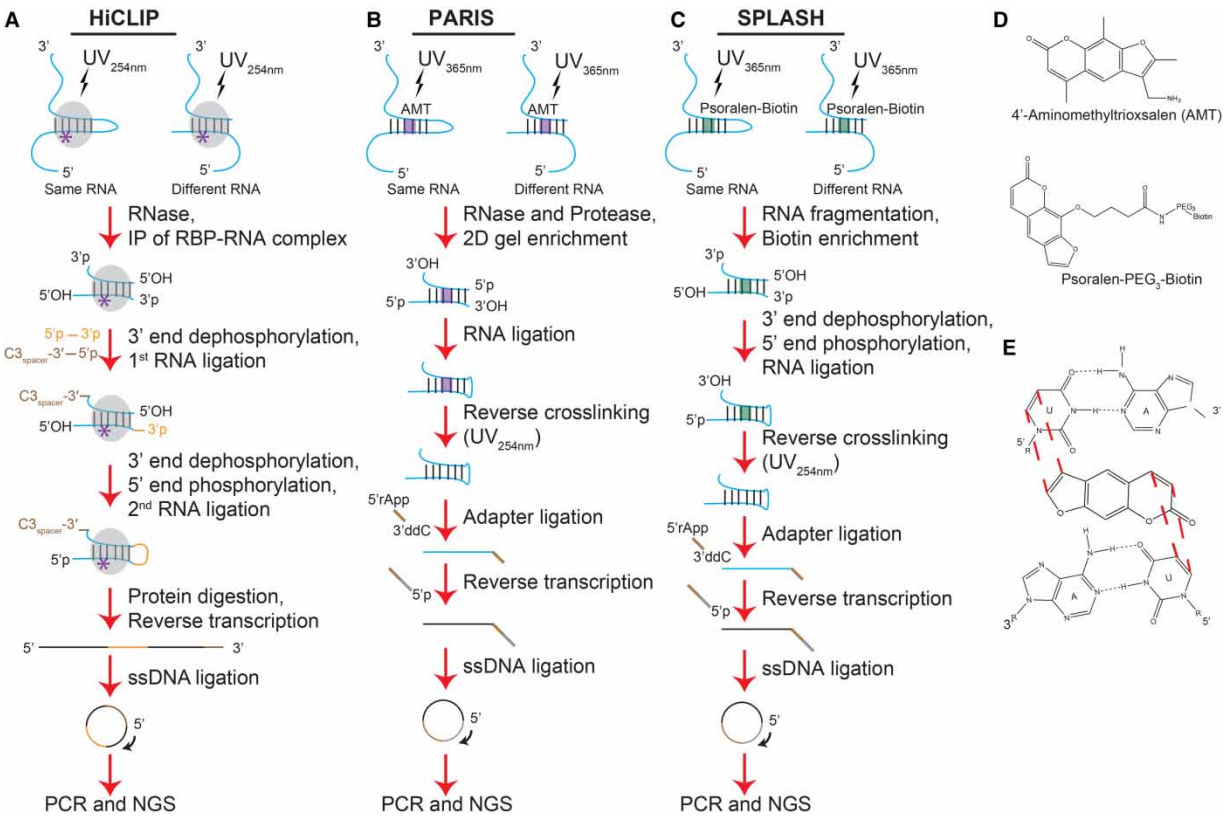


Figure 3. Experimental pipeline for hiCLIP, PARIS, and SPLASH.

The major steps are shown. (**A**) For hiCLIP [81], RNAs (blue) that are bound by protein of interest (gray oval) are UV_{254nm}-cross-linked (purple asterisk). Cross-linked RNAs are then trimmed by RNase and then enriched by immunoprecipitation (IP). The enriched RNA is 3′-dephosphorylation, followed by first RNA ligation (orange and brown). The ligated RNA is 3′-dephosphorylated and 5′-phosphorylated before second RNA ligation (proximity ligation of the orange). Next, the protein is digested by protease, followed by RT to generate cDNA (black). The cDNA is circularized, and subjected to PCR and NGS. (**B**) For PARIS [82], RNAs are UV_{365nm}-cross-linked (purple) with the psoralen derivative, AMT. Cross-linked RNAs are then RNase and protease-treated, and enriched by 2D gel electrophoresis. The enriched RNA is ligated (proximity ligation), followed by reverse UV_{254nm}-cross-linking. Next, ligation is performed with a pre-adenylated adapter (brown). An adapter-specific RT is then performed to generate cDNA (black). The cDNA is circularized, and subjected to PCR and NGS. (**C**) For SPLASH [83], RNAs are UV_{365nm}-cross-linked (green) with biotinylated psoralen. Next, cross-linked RNA is subjected to random fragmentation by Zn^{2+}-mediated hydrolysis, followed by biotin enrichment using streptavidin magnetic beads. The enriched, cross-linked RNAs are then 3′-dephosphorylated and 5′-phosphorylated before RNA ligation (proximity ligation). Reverse UV_{254nm}-cross-linking is then performed. Next, ligation is performed with a pre-adenylated adapter (brown). An adapter-specific RT is then performed to generate cDNA (black). The cDNA is circularized, and subjected to PCR and NGS. (**D**) Chemical structure of AMT and biotinylated psoralen. (**E**) Reaction interface of psoralen and UA base pairs. Psoralen (or its derivatives) intercalates into RNA helices (e.g. UA base pairs) and reacts with Us to generate interstrand cross-links upon 365 nm UV irradiation.

utilizing chemical cross-linking reagents (psoralen or its derivatives), which circumvent the need of an RNA-binding protein for cross-linking. Enrichment steps were designed in each method to capture the cross-linked RNA hybrids (Figure 3), allowing more robust identification of the much-needed RNA base-pairing information. Caution is needed for the UV treatment steps, especially the reverse cross-link step using UV 254 nm, as they can damage RNA [92]. In addition, the proximity ligation yield is generally low and so requires further optimization. As psoralen and its derivatives prefer cross-linking at pyrimidines, especially at Us (Figure 3), future development may focus on utilizing or designing new chemical cross-linking which that can complement this psoralen-based approach.

