Chemical RNA Crosslinking: Mechanisms, Computational Analysis and Biological Applications

Willem A. Velema[1*] and Zhipeng Lu[2*]
[1]Institute for Molecules and Materials, Radboud University Nijmegen, The Netherlands
[2]Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90033
*To whom correspondence should be addressed: willem.velema@ru.nl, zhipengl@usc.edu

Abstract.
In recent years, RNA has emerged as a multifaceted biomolecule that is involved in virtually every function of the cell and is critical for human health. This has led to a substantial increase in research efforts to uncover the many chemical and biological aspects of RNA and target RNA for therapeutic purposes. In particular, analysis of RNA structures and interactions in cells has been critical for understanding their diverse functions and druggability. In the last 5 years, several chemical methods have been developed to achieve this goal, using chemical crosslinking combined with high-throughput sequencing and computational analysis. Applications of these methods resulted in important new insights into RNA functions in a variety of biological contexts. Given the rapid development of new chemical technologies, a thorough review on the past and future of this field is provided. In particular, the various RNA crosslinkers and their mechanisms, the computational analysis and challenges and illustrative examples from recent literature are discussed.

1. Introduction.

RNA is increasingly being recognized as a versatile biomacromolecule involved in virtually every process in the cell.[1,2] In addition to serving as a passive carrier of genetic information like DNA, RNA can fold into complex structures like proteins.[3] The various structures formed by RNA, either within a single RNA molecule, or between different RNAs, carry out diverse and active cellular functions, such as catalysis, scaffolding, and guiding.[1,4]. Normal and abnormal functions of RNA underlie a variety of human pathologies, such as virus infections, nucleotide repeat disorders and splicing diseases.[5] Technological advancements in several fields have culminated in new RNA-based and RNA-targeting therapeutic approaches in recent years, including small molecule drugs that target highly structured regions,[6–8] antisense oligos that target low structured regions or alter RNA structures,[9] and mRNA therapies based on modification chemistries that stabilize RNA and minimize innate immunogenicity.[10]

To dissect the critical roles of RNA in biology, study disease mechanisms and develop RNA-targeting therapeutics, structural analysis of RNA has played an essential role.[11] While conventional physical methods, such as NMR, X-ray crystallography and cryo-EM have been instrumental in protein structure analysis, their applications in RNA have been more limited, primarily due to RNA's large size, high flexibility, and strong dependence on physiological environments.[11,12] Computational structure modeling based on minimal free energy calculations and phylogenetic analysis of conservation and covariation also suffer from multiple limitations, such as lack of understanding of RNA folding rules and high computational cost. Therefore, direct measurements of RNA structures in cells have been critical for understanding RNA behavior in various biological and pathological processes.[13] A variety of chemical reactions have been developed and exploited that can modify RNA at certain positions (**Fig. 1**) depending on nucleotide reactivity, flexibility, or accessibility, which correlate with RNA structural constraints.[14] For example, dimethyl sulfate (DMS) selectively alkylates the N1 position of adenine and N3 position of cytosine on unpaired nucleotides[15] and the 2'-OH in flexible regions can be acylated with Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) reagents.[16–18] Resulting reactivity profiles have been useful to improve secondary and tertiary structure modeling,[19–21] however, the 1D information obtained with these experiments is not necessarily definitive evidence for specific structures. More recently, correlated chemical probing coupled with computational deconvolution has been used to discover potential contacts and alternative conformations in relatively short RNA regions and simple conformations.[22–24]

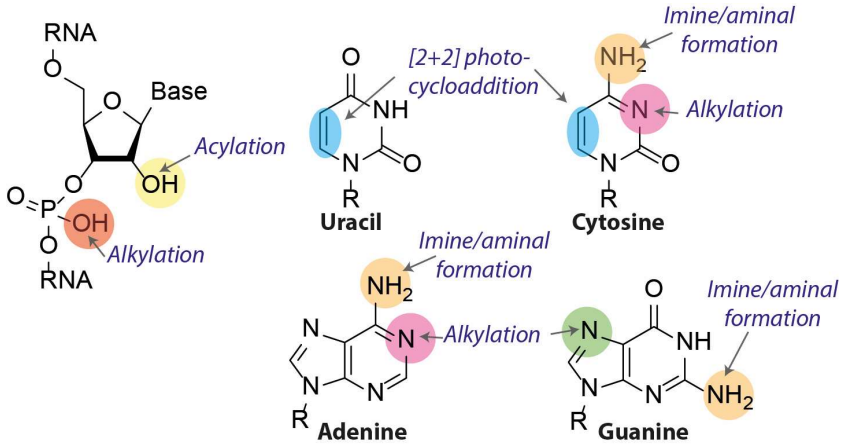

***Figure 1**. The molecular structure and reactive sites of RNA. The phosphate group (red) can be alkylated with diazo compounds. The 2'-OH (yellow) is prone to acylation. The C5-C6 double bond in uracil and cytosine (blue) can undergo a photocycloaddition. The exocyclic amines (orange) can react with aldehydes to form an imine or aminal. The N7 position of guanine (green) can be alkylated as well as the N1 of adenine and N3 of cytosine (pink).*

In theory, the structure of any object can be uniquely determined by the coordinates of their components, which should be equivalent to the distances among the components. Therefore, measuring the spatial distances between nucleotides should allow de novo determination of RNA structures in any biological sample. Chemical crosslinking of RNA in cells using compounds of defined sizes and subsequent identification of crosslinked nucleotides represent a practical path towards this goal. Pioneering work by Hearst and colleagues in the 1970s and 1980s[25–28] demonstrated this principle using psoralen, a duplex-specific agent that crosslinks opposing pyrimidine nucleobases with two consecutive [2+2] photocycloadditions.[29] Subsequent applications provided the first physical evidence for snRNA-dependent pre-mRNA splicing, and snoRNA-guided ribosomal RNA processing.[12,30–32] Rapid progress in both crosslinking chemistry and sequencing technology in the last few years have led to several new high-throughput methods for the analysis of RNA structures and interactions.

While other recent reviews have summarized these methods,[12,14,33–35] a critical analysis and evaluation of the chemical mechanisms, computational tools and applications in biology have been lacking. Simultaneous consideration of these three aspects is necessary to solve increasingly challenging problems in RNA chemistry and biology. In this review, we focus on these topics together, to inspire chemists to further develop specialized chemical crosslinking

agents necessary to tackle outstanding questions in RNA biology and provide a guide for biologists to choose the most appropriate methods for addressing specific biological questions.

2. Chemistry of current technologies.

2.1. RNA's chemical reactivity.

The molecular structure of RNA contains multiple functional groups that can potentially be targeted with (photo)chemical reagents to effectuate crosslinks (**Fig. 1**). The phosphodiester in the RNA backbone can be alkylated with diazo compounds (red, **Fig. 1**),[36] which has been successfully demonstrated with caging agents to control RNA function.[37,38] Gillingham and coworkers recently showed that terminal phosphates are more prone to *O*-alkylation with diazo compounds than internal phosphate diesters, for which usually a large excess of reagent is required. This was attributed to the substantial difference in pK_a (~6-7 for phosphate mono-ester and ~1 for phosphate diester).[36] The 2'-OH position (yellow, **Fig. 1**) can be acylated with activated carbonyl reagents, which forms the basis of Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE), a widely used method to determine RNA secondary structure.[16,17] More recent bifunctional acylators target the 2'-OH position to establish crosslinking.[39–41] The unusual nucleophilicity of the 2'-OH position has been attributed to oxyanion formation at this position due to inductive effects provided by neighboring 3' and 4' oxygens and the nucleobase nitrogen.[42,43] Pyrimidine bases can undergo [2+2] photocycloadditions using their C5-C6 double bond (blue, **Fig. 1**).[28] This feature has been extensively exploited with psoralen crosslinkers.[29] More recent photochemical crosslinkers such as carbazoles and coumarins undergo photocycloadditions as well with improved efficiencies.[44] The exocyclic amines of cytosine, adenine and guanine (orange, **Fig. 1**) can react with aldehydes to form imines.[45] When using formaldehyde an aminal bond is formed that crosslinks opposing nucleobases.[45] The N7 position of guanine (green, **Fig. 1**) can act as a nucleophile towards nitrogen mustards.[46] Interestingly, the N7 position of adenine is unreactive towards nitrogen mustards, which is mainly attributed to the lower nucleophilicity as compared to the N7 of guanine.[46,47] The N1 of adenine and N3 of cytosine (pink, **Fig. 1**) can act as nucleophiles towards appropriately electrophilic compounds.[15] This is mainly exploited with DMS footprinting[15] to elucidate RNA secondary structures. Moreover, the N3 position of guanine and uracil can also be methylated by DMS, but at lower rates causing this reactivity to be mainly ignored.[48]