RNA often folds into an architecture that contains complex structural motifs, such as RNA pseudoknots and G-quadruplexes, which have traditionally been difficult to predict and detect. Recent studies have shown some progress in the detection and validation of several new pseudoknots in the HIV-1 RNA genome *in vitro* using

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

SHAPE-MaP [93]. Also, PARIS was able to detect the well-studied pseudoknot in the human telomerase RNA *in vivo* [82]. It is anticipated that these or other approaches may be applied more broadly to discover functional RNA pseudoknots in other organisms and help improve the pseudoknot structure prediction. Another challenging structural motif, the RNA G-quadruplex, has conventionally been predicted using sequence motif search algorithms [94,95], which do not generally consider the effect of the flanking sequence on RNA G-quadruplex formation. RNA G-quadruplexes have been recently shown to exist in human cells [96,97] and be biologically significant [98–100]; however, the limited experimental data on RNA G-quadruplexes have restricted the advancement of structure-based prediction algorithms for RNA G-quadruplexes [101], as well as our understanding of their roles in living cells. Current experimental methods to identify and characterize RNA G-quadruplexes are low-throughput [102–107], and thus great efforts are needed to develop novel methods to discover new and important RNA G-quadruplex on a transcriptome-wide scale. Recently, Kwok et al. [108] has developed the first *in vitro* transcriptome-wide RNA G-quadruplex profiling method, rG4-seq, and revealed widespread formation of RNA G-quadruplexes in the human transcriptome.

RNA chemical probing is performed under conditions of single-hit kinetics (about one modification every 200 nt) as over-modification can affect the propensity of base-pairing and probably induce changes in the native RNA structure. Under these criteria, only a small proportion of the sequencing data (10–20%) reports on modified RNAs that contain useful structural information (the rest are from unmodified RNAs), resulting in low signal-to-noise ratio. One straightforward way to resolve this issue is to sequence deeper, although the associated cost will be high. Alternative approaches have been reported to alleviate this issue. Mod-seq [68] introduces adapter selection and subtraction steps in the protocol to enrich for modified RNA pools, whereas icSHAPE [70] employs a novel bifunctional SHAPE reagent, NAI-N3, which is compatible with performing click chemistry [109] with a biotin-alkyne molecule after SHAPE treatment. The clicked, modified RNAs are subsequently enriched by streptavidin pull-down. While these strategies allow the enrichment of modified RNA and thus increase the signal-to-noise ratio, they come at a cost of requiring more handling steps at the RNA level when compared with Structure-seq and SHAPE-MaP (Figure 2). The additional handling might lead to undesired RNA degradation, as well as loss of precious RNA sample, making *in vivo* RNA structural study on rare RNAs and on smaller number of cells challenging. Future development may focus on simplifying the library preparation workflow and meanwhile optimizing the coverage of useful sequencing data that contain structural information.

The dynamic range of RNA expression level is vast. Similar to RNA-seq, low abundance transcripts are difficult to detect in any of the RNA structure-sequencing methods (Table 2), as the majority of sequencing reads will contribute to high or medium abundance transcripts. One way to study the structure of low abundance transcript of interest is to employ targeted RNA structure probing approach such as DMS/SHAPE-LMPCR [27]. Recently, targeted Structure-seq [110], which couples *in vivo* DMS and transcript-specific reverse transcription with sequencing-based read out, was developed and applied to probe the structure of low abundance, long ncRNA *in vivo* [110]. This method provides a useful tool for studying low abundance RNA in great coverage depth without RNA length limitation. Another promising approach, CaptureSeq [111], utilizes oligonucleotides to capture selected transcripts or regions of interest for targeted sequencing. This selective enrichment strategy applies robustly to RNA-seq experiment to detect rare transcripts and isoforms [112], and may be applicable for RNA structurome studies as well. One can also consider overexpressing these low abundance RNAs in cells to increase the RNA's copy number (concentration) for easier structural investigation and detection; however, caution is needed for this approach as solely overexpressing the low abundance RNA of interest may affect the stoichiometry between the RNA and its interaction partners inside cells which potentially affects the cell environment.

Most RNA structurome methods measure an ensemble of RNA structures that are averaged over the reaction timeframe, as such we so far can only reveal and appreciate consistently strong and predominant structural features that persist over this reaction window (for other reviews, see refs [7–12]). Providing quick RNA structural snapshots at specific time intervals in a controlled manner will aid in deciphering the folding pathways of RNAs, such as RNP assembly and co-transcription folding, in a time-resolved manner. As RNA is dynamic in nature, they can exist in multiple structural conformations in cells and interconvert in response to change in cellular environment. Discerning the functional RNA folds from the nonfunctional RNA folds is a big challenge for both experimentalists and bioinformaticians. Integrative analysis of the results from enzymatic/chemical probing-sequencing (Table 2), RNA hybrid ligation-sequencing (Table 2), and phylogenetic analysis may help to distinguish and validate the functional RNA fold. In addition, targeted approaches that are based on

random mutagenesis of RNA target followed by functional selection and NGS [113–115] have been developed to identify primary, secondary, and tertiary structures of RNA that are crucial for its function. Potentially, selected RNA target from the structurome and interactome studies can be subjected to these targeted approaches for in-depth variation and co-variation analysis to dissect different structural conformations in RNA. With this in mind, developing novel targeted and global RNA structure-sequencing methods and robust computational algorithms to deconvolute ensemble RNA structures and interactions into their distinct conformational state, for example, by cluster analysis [54,55,116], will undoubtedly allow us to examine the RNA structures, dynamics, and functions in many cellular processes, such as translation, RNA localization, and RNA decay, in unprecedented level of details.

Conclusions

The emergence of RNA structurome and interactome research, thanks to the marriage of RNA structure probing and NGS, has revolutionized the field by stimulating new ways to study RNA structure and interaction on a transcriptome-wide scale. We are entering an exciting phase of establishing a more complete toolset to decipher the *in vivo* RNA structurome and interactome. Application of these RNA structure-sequencing methods (Table 2) in different experimental settings will enable us to understand the way RNA, an ancient biomolecule, can be a sensor to perceive the physiological stimuli (Figure 1), and at the same time an effector (either directly or indirectly) to regulate diverse biological processes including gene expression. Accumulating evidence suggests that RNAs have potential links with the development of diseases; therefore, revealing the structure and interaction of RNAs may help elucidating the molecular basis of diseases and facilitate the development of therapeutic and diagnostic tools. We look forward to the sunrise of the *in vivo* RNA structurome and interactome with great anticipation.

Abbreviations

1M7, 1-methyl-7-nitroisatoic anhydride; 3′-UTRs 3′-untranslated regions; AMT, 4′-aminomethyltrioxsalen; BzCN, benzoyl cyanide; CMCT, 1-cyclohexyl-(2-morpholinoethyl)carbodiimide metho-*p*-toluene sulfonate; CRAC, cross-linking and analysis of cDNAs; DIBO, dibenzocyclooctynes; DMS, dimethyl sulfate; DMS-seq, DMS-sequencing; ds, double-stranded; hiCLIP, hybrid iCLIP; HRF-seq, hydroxyl radical footprinting-sequencing; iCLIP, individual cross-linking immunoprecipitation; LIGR-seq, LIGation of interacting RNA followed by high-throughput sequencing; lncRNAs, long non-coding RNAs; Mod-seq, modification-sequencing; MOHCA-seq, massive ·OH cleavage analysis-sequencing; NAI, 2-methylnicotinic acid imidazolide; ncRNAs, non-coding RNAs; NGS, next-generation sequencing; NMIA, *N*-methylisatoic anhydride; PARIS, psoralen analysis of RNA interactions and structures; PARS, parallel analysis of RNA structures; RNases, ribonucleases; RPL, RNA proximity ligation; RNP, ribonucleoprotein; rRNAs, ribosomal RNAs; RT, reverse transcription; SHAPE, selective 2′-hydroxyl acylation analyzed by primer extension; SHAPE-MaP, SHAPE-mutational profiling; snoRNAs, small nucleolar RNAs; SNP, single-nucleotide polymorphism; SPLASH, sequencing of psoralen cross-linked, ligated, and selected hybrids; ss, single-stranded; Structure-seq, structure-sequencing. tRNAs, transfer RNAs; rRNAs, ribosomal RNAs; miRNAs, microRNAs.

Funding

This work is supported by an Advanced Grant from the European Research Council (Grant No. [339778]). C.K.K. received some support from the Croucher Foundation.

Acknowledgements

I thank Prof. Shankar Balasubramanian, Prof. Philip C. Bevilacqua, Ms Vicki S. Chambers, and Dr Omer Ziv for critical reading and comments of the manuscript. I apologize to colleagues whose works were not cited due to space limitations.

Competing Interests

The Author declares that there are no competing interests associated with the manuscript.

References

1 Elliott, D. and Ladomery, M. (2015) Molecular Biology of RNA, Oxford University Press
2 Sharp, P.A. (2009) The centrality of RNA. *Cell* **136**, 577–580 doi:10.1016/j.cell.2009.02.007