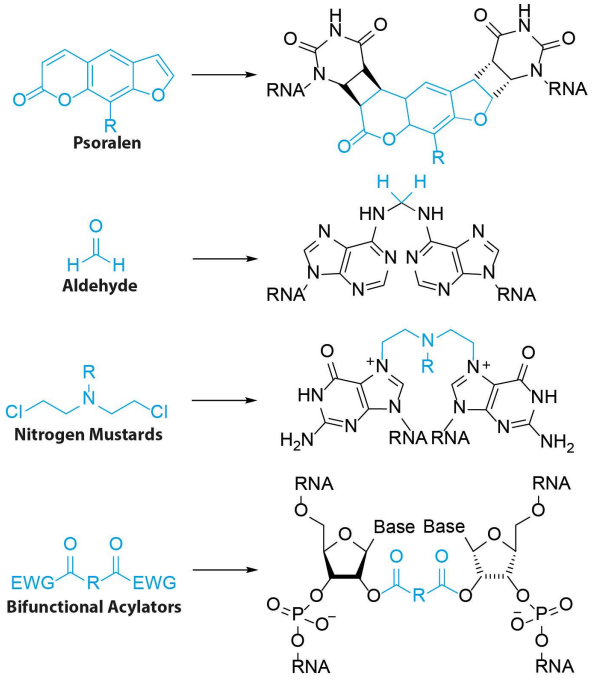

***Figure 2**. The molecular structure and corresponding RNA adducts of the most commonly employed external RNA crosslinking reagents.*

We make a distinction between external and internal crosslinking reagents. External crosslinking reagents are the main focus of this review and can be added exogenously to samples without the need of first modifying the RNA under investigation allowing to study native RNA. Other advantages are that they are generally more accessible and pan acting, enabling transcriptome-wide interrogations. Conversely, internal crosslinking reagents are installed in the RNA scaffold prior to the study either chemically or enzymatically and include coumarins,[49] carbozeles,[44] thionucleotides,[50] diazirines[51] and platinum complexes[52] among others and have recently been reviewed.[53]

Four classes of external crosslinking reagents (**Fig. 2**) are predominantly used and each offer unique advantages and react with distinct functional groups on the RNA scaffold to establish intra- or intermolecular crosslinks.

2.2. Psoralen crosslinkers.

One of the most commonly employed RNA crosslinkers for structural studies is psoralen.[25] This natural product can undergo two consecutive [2+2] photocycloadditions to crosslink opposing pyrimidine bases (**Fig. 3a**).[28] Psoralens first intercalate into duplex regions, placing the reactive 3,4 and 4',5' double bonds in a favorable position towards the C5-C6 double bond of pyrimidine bases. Exposure to 365 nm light results in a photocycloaddition.[28] Both the 3,4 and 4',5' double bonds can in principle react first, however only 4',5' adducts still absorb at 365 nm and can therefore undergo a second photocycloaddition with the 3,4 double bond resulting in a successful crosslink.[25,54] Importantly, this crosslink can be reversed by exposure to ~250 nm light, which is heavily exploited in RNA structural studies to simplify data analysis. This reversal reaction has reportedly quantum yields between 0.16-0.30, depending on the wavelength.[26] The furan-side was more readily reversed than the pyrone-side with a ~2-fold difference in rate at pH 2.2 and ~20-fold difference at pH 7.5.[26] In practice, the reversal reaction has been reported with limited efficiency,[39] which in part was attributed to significant RNA degradation.[55] A base-catalyzed rearrangement has been reported with higher efficiency that selectively cleaves the crosslink at the pyrone side of psoralen.[56] To overcome RNA damage including potential cyclobutane dimers and (6-4) products (**Fig. 3c**) due to UV exposure caused by psoralen (un)crosslinking, Lu and colleagues developed a protocol[57] based on acridine orange (**Fig. 3c**) singlet-state quenching.[58] They demonstrated that 30% of RNA remains intact after 30 min irradiation with 254 nm at 4 mW/cm^2 in the presence of acridine orange, compared to less than 0.5% in its absence, enabling efficient application of psoralen crosslinking for RNA structure determination.

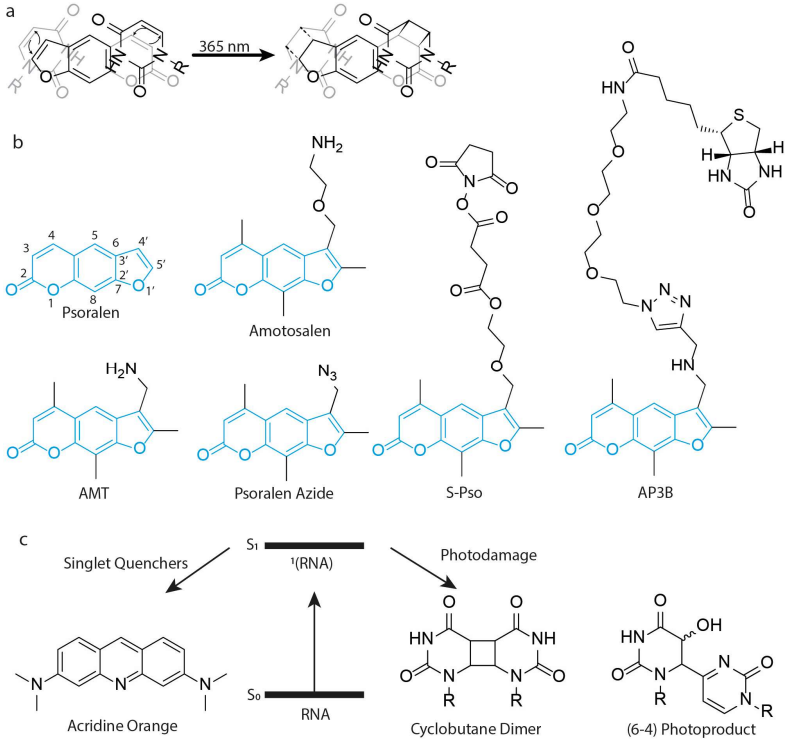

***Figure 3**. Psoralen crosslinkers. (a) Crosslinking mechanism. Psoralen intercalates between opposing pyrimidines. UV exposure initiates a [2+2] photocycloaddition between the C5-C6 double bond of uracil or cytosine and the 3,4 and 4',5' double bond of psoralen to form a cyclobutane and crosslinking the pyrimidines. (b) Molecular structure of psoralen derivatives. (c) Singlet quenchers like acridine orange can protect RNA from photodamage.*