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

3 Cech, T.R. and Steitz, J.A. (2014) The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **157**, 77–94 doi:10.1016/j.cell.2014.03.008
4 Morris, K.V. and Mattick, J.S. (2014) The rise of regulatory RNA. *Nat. Rev. Genet.* **15**, 423–437 doi:10.1038/nrg3722
5 Serganov, A. and Patel, D.J. (2007) Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet.* **8**, 776–790 doi:10.1038/nrg2172
6 Hendrix, D.K., Brenner, S.E. and Holbrook, S.R. (2005) RNA structural motifs: building blocks of a modular biomolecule. *Q. Rev. Biophys.* **38**, 221–243 doi:10.1017/S0033583506004215
7 Wan, Y., Kertesz, M., Spitale, R.C., Segal, E. and Chang, H.Y. (2011) Understanding the transcriptome through RNA structure. *Nat. Rev. Genet.* **12**, 641–655 doi:10.1038/nrg3049
8 Mortimer, S.A., Kidwell, M.A. and Doudna, J.A. (2014) Insights into RNA structure and function from genome-wide studies. *Nat. Rev. Genet.* **15**, 469–479 doi:10.1038/nrg3681
9 Kwok, C.K., Tang, Y., Assmann, S.M. and Bevilacqua, P.C. (2015) The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem. Sci.* **40**, 221–232 doi:10.1016/j.tibs.2015.02.005
10 Kubota, M., Tran, C. and Spitale, R.C. (2015) Progress and challenges for chemical probing of RNA structure inside living cells. *Nat. Chem. Biol.* **11**, 933–941 doi:10.1038/nchembio.1958
11 Lu, Z. and Chang, H.Y. (2016) Decoding the RNA structurome. *Curr. Opin. Struct. Biol.* **36**, 142–148 doi:10.1016/j.sbi.2016.01.007
12 Vandivier, L.E., Anderson, S.J., Foley, S.W. and Gregory, B.D. (2016) The conservation and function of RNA secondary structure in plants. *Annu. Rev. Plant Biol.* **67**, 463–488 doi:10.1146/annurev-arplant-043015-111754
13 Gutell, R.R. (2014) Ten lessons with Carl Woese about RNA and comparative analysis. *RNA Biol.* **11**, 254–272 doi:10.4161/rna.28718
14 Parsch, J., Braverman, J.M. and Stephan, W. (2000) Comparative sequence analysis and patterns of covariation in RNA secondary structures. *Genetics* **154**, 909–921 PMID: 10655240
15 Gutell, R.R. (1993) Comparative studies of RNA: inferring higher-order structure from patterns of sequence variation. *Curr. Opin. Struct. Biol.* **3**, 313–322 doi:10.1016/S0959-440X(05)80101-0
16 Gutell, R.R., Larsen, N. and Woese, C.R. (1994) Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol. Rev.* **58**, 10–26 PMID: 8177168
17 Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y. et al. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 2 doi:10.1186/1471-2105-3-2
18 Gutell, R.R., Lee, J.C. and Cannone, J.J. (2002) The accuracy of ribosomal RNA comparative structure models. *Curr. Opin. Struct. Biol.* **12**, 301–310 doi:10.1016/S0959-440X(02)00339-1
19 Johnsson, P., Lipovich, L., Grandér, D. and Morris, K.V. (2014) Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim. Biophys. Acta* **1840**, 1063–1071 doi:10.1016/j.bbagen.2013.10.035
20 Lengyel, J., Hnath, E., Storms, M. and Wohlfarth, T. (2014) Towards an integrative structural biology approach: combining cryo-TEM, X-ray crystallography, and NMR. *J. Struct. Funct. Genomics* **15**, 117–124 doi:10.1007/s10969-014-9179-9
21 Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P. and Ehresmann, B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**, 9109–9128 doi:10.1093/nar/15.22.9109
22 Weeks, K.M. (2010) Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.* **20**, 295–304 doi:10.1016/j.sbi.2010.04.001
23 Knapp, G. (1989) Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol.* **180**, 192–212 doi:10.1016/0076-6879(89)80102-8
24 Woese, C.R., Magrum, L.J., Gupta, R., Siegel, R.B., Stahl, D.A., Kop, J. et al. (1980) Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res.* **8**, 2275–2294 doi:10.1093/nar/8.10.2275
25 Aultman, K.S. and Chang, S.H. (1982) Partial P1 nuclease digestion as a probe of tRNA structure. *Eur. J. Biochem.* **124**, 471–476 doi:10.1111/j.1432-1033.1982.tb06617.x
26 Guerrier-Takada, C. and Altman, S. (1984) Structure in solution of M1 RNA, the catalytic subunit of ribonuclease P from *Escherichia coli*. *Biochemistry* **23**, 6327–6334 doi:10.1021/bi00321a006
27 Kwok, C.K., Ding, Y., Tang, Y., Assmann, S.M. and Bevilacqua, P.C. (2013) Determination of in vivo RNA structure in low-abundance transcripts. *Nat. Commun.* **4**, 2971 PMID: 24336128
28 Tyrrell, J., McGinnis, J.L., Weeks, K.M. and Pielak, G.J. (2013) The cellular environment stabilizes adenine riboswitch RNA structure. *Biochemistry* **52**, 8777–8785 doi:10.1021/bi401207q
29 Lavery, R. and Pullman, A. (1984) A new theoretical index of biochemical reactivity combining steric and electrostatic factors. An application to yeast tRNAPhe. *Biophys. Chem.* **19**, 171–181 PMID: 6372881
30 Moazed, D., Stern, S. and Noller, H.F. (1986) Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. *J. Mol. Biol.* **187**, 399–416 doi:10.1016/0022-2836(86)90441-9
31 Brunel, C., Romby, P., Westhof, E., Ehresmann, C. and Ehresmann, B. (1991) Three-dimensional model of *Escherichia coli* ribosomal 5S RNA as deduced from structure probing in solution and computer modeling. *J. Mol. Biol.* **221**, 293–308 doi:10.1016/0022-2836(91)80220-O
32 Wassarman, K.M. and Steitz, J.A. (1992) The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol. Cell. Biol.* **12**, 1276–1285 doi:10.1128/MCB.12.3.1276
33 Zaug, A.J. and Cech, T.R. (1995) Analysis of the structure of tetrahymena nuclear RNAs in vivo: telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. *RNA* **1**, 363–374 PMID: 7493315
34 Wells, S.E., Hughes, J.M.X., Haller Igel, A., Ares, M. (2000) Use of dimethyl sulfate to probe RNA structure in vivo. *Methods Enzymol.* **318**, 479–493 doi:10.1016/S0076-6879(00)18071-1
35 Merino, E.J., Wilkinson, K.A., Coughlan, J.L. and Weeks, K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223–4231 doi:10.1021/ja043822v

36 McGinnis, J.L. Dunkle, J.A., Cate, J.H.D. and Weeks, K.M. (2012) The mechanisms of RNA SHAPE chemistry. *J. Am. Chem. Soc.* **134**, 6617–6624 doi:10.1021/ja2104075
37 Spitale, R.C., Flynn, R.A., Torre, E.A., Kool, E.T. and Chang, H.Y. (2014) RNA structural analysis by evolving SHAPE chemistry. *Wiley Interdiscip. Rev. RNA* **5**, 867–881 doi:10.1002/wrna.1253
38 Mortimer, S.A. and Weeks, K.M. (2009) Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. *Nat. Protoc.* **4**, 1413–1421 doi:10.1038/nprot.2009.126
39 McGinnis, J.L. and Weeks, K.M. (2014) Ribosome RNA assembly intermediates visualized in living cells. *Biochemistry* **53**, 3237–3247 doi:10.1021/bi500198b
40 Spitale, R.C., Crisalli, P., Flynn, R.A., Torre, E.A., Kool, E.T. and Chang, H.Y. (2013) RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* **9**, 18–20 doi:10.1038/nchembio.1131
41 Latham, J.A. and Cech, T.R. (1989) Defining the inside and outside of a catalytic RNA molecule. *Science* **245**, 276–282 doi:10.1126/science.2501870
42 Shcherbakova, I. and Mitra, S. (2009) Hydroxyl-radical footprinting to probe equilibrium changes in RNA tertiary structure. *Methods Enzymol.* **468**, 31–46 doi:10.1016/S0076-6879(09)68002-2
43 Tullius, T.D. and Greenbaum, J.A. (2005) Mapping nucleic acid structure by hydroxyl radical cleavage. *Curr. Opin. Chem. Biol.* **9**, 127–134 doi:10.1016/j.cbpa.2005.02.009
44 Sclavi, B., Chance, M.R., Brenowitz, B. and Woodson, S.A. (1998) RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. *Science* **279**, 1940–1943 doi:10.1126/science.279.5358.1940
45 Adilakshmi, T., Soper, S.F.C. and Woodson, S.A. (2009) Structural analysis of RNA in living cells by in vivo synchrotron X-ray footprinting. *Methods Enzymol.* **468**, 239–258 doi:10.1016/S0076-6879(09)68012-5
46 Clatterbuck Soper, S.F., Dator, R.P., Limbach, P.A. and Woodson, S.A. (2013) In vivo X-ray footprinting of pre-30S ribosomes reveals chaperone-dependent remodeling of late assembly intermediates. *Mol. Cell* **52**, 506–516 doi:10.1016/j.molcel.2013.09.020
47 Zito, K., Hűttenhofer, A. and Pace, N.R. (1993) Lead-catalyzed cleavage of ribonuclease P RNA as a probe for integrity of tertiary structure. *Nucleic Acids Res.* **21**, 5916–5920 doi:10.1093/nar/21.25.5916
48 Lindell, M., Romby, P. and Wagner, E.G.H. (2002) Lead(II) as a probe for investigating RNA structure in vivo. *RNA* **8**, 534–541 doi:10.1017/S1355838201020416
49 Reuter, J.S. and Mathews, D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**, 129 doi:10.1186/1471-2105-11-129
50 Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. et al. (2011) ViennaRNA package 2.0. *Algorithms Mol. Biol.* **6**, 26 doi:10.1186/1748-7188-6-26
51 Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 doi:10.1093/nar/gkg595
52 Eddy, S.R. (2014) Computational analysis of conserved RNA secondary structure in transcriptomes and genomes. *Annu. Rev. Biophys.* **43**, 433–456 doi:10.1146/annurev-biophys-051013-022950
53 Sloma, M.F. and Mathews, D.H. (2015) Improving RNA secondary structure prediction with structure mapping data. *Methods Enzymol.* **553**, 91–114 doi:10.1016/bs.mie.2014.10.053
54 Rogers, E. and Heitsch, C. (2016) New insights from cluster analysis methods for RNA secondary structure prediction. *Wiley Interdiscip. Rev. RNA* **7**, 278–294 doi:10.1002/wrna.1334
55 Lorenz, R., Wolfinger, M.T., Tanzer, A. and Hofacker, I.L. (2016) Predicting RNA secondary structures from sequence and probing data. *Methods* **103**, 86–98 doi:10.1016/j.ymeth.2016.04.004
56 Ouyang, Z., Snyder, M.P. and Chang, H.Y. (2013) Seqfold: genome-scale reconstruction of RNA secondary structure integrating high-throughput sequencing data. *Genome Res.* **23**, 377–387 doi:10.1101/gr.138545.112
57 Li, F., Ryvkin, P., Childress, D.M., Valladares, O., Gregory, B.D., Wang, L.-S. et al. (2012) SAVor: a server for sequencing annotation and visualization of RNA structures. *Nucleic Acids Res.* **40**, W59–W64 doi:10.1093/nar/gks310
58 Tang, Y., Bouvier, E., Kwok, C.K., Ding, Y., Nekrutenko, A., Bevilacqua, P.C. et al. (2015) Structurefold: genome-wide RNA secondary structure mapping and reconstruction in vivo. *Bioinformatics* **31**, 2668–2675 doi:10.1093/bioinformatics/btv213
59 Smola, M.J., Rice, G.M., Busan, S., Siegfried, N.A. and Weeks, K.M. (2015) Selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.* **10**, 1643–1669 doi:10.1038/nprot.2015.103
60 Incarnato, D., Neri, F., Anselmi, F. and Oliviero, S. (2016) RNA structure framework: automated transcriptome-wide reconstruction of RNA secondary structures from high-throughput structure probing data. *Bioinformatics* **32**, 459–461 doi:10.1093/bioinformatics/btv571
61 Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y. et al. (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467**, 103–107 doi:10.1038/nature09322
62 Wan, Y., Qu, K., Ouyang, Z., Kertesz, M., Li, J., Tibshirani, R. et al. (2012) Genome-wide measurement of RNA folding energies. *Mol. Cell* **48**, 169–181 doi:10.1016/j.molcel.2012.08.008
63 Righetti, F., Nuss, A.M., Twittenhoff, C., Beele, S., Urban, K., Will, S. et al. (2016) Temperature-responsive in vitro RNA structurome of *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* **113**, 7237–7242 doi:10.1073/pnas.1523004113
64 Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z. et al. (2014) Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**, 706–709 doi:10.1038/nature12946
65 Corley, M., Solem, A., Qu, K., Chang, H.Y. and Laederach, A. (2015) Detecting riboSNitches with RNA folding algorithms: a genome-wide benchmark. *Nucleic Acids Res.* **43**, 1859–1868 doi:10.1093/nar/gkv010
66 Ding, Y., Tang, Y., Kwok, C.K., Zhang, Y., Bevilacqua, P.C., Assmann, S.M. et al. (2014) In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**, 696–700 doi:10.1038/nature12756
67 Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. and Weissman, J.S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **505**, 701–705 doi:10.1038/nature12894
68 Talkish, J., May, G., Lin, Y., Woolford, J.L. and McManus, C.J. (2014) Mod-seq: high-throughput sequencing for chemical probing of RNA structure. *RNA* **20**, 713–720 doi:10.1261/rna.042218.113

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

69 Aviran, S. and Pachter, L. (2014) Rational experiment design for sequencing-based RNA structure mapping. *RNA* **20**, 1864–1877 doi:10.1261/rna.043844.113