Over 100 psoralen derivatives have been reported with improved photophysical, physicochemical and biochemical properties as compared to the parent compound.[25] For example, AMT (**Fig. 3b**) bears a primary amine that improves solubility in water at physiological conditions to ~1 mg/mL, enabling application for RNA structural studies, first reported by Calvet and Pederson to study heterogenous nuclear RNA (hnRNA) in live cells.[59] Using AMT they found that hnRNA contains double stranded regions that are organized in an accessible manner within the nucleus. AMT efficiently crosslinked double stranded regions in cells with a 7.2 fold increase as compared to UV light alone.[59] With recent revolutionizing developments in Next Generation Sequencing and bioinformatic analysis the use of AMT to determine RNA structure on a transcriptome-wide level has exponentially increased. Psoralen Analysis of RNA Interactions and Structures (PARIS) and LIGation of interacting RNA followed by high-throughput sequencing (LIGR-

seq) used AMT crosslinking to map transcriptome-wide RNA structures and discover previously unknown RNA-RNA interactions.[60,61]

To further improve the analysis of dynamic RNA structures and interactions, PARIS2 uses amotosalen (**Fig. 3b**) instead of AMT for crosslinking.[57] The limited solubility of AMT at ~1 mg/mL in water was hampering the efficiency in capturing RNA-RNA interactions. Amotosalen bears a primary amine and ether group and exhibits markedly increased solubility in water of 230 mg/mL. In vivo, it was found that a 10 fold increase in concentration (0.5 mg/mL AMT vs. 5.0 mg/mL amotosalen) resulted in a 7 fold increase in crosslinked RNA. This overall improved crosslinking allowed for establishing the first whole genome structure of enterovirus D68 and dynamic interaction networks for the U8 snoRNA, where genetic mutations cause a neurological disorder.

To increase the functionality of psoralens, azido modified derivatives were designed that can undergo azide-alkyne cycloadditions. For example, Hall and coworkers used Psoralen Azide (**Fig. 3b**) to append to the 3'-end of alkyne modified pre-microRNAs with up to 94% conjugation yield.[62] A different azido derivative bears the azide group at the end of a triethylene glycol linker attached to the primary amine of AMT and forms the basis of Crosslinking Of Matched RNA And Deep Sequencing (COMRADES).[63] Using this psoralen derivative, crosslinked RNA can be selectively captured and enriched using a biotin ligation and Streptavidin pulldown.[64–66] The presence of the azide group did not affect the crosslink efficiency and using this method the researchers determined the architecture of Zika virus inside cells.[63]

To conjugate psoralen to an RNA of interest, Rana and coworkers designed the NHS-ester bearing derivative S-Pso (**Fig. 3b**).[67] This was ligated to an amine bearing miRNA-29a mimic using standard amide-coupling conditions to afford conjugated RNA in quantitative yields. After transfection into HeLa cells, S-Pso labeled miRNA-29a efficiently silenced luciferase gene expression of a reporter plasmid containing a miRNA-29a target site in its 3' UTR to ~15%, showing that the psoralen group did not affect miRNA silencing. The miRNA mimic was then applied to identify targets in live cells. After photocrosslinking, captured transcripts were quantified by RT-qPCR. Tet2 RNA was identified as a miRNA-29a target with 20-fold enrichment after crosslinking. This study highlights the potential use of psoralen crosslinkers in RNA target identification.

Biotinylated psoralen would allow for direct enrichment of crosslinked RNA using Streptavidin beads. Taking advantage of this, Nagarajann, Wan and coworkers[68] developed Sequencing of Psoralen crosslinked, ligated and Selected Hybrids (SPLASH). In particular the sensitivity of their method increased from ~0.45 with psoralen to ~0.75 with biotinylated psoralen. Using SPLASH the authors identified hundreds of known and unknown snoRNA-rRNA binding sites. One disadvantage of this protocol is that cellular uptake of biotinylated psoralen was low, and 0.01% digitonin was added to increase uptake. The psoralen derivative used in this study was modified at the C8 position with biotin. Lin and coworkers found that when biotin is appended to the 4'-amine group of AMT (AP3B, **Fig. 3b**), the efficiency and crosslinking is significantly improved.[69] Using a gel-shift assay and dsDNA *in vitro*, a 100 fold increase in efficiency was observed. In cells a 5-fold increase in biotinylation of DNA was observed. The crosslinking efficiency for RNA was not determined.

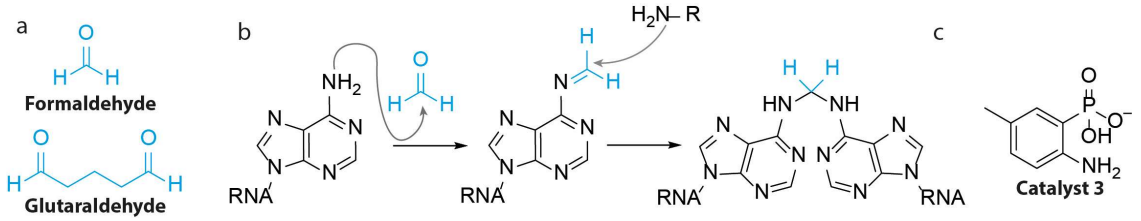

***Figure 4**. Aldehyde crosslinking. (a) Structure of formaldehyde and glutaraldehyde. (b) Crosslinking mechanism of formaldehyde. (c) Molecular structure of **catalyst 3** that reverses formaldehyde crosslinks.*

2.3. Aldehyde crosslinkers.

The use of aldehyde to crosslink biological samples dates back to the 19th century, when Blum reported successful tissue fixation with formaldehyde (**Fig. 4a**).[70] Formalin-Fixed Paraffin Embedded (FFPE) treatment is used for preservation for long-term storage and preservation of patient samples. Glutaraldehyde (**Fig. 4a**) has been used extensively for crosslinking proteins,[71] but studies with RNA are scarce, although it is likely that it can crosslink nucleic acids. The dialdehyde glyoxal has recently been used as well for temporary caging of nucleobases[72] and when incorporated in the backbone of DNA it was shown to crosslink to proteins.[73] Formaldehyde forms adducts and crosslinks with biomolecules that protects them from degradation. Specifically, formaldehyde can react with the exocyclic amines of adenine, guanine and cytosine to form imine and hemiaminal adducts and aminal crosslinks (**Fig. 4b**).[45] The first step is rapid, but the formed imine inhibits base pairing interactions, significantly slowing down crosslinking, which requires the second nucleophile to be in close proximity.[74]

The formed adducts are in principle reversible and allows for analysis of genetic material after long term storage.[75] After RNA extraction, samples are typically incubated at 80-90 °C in Tris buffer for several hours, which reverses hemiaminal formation. These relatively harsh conditions have been shown to affect the integrity of RNA, impeding meaningful quantification of RNA levels. Kool and coworkers designed catalysts that can speed up adduct removal, yielding higher quality RNA. In particular phosphanilate **catalyst 3** (**Fig. 4c**) efficiently reversed hemiaminal adducts. Incubation at 5 mM for 2 hours reversed ~50% of adducts compared to 11% without catalyst. While recovering

adducts from FFPE prepared cell samples, up to 25-fold enhancement in recovery was found for **catalyst 3** compared to no catalyst, as was quantified with qRT-PCR. Longer RNA amplicons benefited most from catalytic adduct removal. It was suggested that the mechanism of catalyst-assisted reversal is based on general acid catalysis with possible nucleophilic catalysis.

Apart from RNA preservation, aldehyde crosslinking can be used for mapping RNA-RNA interactions as well. Guttman, Lander and coworkers exploited this to study U1 small nuclear RNA and the large ncRNA Malat1.[76] The authors combined AMT (**Fig. 3b**) crosslinking with formaldehyde crosslinking and found that both methods yield different RNA-RNA interactions: AMT provided information on duplexed interactions at high resolution, while formaldehyde could capture a broader range of interactions. After crosslinking, RNAs of interest were captured with antisense oligonucleotides and analyzed with high-throughput sequencing. Strong enrichment of target RNAs (>1000 fold) was obtained with both AMT and formaldehyde crosslinking. Using this approach, it was found that U1 RNA targets 5' splice sites throughout introns and Malat1 interacts with pre-mRNAs mostly through protein intermediates.