70 Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.-W. et al. (2015) Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **519**, 486–490 doi:10.1038/nature14263

71 Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., Pan, T. et al. (2015) N(6)- Methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. *Nature* **518**, 560–564 doi:10.1038/nature14234

72 Roost, C., Lynch, S.R., Batista, P.J., Qu, K., Chang, H.Y., Kool, E.T. et al. (2015) Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. *J. Am. Chem. Soc.* **137**, 2107–2115 doi:10.1021/ja513080v

73 Zhou, K.I., Parisien, M., Dai, Q., Liu, N., Diatchenko, L., Sachleben, J.R. et al. (2016) N(6)-Methyladenosine modification in a long noncoding RNA hairpin predisposes its conformation to protein binding. *J. Mol. Biol.* **428**, 822–833 doi:10.1016/j.jmb.2015.08.021

74 Smola, M.J., Calabrese, J.M. and Weeks, K.M. (2015) Detection of RNA–protein interactions in living cells with SHAPE. *Biochemistry* **54**, 6867–6875 doi:10.1021/acs.biochem.5b00977

75 Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B. et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17**, 909–915 doi:10.1038/nsmb.1838

76 Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B. et al. (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 doi:10.1038/nmeth.3810

77 Zarnegar, B.J., Flynn, R.A., Shen, Y., Do, B.T., Chang, H.Y., Khavari, P.A. et al. (2016) irCLIP platform for efficient characterization of protein–RNA interactions. *Nat. Methods* **13**, 489–492 doi:10.1038/nmeth.3840

78 Granneman, S., Kudla, G., Petfalski, E. and Tollervey, D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl Acad. Sci. USA* **106**, 9613–9618 doi:10.1073/pnas.0901997106

79 Kudla, G., Granneman, S., Hahn, D., Beggs, J.D. and Tollervey, D. (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. *Proc. Natl Acad. Sci. USA* **108**, 10010–10015 doi:10.1073/pnas.1017386108

80 Helwak, A., Kudla, G., Dudnakova, T. and Tollervey, D. (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **153**, 654–665 doi:10.1016/j.cell.2013.03.043

81 Sugimoto, Y., Vigilante, A., Darbo, E., Zirra, A., Militti, C., D'Ambrogio, A. et al. (2015) hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* **519**, 491–494 doi:10.1038/nature14280

82 Lu, Z., Zhang, Q.C., Lee, B., Flynn, R.A., Smith, M.A., Robinson, J.T. et al. (2016) RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279 doi:10.1016/j.cell.2016.04.028

83 Aw, J.G., Shen, Y., Wilm, A., Sun, M., Lim, X.N., Boon, K.-L. et al. (2016) In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell* **62**, 603–617 doi:10.1016/j.molcel.2016.04.028

84 Sharma, E., Sterne-Weiler, T., O'Hanlon, D. and Blencowe, B.J. (2016) Global mapping of human RNA-RNA interactions. *Mol. Cell* **62**, 618–626 doi:10.1016/j.molcel.2016.04.030

85 Ding, F., Lavender, C.A., Weeks, K.M. and Dokholyan, N.V. (2012) Three-dimensional RNA structure refinement by hydroxyl radical probing. *Nat. Methods* **9**, 603–608 doi:10.1038/nmeth.1976

86 Hulscher, R.M., Bohon, J., Rappé, M.C., Gupta, S., D'Mello, R., Sullivan, M. et al. (2016) Probing the structure of ribosome assembly intermediates in vivo using DMS and hydroxyl radical footprinting. *Methods* **103**, 49–56 doi: 10.1016/j.ymeth.2016.03.012

87 Kielpinski, L.J. and Vinther, J. (2014) Massive parallel-sequencing-based hydroxyl radical probing of RNA accessibility. *Nucleic Acids Res.* **42**, e70 doi:10.1093/nar/gku167

88 Cheng, C.Y., Chou, F.C., Kladwang, W., Tian, S., Cordero, P., Das, R. et al. (2015) Consistent global structures of complex RNA states through multidimensional chemical mapping. *eLife* **4**, e07600 PMID: 26035425

89 Helwak, A. and Tollervey, D. (2014) Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). *Nat. Protoc.* **9**, 711–728 doi:10.1038/nprot.2014.043

90 Fecko, C.J., Munson, K.M., Saunders, A., Sun, G., Begley, T.P., Lis, J.T. et al. (2007) Comparison of femtosecond laser and continuous wave UV sources for protein–nucleic acid crosslinking. *Photochem. Photobiol.* **83**, 1394–1404 doi:10.1111/j.1751-1097.2007.00179.x

91 Ramani, V., Qiu, R. and Shendure, J. (2015) High-throughput determination of RNA structure by proximity ligation. *Nat. Biotechnol.* **33**, 980–984 doi:10.1038/nbt.3289

92 Kladwang, W., Hum, J. and Das, R. (2012) Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. *Sci. Rep.* **2**, 517 doi:10.1038/srep00517

93 Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A.E. and Weeks, K.M. (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965 doi:10.1038/nmeth.3029

94 Kikin, O., D'Antonio, L. and Bagga, P.S. (2006) QGRS mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **34**, W676–W682 doi:10.1093/nar/gkl253

95 Huppert, J.L., Bugaut, A., Kumari, S. and Balasubramanian, S. (2008) G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* **36**, 6260–6268 doi:10.1093/nar/gkn511

96 Biffi, G., Di Antonio, M., Tannahill, D. and Balasubramanian, S. (2014) Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **6**, 75–80 doi:10.1038/nchem.1805

97 Laguerre, A., Hukezalie, K., Winckler, P., Katranji, F., Chanteloup, G., Pirrotta, M. et al. (2015) Visualization of RNA-quadruplexes in live cells. *J. Am. Chem. Soc.* **137**, 8521–8525 doi:10.1021/jacs.5b03413

98 Millevoi, S., Moine, H. and Vagner, S. (2012) G-quadruplexes in RNA biology. *Wiley Interdiscip. Rev. RNA* **3**, 495–507 doi:10.1002/wrna.1113

99 Crenshaw, E., Leung, B.P., Kwok, C.K., Sharoni, M., Olson, K., Sebastian, N.P. et al. (2015) Amyloid precursor protein translation is regulated by a 3′UTR guanine quadruplex. *PLoS ONE* **10**, e0143160 PMID: 26618502

100 Simone, R., Fratta, P., Neidle, S., Parkinson, G.N. and Isaacs, A.M. (2015) G-quadruplexes: emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS Lett.* **589**, 1653–1668 doi:10.1016/j.febslet.2015.05.003

101 Lorenz, R., Bernhart, S.H., Externbrink, F., Qin, J., Höner zu Siederdissen, C., Amman, F. et al. (2012) RNA folding algorithms with G-quadruplexes. In *Advances in Bioinformatics and Computational Biology* (Souto, M.C.P.d. and Kann, M.G., eds), pp. 49–60, Springer, Berlin, Heidelberg doi:10.1007/978-3-642-31927-3_5
102 Mergny, J.L., Phan, A.-T. and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.* **435**, 74–78 doi:10.1016/S0014-5793(98)01043-6
103 Vorlickova, M., Kejnovská, I., Sagi, J., Renčiuk, D., Bednářová, K., Motlová, J. et al. (2012) Circular dichroism and guanine quadruplexes. *Methods* **57**, 64–75 doi:10.1016/j.ymeth.2012.03.011
104 Beaudoin, J.D., Jodoin, R. and Perreault, J.-P. (2013) In-line probing of RNA G-quadruplexes. *Methods* **64**, 79–87 doi:10.1016/j.ymeth.2013.02.017
105 Kwok, C.K., Ding, Y., Shahid, S., Assmann, S.M. and Bevilacqua, P.C. (2015) A stable RNA G-quadruplex within the 5′-UTR of *Arabidopsis thaliana* ATR mRNA inhibits translation. *Biochem. J.* **467**, 91–102 doi:10.1042/BJ20141063
106 Kwok, C.K. and Balasubramanian, S. (2015) Targeted detection of G-quadruplexes in cellular RNAs. *Angew. Chem. Int. Ed.* **54**, 6751–6754 doi:10.1002/anie.201500891
107 Kwok, C.K., Sahakyan, A.B. and Balasubramanian, S. (2016) Structural analysis using SHALiPE to reveal RNA G-quadruplex formation in human precursor microRNA. *Angew. Chem. Int. Ed* **55**, 8958–8961 doi:10.1002/anie.201603562
108 Kwok, C.K. et al. (2016) rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat. Methods.* doi:10.1038/nmeth.3965
109 Kolb, H.C., Finn, M.G. and Sharpless, K.B. (2001) Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed.* **40**, 2004–2021 doi:10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.3.CO;2-X
110 Fang, R., Moss, W.N., Rutenberg-Schoenberg, M. and Simon, M.D. (2015) Probing Xist RNA structure in cells using targeted Structure-seq. *PLoS Genet.* **11**, e1005668 doi:10.1371/journal.pgen.1005668
111 Mercer, T.R., Clark, M.B., Crawford, J., Brunck, M.E., Gerhardt, D.J., Taft, R.J. et al. (2014) Targeted sequencing for gene discovery and quantification using RNA CaptureSeq. *Nat. Protoc.* **9**, 989–1009 doi:10.1038/nprot.2014.058
112 Clark, M.B., Mercer, T.R., Bussotti, G., Leonardi, T., Haynes, K.R., Crawford, J. et al. (2015) Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. *Nat. Methods* **12**, 339–342 doi:10.1038/nmeth.3321
113 Smyth, R.P., Despons, L., Huili, G., Bernacchi, S., Hijnen, M., Mak, J. et al. (2015) Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat. Methods* **12**, 866–872 doi:10.1038/nmeth.3490
114 Puchta, O., Cseke, B., Czaja, H., Tollervey, D., Sanguinetti, G., Kudla, G. et al. (2016) Network of epistatic interactions within a yeast snoRNA. *Science* **352**, 840–844 doi:10.1126/science.aaf0965
115 Li, C., Qian, W., Maclean, C.J. and Zhang, J. (2016) The fitness landscape of a tRNA gene. *Science* **352**, 837–840 doi:10.1126/science.aae0568
116 Mathews, D.H., Moss, W.N. and Turner, D.H. (2010) Folding and finding RNA secondary structure. Cold Spring Harb. Perspect. **2**, a003665 doi: 10.1101/cshperspect.a003665
117 Ding, Y., Kwok, C.K., Tang, Y., Bevilacqua, P.C. and Assmann, S.M. (2015) Genome-wide profiling of in vivo RNA structure at single-nucleotide resolution using structure-seq. *Nat. Protoc.* **10**, 1050–1066 doi:10.1038/nprot.2015.064
118 Flynn, R.A., Zhang, Q.C., Spitale, R.C., Lee, B., Mumbach, M.R., Chang, H.Y. et al. (2016) Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.* **11**, 273–290 doi:10.1038/nprot.2016.011
119 Harris, K.A., Crothers, D.M. and Ullu, E. (1995) In-vivo structural-analysis of spliced leader RNAs in *Trypanosoma brucei* and *Leptomonas collosoma* — a flexible structure that is independent of Cap4 methylations. *RNA* **1**, 351–362 PMID: 7493314
120 Noller, H.F. and Chaires, J.B. (1972) Functional modification of 16S ribosomal RNA by kethoxal. *Proc. Natl Acad. Sci. USA* **69**, 3115–3118 doi:10.1073/pnas.69.11.3115
121 Wilkinson, K.A., Merino, E.J. and Weeks, K.M. (2006) Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* **1**, 1610–1616 doi:10.1038/nprot.2006.249
122 Mortimer, S.A. and Weeks, K.M. (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J. Am. Chem. Soc.* **129**, 4144–4145 doi:10.1021/ja0704028
123 Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, Jr, J.W., , Swanstrom, R. et al. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**, 711–716 doi:10.1038/nature08237
124 Mortimer, S.A. and Weeks, K.M. (2008) Time-resolved RNA SHAPE chemistry. *J. Am. Chem. Soc.* **130**, 16178–16180 doi:10.1021/ja8061216
125 Steen, K.A., Rice, G.M. and Weeks, K.M. (2012) Fingerprinting noncanonical and tertiary RNA structures by differential SHAPE reactivity. *J. Am. Chem. Soc.* **134**, 13160–13163 doi:10.1021/ja304027m
126 Rice, G.M., Leonard, C.W. and Weeks, K.M. (2014) RNA secondary structure modeling at consistent high accuracy using differential SHAPE. *RNA* **20**, 846–854 doi:10.1261/rna.043323.113
127 Poulsen, L.D., Kielpinski, L.J., Salama, S.R., Krogh, A. and Vinther, J. (2015) SHAPE selection (SHAPES) enrich for RNA structure signal in SHAPE sequencing-based probing data. *RNA* **21**, 1042–1052 doi:10.1261/rna.047068.114
128 Wan, Y., Qu, K., Ouyang, Z. and Chang, H.Y. (2013) Genome-wide mapping of RNA structure using nuclease digestion and high-throughput sequencing. *Nat. Protoc.* **8**, 849–869 doi:10.1038/nprot.2013.045
129 Del Campo, C., Bartholomäus, A., Fedyunin, I. and Ignatova, Z. (2015) Secondary structure across the bacterial transcriptome reveals versatile roles in mRNA regulation and function. *PLoS Genet.* **11**, e1005613 doi:10.1371/journal.pgen.1005613
130 Zheng, Q., Ryvkin, P., Li, F., Dragomir, I., Valladares, O., Yang, J. et al. (2010) Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. *PLoS Genet.* **6**, e1001141 doi:10.1371/journal.pgen.1001141
131 Li, F., Zheng, Q., Ryvkin, P., Dragomir, I., Desai, Y., Aiyer, S. et al. (2012) Global analysis of RNA secondary structure in two metazoans. *Cell Rep.* **1**, 69–82 doi:10.1016/j.celrep.2011.10.002
132 Li, F., Zheng, Q., Vandivier, L.E., Willmann, M.R., Chen, Y., Gregory, B.D. et al. (2012) Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. *Plant Cell* **24**, 4346–4359 doi:10.1105/tpc.112.104232
133 Underwood, J.G., Uzilov, A.V., Katzman, S., Onodera, C.S., Mainzer, J.E., Mathews, D.H. et al. (2010) Fragseq: transcriptome-wide RNA structure probing using high-throughput sequencing. *Nat. Methods* **7**, 995–1001 doi:10.1038/nmeth.1529