2.4. Nitrogen Mustards.

Originally studied in the 1940's for their anticancer properties, nitrogen mustards have been found to be potent DNA crosslinkers and as such have found wide clinical use. This extraordinary feature was quickly realized to be applicable to crosslink RNA to investigate RNA-RNA interactions.[77,78] Nitrogen Mustards are first activated by forming an aziridinium cation and elimination of chloride (**Fig. 5a**).[46] The aziridinium cation is sequentially attacked by the N7 of guanine resulting in alkylation. Two subsequent attacks lead to a bifunctional adduct, crosslinking RNA (**Fig. 5a**).

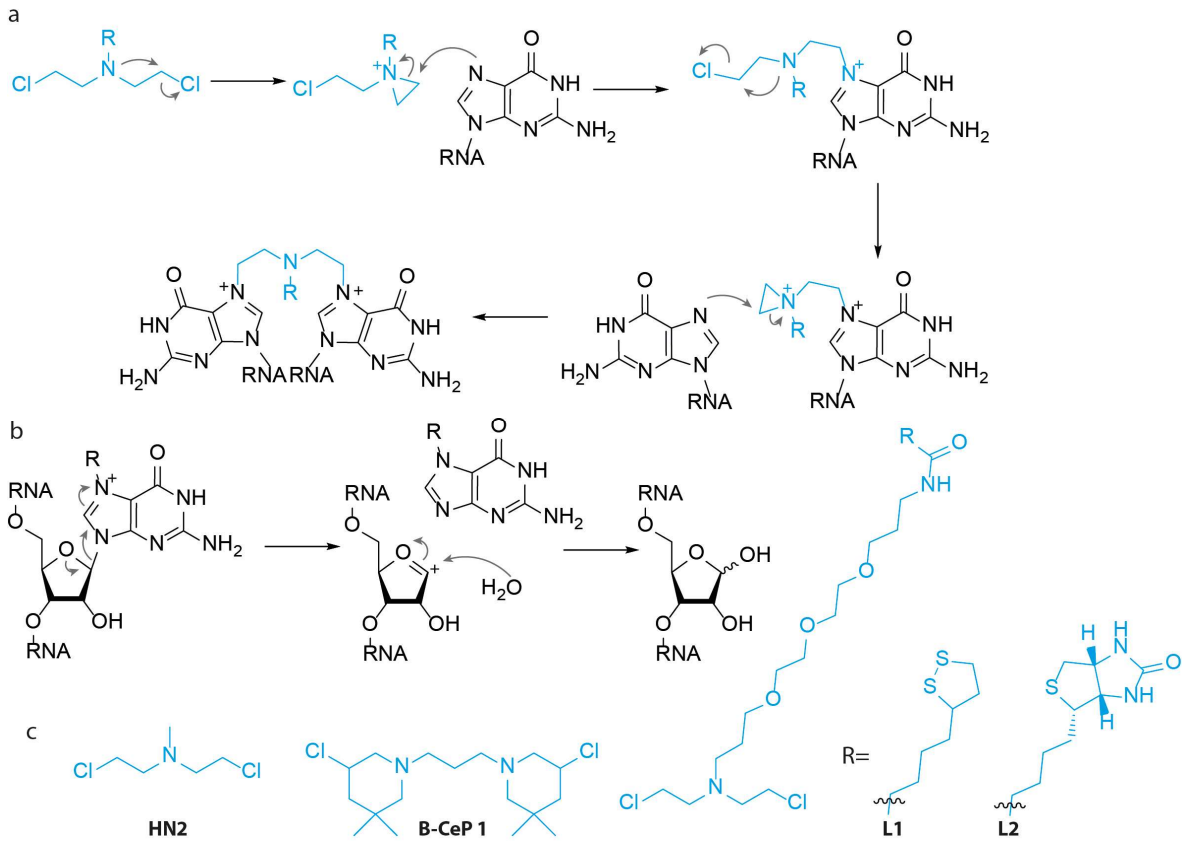

Figure 5. *Nitrogen Mustards. (a) Crosslinking mechanism. An aziridinium is formed that is attacked by the N7 nitrogen of guanine. Two consecutive reactions result in intra- or intermolecular crosslinks. (b) Depurination mechanism. N7-alkylated guanine is electron poor, promoting cleavage of the N-glycosidic bond. (c) Molecular structure of Nitrogen Mustard derivatives.*

One potential consequence of N7 alkylation is depurination. Many studies have investigated this mechanism in DNA,[79] which involves breakage of the N-glycosidic bond promoted by the positive charge and subsequent reaction with water (**Fig. 5b**).[80] N7 alkylated guanine was shown to depurinate 10^6 more rapidly than guanine under physiological conditions.[81] RNA is considered less prone to depurination, because the 2'-OH destabilizes the oxocarbonium intermediate,[80] but several cases in RNA have been reported,[82] which could hamper interpretation of obtained data. There appears to be a physiological relevance to this reaction and enzymatic repair pathways have been identified.[83]

In one of the first reports, the authors applied **HN2** (**Fig. 5c**) to study ribosomal subunits in *Escherichia coli* (*E. coli*). It was first attempted to crosslink 16S and 23S RNA using UV crosslinking, but no interactions were found. **HN2** did efficiently crosslink 16S-23S RNA as analyzed by gel electrophoresis, which led to the conclusion that there are several RNA-RNA interactions within the ribosomal particle and that these can be explored with chemical crosslinking methods. Datta and Weiner applied this principle to investigate higher order RNA structures and tertiary interactions.[84] Nuclear extracts were subjected to 20 mM solutions of **HN2** and analyzed using sequencing gels. Intramolecular crosslinks in U2 snRNA were apparent and could be localized within regions of a few nucleotides. Exact pinpointing of crosslink sites was hampered by monoadducts. Nevertheless, the data clearly supported a tertiary structure model for U2 snRNA. Using mass spectrometry (MS), Farbis and coworkers analyzed **HN2** crosslinking sites in HIV-1 SL1A RNA.[85] This structured RNA contains a large flexible loop and an internal bulge. Crosslinking was performed with 250 µM HN2 and subsequently digested with RNase A. Using MS, crosslinked fragments were observed that originate from both the loop and bulge region. Interestingly, Guanine-Adenine crosslinks were observed as well, implying that nitrogen mustards do not exclusively react with guanine.