Biochemical Society Transactions (2016) **44** 1395–1410
DOI: 10.1042/BST20160075

134 Lucks, J.B., Mortimer, S.A., Trapnell, C., Luo, S., Aviran, S., Schroth, G.P. et al. (2011) Multiplexed RNA structure characterization with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc. Natl Acad. Sci. USA* **108**, 11063–11068 doi:10.1073/pnas.1106501108

135 Mortimer, S.A., Trapnell, C., Aviran, S., Pachter, L. and Lucks, J.B. (2012) SHAPE-Seq: high-throughput RNA structure analysis. *Curr. Protoc. Chem. Biol.* **4**, 275–297 PMID: 23788555

136 Seetin, M.G., Kladwang, W., Bida, J.P. and Das, R. (2014) Massively parallel RNA chemical mapping with a reduced bias MAP-seq protocol. *Methods Mol. Biol.* **1086**, 95–117 doi:10.1007/978-1-62703-667-2_6

137 Lin, Y., May, G.E. and Joel McManus, C. (2015) Mod-seq: A high-throughput method for probing RNA secondary structure. *Methods Enzymol.*, **558** 125-152. doi:10.1016/bs.mie.2015.01.012

138 Lavender, C.A., Gorelick, R.J. and Weeks, K.M. (2015) Structure-based alignment and consensus secondary structures for three HIV-related RNA genomes. *PLoS Comput. Biol.* **11**, e1004230 doi:10.1371/journal.pcbi.1004230

139 Mauger, D.M., Golden, M., Yamane, D., Williford, S., Lemon, S.M., Martin, D.P. et al. (2015) Functionally conserved architecture of hepatitis C virus RNA genomes. *Proc. Natl Acad. Sci. USA.* **112**, 3692–3697 doi: 10.1073/pnas.1416266112

140 Hector, R.D., Burlacu, E., Aitken, S., Bihan, T.L. and Tuijtel, M. (2014) Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. *Nucleic Acids Res.* **42**, 12138–12154 doi:10.1093/nar/gku815

141 Homan, P.J., Favorov, O.V., Lavender, C.A., Kursun, O., Ge, X., Busan, S. et al. (2014) Single-molecule correlated chemical probing of RNA. *Proc. Natl Acad. Sci. USA* **111**, 13858–13863 doi:10.1073/pnas.1407306111

142 Loughrey, D., Watters, K.E., Settle, A.H. and Lucks, J.B. (2014) SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. *Nucleic Acids Res.* **42**, e165 doi:10.1093/nar/gku909

143 Watters, K.E., Yu, A.M., Strobel, E.J., Settle, A.H. and Lucks, J.B. (2016) Characterizing RNA structures in vitro and in vivo with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods* **103**:34–48

144 Incarnato, D., Neri, F., Anselmi, F. and Oliviero, S. (2014) Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. *Genome Biol.* **15**, 491 doi:10.1186/s13059-014-0491-2

145 Watters, K.E., Abbott, T.R. and Lucks, J.B. (2016) Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq. *Nucleic Acids Res.* **44**, e12 doi:10.1093/nar/gkv879