More recently, a set of bis-3-chloropiperidines were reported based on the natural product 593A,[86] a naturally occurring nitrogen mustard isolated from *Streptomyces griseoluteus*.[87] Bis-3-chloropiperidines, including **B-CeP 1** (**Fig. 5c**), showed efficient alkylation of a model DNA strand and crosslinks were observed with only modest concentrations of 50 µM. Inspired by this, Sosic and coworkers applied **B-CeP 1** to investigate RNA tertiary structures.[88] When tested on a model RNA construct, **B-CeP 1** rapidly alkylated RNA and substantial alkylation was observed after only 1 hour incubation at 50 µM. Interestingly, no single strand breaks were observed in RNA, whereas a similar experiment with DNA yielded extensive backbone cleavage.[86] To study RNA-RNA interactions, **B-CeP 1** was applied to the HIV-1 dimerization initiation site. Several inter- and intramolecular crosslinks were detected using MS. Experiments were performed at both 25 °C and 95 °C to discriminate between these two types of crosslinks. The apparent ability of **B-CeP 1** to crosslink duplex regions, renders it an interesting tool to elucidate higher order structures.[88]

Modified nitrogen mustards have been developed to increase the functionality of these crosslinkers. In a recent study,[89] a nitrogen mustard derivative was prepared bearing a cyclic disulfide group (**L1, Fig. 5c**) that can be immobilized on a gold surface and used in Surface Plasmon Resonance (SPR). **L1** crosslinked a model DNA strand with ~100% efficiency and was then used to measure cytosine methylation (mC) with SPR. Anti mC antibodies do not recognize 5-mC in duplex DNA, but do bind to 5-mC located in bulged dsDNA and the k_D was determined to be $7.70*10^{-3}$ using SPR. Interestingly, **L1** crosslinked and immobilized DNA showed a similar k_D of $5.60*10^{-3}$ towards anti mC antibody enabling the analysis of 5-mC in genomic DNA. In a later study,[90] the authors reported the design and synthesis of a biotin bearing nitrogen mustard **L2** (**Fig. 5c**). The biotin provided enhanced functionality and crosslinked DNA was captured to Streptavidin coated microtiter plates. Methylated cytosine was now quantified with an anti mC antibody and secondary antibody labeled with horseradish peroxidase. Using biological samples, the amount of 5-mC in mouse brain and intestine was determined to be 0.65% and 0.68% respectively. Although these examples were applied to DNA, we believe that they could find use in analysis of RNA structure as well.

2.5. Bifunctional Acylators.

Early work by Knorre and coworkers[91] showed that the 2'-OH of RNA is readily acylated when reacted with acetic anhydride in water at 0.25 M. Taking advantage of this, Weeks and coworkers invented SHAPE to deduce RNA secondary structure.[16] When reacted with *N*-methylisatoic anhydride (13 mM), acyl adducts were left on flexible positions of RNA, that were shown to stall reverse transcriptase. Using primer extension, the exact position of these adducts could be determined, which helped to deduce RNA secondary structures.[16] This pioneering method has become widespread in the field and several new versions of acylating SHAPE reagents have been reported in the last few years.[18,92]

Two consecutive acylating reactions on opposing nucleotides should result in a crosslink (**Fig. 6**). Kool and coworkers explored this possibility using BINARI probes (**Fig. 6b**).[39] The bifunctional probes bear two carbonylimidazoles that can react with the 2'-OH of opposing nucleotides. Depending on linker length between the reactive groups, crosslinking efficiencies between 45% and 84% were observed on a model self-complementary RNA duplex. To enable downstream analysis of crosslinked RNA, azide trigger groups were installed to reverse the crosslink. Azides could be reduced to amines using phosphines, promoting lactam formation and crosslink reversal. Reversal efficiencies ranged between 2% and 70% depending on phosphine, with 20 mM tris(hydroxypropyl) phosphine (THPP) being most efficient. The applicability of this crosslinking method was demonstrated by protecting RNA from nuclease digestion. A model RNA was fully degraded by S1 nuclease and RNase T1, whereas BINARI crosslinked RNA remained intact. Crosslink reversal by THPP liberated the model RNA strand, illustrating the potential use for chemical crosslinking for temporary RNA preservation.

In 2022, Lu, Velema and colleagues explored the possibility to exploit bifunctional acylators for RNA structure determination.[40] This method, named Spatial 2'-Hydroxyl Acylation Reversible Crosslinking (SHARC), used simple dicarboxylic acids that could be activated with CDI in a single step to crosslink RNA. In particular, DPI (**Fig. 6b**) efficiently crosslinked RNA up to 97%. To reverse the crosslink, mild alkaline conditions were used to hydrolyze the formed carboxylate ester, but not affect RNA integrity. Full crosslink reversal was observed in 100 mM Borate buffer pH 10.0 for 2 h at 37 °C, with no apparent RNA degradation. Combining this with exonuclease trimming and proximity ligation, allowed for determining RNA tertiary interactions (see section 3).

Weeks and coworkers developed *trans*-bis-isatoic anhydride[41] (TBIA, **Fig. 6b**) to crosslink 2'-OH of nucleotides that are close in space, which formed the basis for SHAPE-JuMP to interrogate RNA tertiary interactions. Applying TBIA to RNase P, ~5-10% of RNA was crosslinked as apparent from lower mobility on PAGE. Using an engineered reverse transcriptase that can 'jump' across the crosslink, the crosslinked sites were permanently recorded in the product cDNA strand, revealing tertiary RNA interactions.

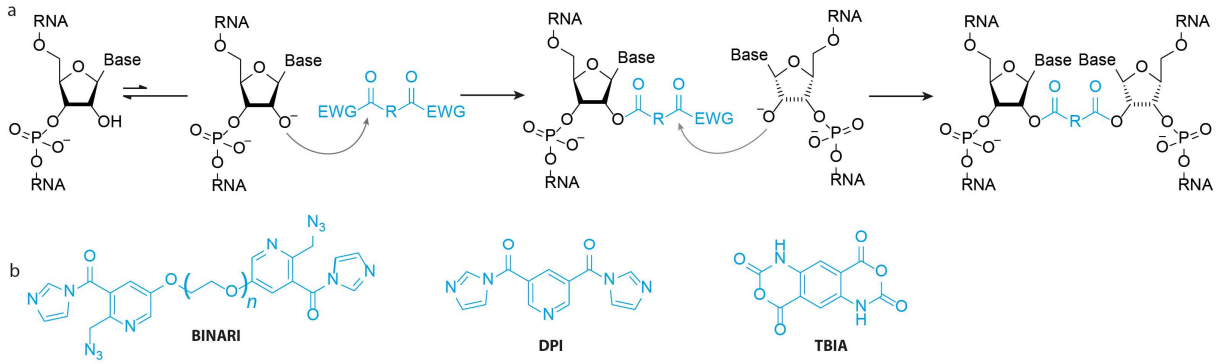

***Figure 6**. Bifunctional acylators. (a) Crosslinking mechanism. The 2'-OH attacks one of the carbonyls of the crosslinking reagent, while a second 2'-OH group that is close through space will attack the other carbonyl, effectuating a crosslink. EWG = electron withdrawing group (b) Molecular structure of bifunctional acylators.*

Taken together, there are several options available when planning RNA crosslinking experiments. Historically, psoralens have been favored and continue to be indispensable tools for RNA structure determination. Advantages include relative selectivity for duplex regions and well characterized chemistry. Recent advances in 2'-OH acylators, provide an attractive alternative for psoralens. High crosslink and reversal efficiencies outperform most psoralen analogues. Current limitations are exclusive acylation of flexible RNA regions and mono-acylation, complicating data analysis. We expect chemical improvements to address these drawbacks. Examples of aldehyde and nitrogen mustard crosslinking for RNA structure determination are scarce and the main drawback appears to be crosslink efficiency and reversibility for aldehydes and depurination and toxicity in the case of nitrogen mustards.

3. Measuring RNA Structures and Interactions with Chemical Crosslinkers

Once RNA has successfully been crosslinked, complexes can be analyzed using multiple different methods to identify the two crosslinked fragments. Classically, low-throughput methods have been used, including electron microscopy, gel electrophoresis and low-throughput enzymatic sequencing.[93] Development of high-throughput sequencing methods has enabled simultaneous measurement of transcriptome-wide RNA structures and interactions. Typical workflows of 'crosslink-ligation-sequencing' methods include the following major steps: in vivo crosslinking, RNA fragmentation, enrichment of crosslinked fragments, proximity ligation, crosslink reversal, adapter ligation, reverse transcription, PCR amplification and sequencing (**Fig. 7a**). The gapped reads obtained after sequencing reveal base paired or spatial proximal RNA fragments, which are incorporated into 2D or 3D structure modeling, using various published computational tools.[22–24] In addition to varying the choice of chemical crosslinkers described above, many variations of this general strategy have been reported. Here, we will discuss these different options to enhance the workflow and focus on the enrichment of RNA types (**Fig. 7b**), fragmentation by enzymes or ions, enrichment of crosslinked fragments (**Fig. 7a**), approaches to increase resolution (**Fig. 7c**), and alternatives to proximity ligation such as template switching (**Fig. 7d**). These variations create a versatile toolbox, where different options can be selected for specific biological applications. Some of the most critical steps and options are discussed as follows.

3.1. Enrichment of RNA subsets and conformations.

Abundance of cellular RNAs spans at least seven orders of magnitude, making it difficult to study low abundance RNAs. In addition, RNA structure conformations can change during the life cycle of RNA biogenesis and function. To ensure sufficient RNA input, enriching subsets of the transcriptome is necessary. Several different approaches can be employed, including rRNA depletion (e.g., SPLASH, LIGR-seq),[61,68] biotinylated antisense oligos targeting specific transcripts (e.g., PARIS, COMRADES),[57,63] or antibody-based immunoprecipitation of protein-bound RNA (e.g., CLASH and hiCLIP) (**Fig. 7b**).[94,95] Subcellular fractionation using gradient centrifugation or the recently developed APEX proximity labeling offers another approach to enrich RNAs through different stages of their biogenesis or different subcellular localization.[96]

Each approach has its own advantages and problems. rRNA depletion or oligo(dT) enrichment of mRNAs may not be sufficient to isolate low-abundance RNAs. Biotinylated antisense oligos are expensive, making it hard to scale up the experiments. Antibody enrichment of RNP complexes results in highly specific RNA conformations but depends on prior knowledge of the complex composition and the availability of high-quality antibodies. Centrifugation based fractionation is very crude and highly variable, and often does not achieve sufficient specificity and purity. APEX-based

enrichment introduces another irreversible chemical modification step that impedes reverse transcription (see further discussion below), reducing the efficiency of detecting proximally ligated RNA fragments.

3.2. Enrichment of crosslinked fragments.

Given the low efficiency of most crosslinking agents, many RNA fragments do not carry structural information and will ideally be removed. Several methods have been developed to enrich crosslinked fragments only, including 2-dimensional electrophoresis (2D gels), exonuclease digestion, and streptavidin selection of biotinylated crosslinkers. The 2D gel method, including both native-denatured 2D and denatured-denatured 2D, separates crosslinked fragments from non-crosslinked based on slower migration of its extended "X"-shape, which in theory provides 100% purity (**Fig. 7a**).[57] The 2D gel method is applicable to any chemical crosslinker, such as psoralens, nitrogen mustards and bifunctional acylators, since the separation only depends on RNA geometry.[57] However, 2D gels are very laborious and difficult to scale up. Exonuclease digestion of non-crosslinked fragments is easier to perform, but also enriches non-crosslinked RNAs that are highly structured, have chemical monoadducts, or are crosslinked to proteins.[61] Biotinylated crosslinkers allow facile enrichment of reacted RNAs using the biotin-streptavidin system (see examples in psoralens and nitrogen mustards in **Figs. 3 and 5**). However, it not only enriches crosslinked fragments, but also monoadducts, which are likely more abundant than the crosslinked ones, leading to high background.

3.3. Exonuclease trimming to improve resolution.

Fragmented RNAs range in length from a few to several hundred nucleotides. While selection of shorter fragments increases the resolution of structure modeling, it only helps duplex modeling, where base pairing rules can be used to build the secondary structure model. Crosslinking of fragments that form tertiary contacts or spatial proximity does not provide sufficient resolution for structure modeling. To resolve this issue, two approaches have been developed. Exonucleases can trim off nucleotides from either end (e.g., RNase R for the 3' end), until blocked by the crosslink, leaving a tail of a defined length (**Fig. 7c**). Such tails can be used to determine the crosslinking sites precisely. However, monoadducts may also block the exonuclease, leading to incomplete trimming, and reducing precision. As an alternative to proximity ligation, sequence information from the two fragments can also be joined by template switching during reverse transcription (RT), especially with engineered reverse transcriptases that can jump across crosslinking junctions at higher frequencies (**Fig. 7d**).[41] This approach also improves the resolution in defining the crosslinked sites. However, this method has only been tested in simple RNA samples in vitro, and the efficiency is even lower than proximity ligation. Given that monoadducts affect both exonuclease trimming and template switching, further development of crosslinkers that produce minimal monoadducts is needed to improve resolution in the identification of crosslinking sites.

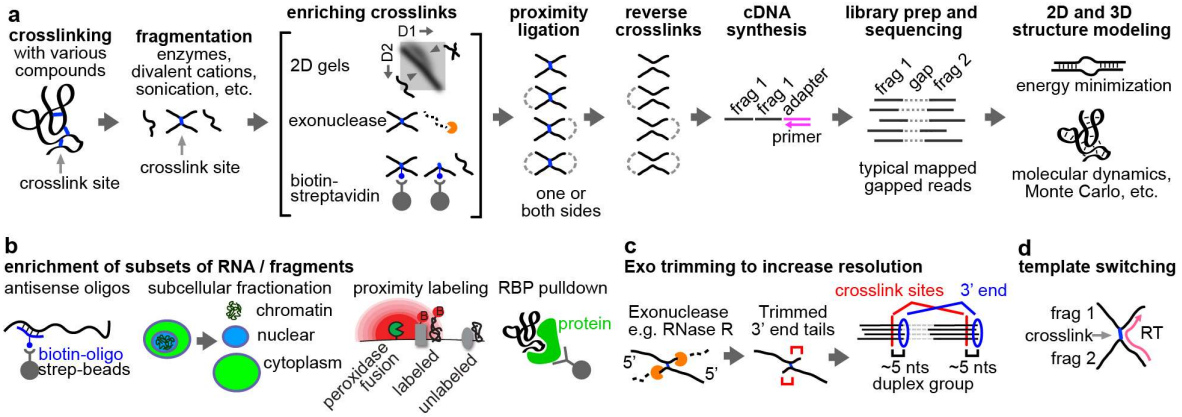

***Figure 7**. The general workflow of crosslink-ligation methods. (a) Basic pipeline. RNAs are crosslinked, fragmented, and enriched for the crosslinked fragments using various methods such as 2D gels, exonucleases, or streptavidin-beads for biotinylated crosslinkers. After proximity ligation and reversal of crosslinks, adapters are ligated to the fragments for cDNA synthesis, which is followed by cDNA amplification and high throughput sequencing. The gapped reads provide spatial constraints for 2D and 3D structure modeling. (b) Specific subsets of RNAs can be enriched for structure analysis after crosslinking and before fragmentation, using various approaches, e.g., antisense oligos, subcellular fractionation, proximity labeling (biotinylation by the APEX system) and antibody pull down of specific RNA binding proteins. (c) Exonuclease trimming can be used to improve resolution. Exonucleases are blocked by monoadducts or crosslinks, leaving a stub of fixed length, and the crosslink sites can be deduced by counting backwards from the 3' ends. (d) Template switching is an alternative of proximity ligation to capture the two fragments in a single read, based on the ability of the reverse transcriptase (RT) to switch templates as it encounters roadblocks.*

3.4. Proximity ligation.

Proximity ligation is the most widely used approach to join two crosslinked fragments. The ligation reaction can occur on either end, leading to different types of sequenced reads. Ligation of both ends leads to circular products that can no longer be ligated to adapters and are therefore lost during library preparation. However, given the low ligation efficiency, such double ligation events are rare and can be mostly ignored. The covalent linkage between the two fragments dramatically increases the likelihood of ligating them, compared to non-crosslinked ones in solution; however, the crosslinked stable structures and shortness of the fragments create steric hindrance, limiting ligation efficiency to typically below 15%. While longer fragments may be subject to lower steric hindrance and more efficiently ligate, they reduce the resolution of structure analysis. Several different protocols have been used in the last 10 years, including direct enzymatic ligation by T4 RNA Rnl1, Rnl2, CircLigase, Mth Rnl, RtcB and the indirect ligation by incorporation of linkers, such as pCp-biotin and short oligos.[57,61,94,97] Chemical proximity ligation, using phosphate activators, such as bromide cyanide, has also been tested.[57] Given that each protocol has several other steps that affect the perceived ligation efficiency, such as the purity of the crosslinked fragments (see the section above on enrichment of crosslinked fragments) and RNA damage levels, none of the improvements have been convincingly demonstrated to outperform others in side-by-side comparisons. New chemical and enzymatic approaches that overcome damages of RNA and steric hindrance of short fragments are needed to improve the ligation efficiency.

3.5. Reverse transcription (RT).

Converting RNA to cDNA fragments is straight forward, but is impacted by side reactions of crosslinking, which reduce efficiency and accuracy of detecting the crosslinking events. Although more resistant to UV damage than DNA,[98] UV-induced lesions in RNA are commonly observed and include base oxidation (primarily 8-oxoguanine),[99] pyrimidine dimers,[99,100] and adducts with proteins[101,102] and other cellular components (**Fig. 8**). Nitrogen mustards are highly reactive alkylating agents that can crosslink nucleic acids to proteins (**Fig. 8**).[103] While not reported yet, it is likely that bifunctional acylators can crosslink RNA to nucleophilic protein residues as well (**Fig. 8**). Both lesions can potentially interfere with reverse transcription.

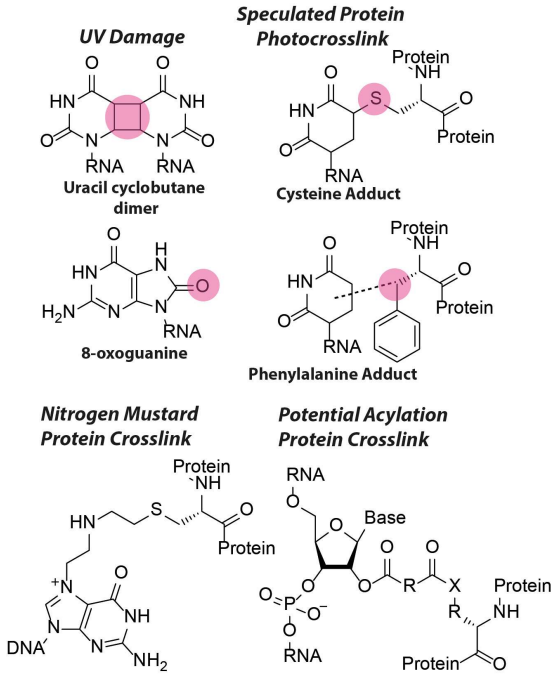

***Figure 8**. Molecular structure of side-products that can potentially interfere with reverse transcription. Multiple UV damage lesions have been reported with the uracil cyclobutane dimer and 8-oxoguanine as the predominant ones. Pyrimidines can photocrosslink to proteins, with cysteine and phenylalanine adducts as speculated products. Nitrogen mustards can crosslink nucleic acids to cysteine residues. It is expected that bifunctional acylators can crosslink RNA to nucleophilic residues (X) on proteins.*

Several crosslinkers, such as nitrogen mustards, and formaldehyde suffer from poor reversibility, reducing the efficiency of cDNA synthesis and decreasing the perceived percentage of gapped reads if the RT is blocked before the ligation junctions. Damaged sites from 254nm UV, psoralen and formaldehyde also lead to mutations and short 1-2 nt deletions, which are artifacts that confound the analysis of gapped reads.[57] Improved RT conditions, such as the use of Mn^{2+} ions, and extended RT time, have been reported, but do not address all the different types of damages.[57,104] In addition to developing better chemical crosslinkers and conditions to minimize damages, direct RNA sequencing may also help overcome some of these problems.[105]

4. Analysis of chemical crosslinking data and structure modeling

Despite the seemingly straight-forward gapped reads from crosslink-ligation experiments, analysis of such data can be challenging. Numerous computational approaches have been described in the past few years. Here, we describe a unified conceptual framework and some of the critical considerations in the data analysis for various types of chemical crosslinking methods, focusing on two major steps: (1) extraction of information from reads including classification and clustering, and (2) inference of structural conformations, including 2D and 3D structure modeling. The first step processes the data, without the need for any prior knowledge or assumptions about the structure, whereas the second step models the structures based on principles of RNA folding.

4.1. Classification read types from crosslink-ligation experiments.

In theory, crosslinking and ligation can occur on any RNA in physical proximity (as long as the chemistry of the crosslinker permits). Therefore, such experiments can join many different types of arrangements of RNA fragments into a single read. The chemical reactivities of the crosslinkers determine the types of structural information obtained (**Fig. 9a**). For example, psoralens can only crosslink base paired regions, while aldehydes, nitrogen mustards and bifunctional acylators can crosslink any spatially proximal nucleotides, either constrained by nearby helices, tertiary contacts, or proteins. In addition, the ability of many RNA crosslinkers to react with proteins further increases the chances of capturing protein-mediated structures (**Fig. 9a**). Psoralens can react with proteins, even though at much lower efficiency than nucleic acids, and lower compared to other crosslinkers.[106] Proteinase treatment after crosslinking can exclude the protein-mediated proximities.[40,57] Careful comparison of data from different types of crosslinkers will provide deeper insights into the organization of the RNP complexes.

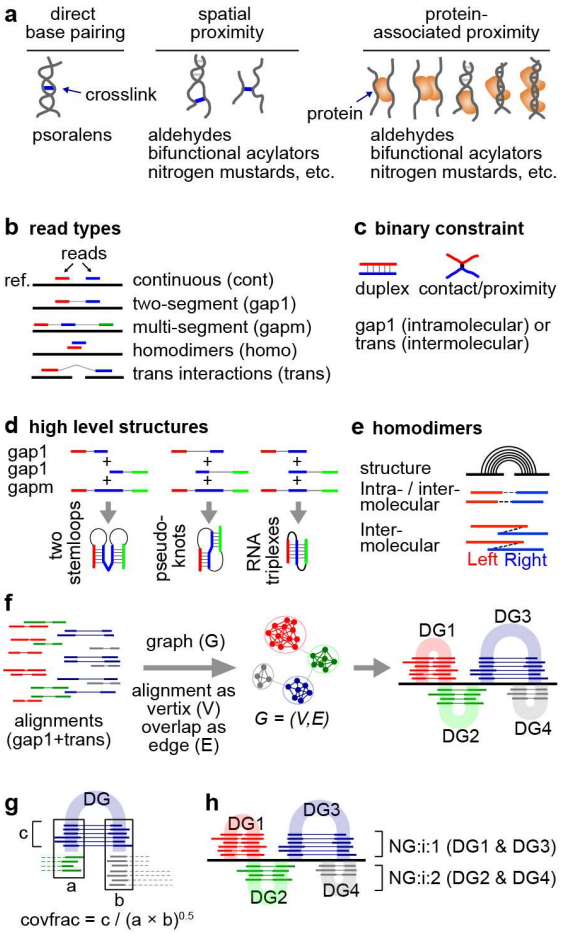

***Figure 9**. Strategies for extracting RNA structures and interactions from crosslink-ligation data. (a) Various types of crosslinked conformations by different chemical crosslinkers. (b) Classification and rearrangement of reads to 5 major categories, mapped to a reference genome (ref.). (c) Two basic types of binary constraints: basepaired duplex or contact/proximity. (d) Detection of high level structures by combining gap1 and gapm reads. (e) Detection of homodimers using reads with overlapped arms. (f) Clustering of gapped reads into duplex groups (DGs) using graph theory (network techniques). (g) Calculation of relative structure strength using the coverage fraction (covfrac) method. (h) Clustering and visualization of DGs into non-overlapping groups (NGs).*

Sequencing reads from crosslink experiments can be aligned to the reference using various types of mappers, such as STAR and Bowtie.[107] Classification of these reads reveals the fragment arrangements and therefore the underlying structural conformations. Recent exhaustive classification results in 5 major read types (**Fig. 9b**): continuous or non-gapped reads that are due to failed crosslinking and/or ligation (cont), two-segment or single-gapped reads due to single ligation events either within one RNA (gap1) or between two RNA molecules (trans), multi-segment (>2 segments) or multi-gapped (>1 gap) reads due to simultaneous multi-crosslink and multi-ligation events on several RNA fragments in spatial proximity (gapm), and overlapping fragments that come from RNA homodimers (homo).[57,108,109]

Reads with a single gap (gap1 and trans) suggest a single constraint in the structure, either a double helix crosslinked by psoralens, or tertiary contact / proximity crosslinked by SHARC or TBIA and other types of crosslinkers (**Fig. 9c**). Reads with more than one gap suggest more complex structures, such as consecutive stemloops, pseudoknots and triplexes, depending on the relative location of the two stem regions (**Fig. 9d**). Co-assembly of the gap1 and gapm reads provide strong evidence for these more complex structures. The percentages of gapm reads are very low in current methods due to the low crosslinking and proximity ligation efficiencies, limiting the discovery of complex conformations. A small percentage of reads have two segments overlapping each other. This is not possible for an intramolecular duplex since the fragmentation should only remove parts of the RNA. The only explanation is that the two fragments came from two identical RNA molecules, in other words, RNA homodimers (**Fig. 9e**).[108] Additional types that are combinations of these 5 types are also possible and indicate even more complex structural conformations.

4.2. Clustering of reads into groups that represent specific structures.

High-throughput sequencing produces very dense gapped reads that come from various structural units and conformations in the transcriptome, making it difficult to interpret the data. Based on the assumption that each structure unit produces a set of reads that are highly similar to each other, the reads can be clustered, where each group can be used to infer a specific contact (or closely positioned contacts, **Fig. 9f**). This approach was first proposed in 2016 and has been systematically benchmarked, optimized and widely applied to various experimental methods.[33,40,60,97,108] The clustering of multiple different DGs provides direct evidence for the existence of alternative or dynamic conformations. Such conformations are easily detected for psoralen crosslinked RNA fragments, given the mutually exclusive base pairing (except the triplexes and G-quadruplexes where each stretch of nucleotides can participate in multiple consecutive contacts). Other clustering approaches have been developed simply based on overlap of arms among reads.[110]

The clustering approach provides a statistical assessment of the underlying structure (**Fig. 9g**) and a method for visualization (**Fig. 9h**). RNA structural conformations exist in variable frequencies, and are crosslinked at different rates, resulting in read groups with a wide range of abundances. To quantify the abundance, coverage fraction can be used, where the read number in each group is normalized by the total coverage of reads across the two arms (**Fig. 9g**). Several different types of statistical tests of the significance for the structure formation can be used to rank and filter data, such as binomial and Fisher's exact test, which typically assess the significance of the ligation event against expression levels of the RNAs.[110,111] For efficient visualization, the DGs are further arranged into non-overlapping groups (NGs), enabling tight packing of the reads in genome browsers (e.g., IGV,[112] **Fig. 9h**). This visualization of the read groups does not make assumptions about the underlying experimental methods and does not require knowledge-based inference of structures, and thus is generally applicable to all types of crosslinking data and free from modeling biases. The initial implementation considers only RNA double helices based on gap1 reads, hence the name duplex group (DG). However, the basic principle has been extended to include multi-gapped reads (gapm), homodimers (homo), which are technically the same as typical hetero-duplexes, and binary non-basepairing spatial proximity or contacts (e.g., from other crosslinkers beyond psoralens).

4.3. Structure modeling assisted by experimental constraints.

Protein-mediated proximities and in situ proximity ligations can join any RNA fragments as long as they are close to each other in space, but does not require direct contact, making it difficult to determine the RNA conformations (**Fig. 9a**). Direct RNA crosslinking, such as by psoralen and bifunctional acylation reagents (e.g., SHARC and TBIA), especially after proteinase treatment to remove proteins, captures direct secondary and tertiary contacts/proximity which are more useful for structure modeling (**Fig. 9a**). At a higher level, many recent studies have revealed modular domains in RNA architectures, defined by frequent crosslinks within each domain and sparse crosslinks between domains (**Fig. 10a-b**).[60,63,113–115] The large domains are linked by more flexible and dynamic regions. The discovery of these modular domains based on many different types of chemical crosslinkers (e.g., psoralen and formaldehyde) demonstrated their authenticity. At a lower level, the inference of individual secondary and tertiary structure units from crosslinking data remains an unsolved problem, for several reasons. First, the low crosslink-ligation efficiency and strong bias led to incomplete data for modeling (e.g., psoralen bias towards uridines, and acylation bias towards unconstrained riboses). Second, conventional computational modeling tools were not optimized for the incorporation of crosslink-ligation data, so it remains unclear how to make best use of experimental constraints. Third, current modeling tools can only be applied to short RNAs due to the prohibitive computational cost. Despite these challenges, several general approaches have been implemented.

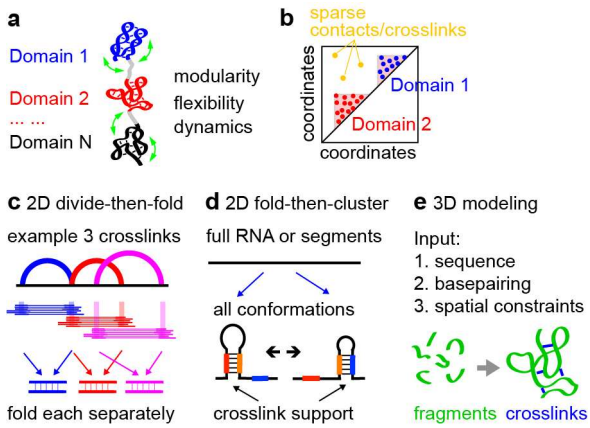

***Figure 10**. Modeling of structures from crosslinking constraints. (a) Conceptual diagram of high-level modular RNA domains that are flexible and dynamic. (b) Clustering of contacts define RNA domains. (c) Divide and then fold approach for secondary structure modeling, exemplified by 3 crosslinks. (d) Fold and then cluster approach for secondary structure modeling, for 2 example alternative conformations. (e) Typical fragment assembly-based approach for 3D modeling, where known 3D fragments are assembled and constrained by crosslinking data (e.g., Rosetta).*

2D structures.

Two different strategies are used for 2D structure modeling based on experimental constraints: divide-then-fold (**Fig. 10c**), and fold-then-cluster (**Fig. 10d**). In the divide-then-fold approach, the gapped reads (and the assembled DGs) carve out pairs of RNA fragments, and minimal free energy structures can be built from the two fragments (**Fig. 10c**).[60] Whole transcript structures can then be established by combining all the structure models. This method is fast, since only crosslinked regions are modeled, however, the resulting global conformations cannot be clearly deconvolved. Typically, one structure map is produced for each RNA as a result, including all the predicted helices, some of which may be mutually exclusive (**Fig. 10c**, blue and red arcs with overlapping arms). Alternatively, in the fold-then-cluster approach, RNA molecules can be folded de novo, producing an ensemble of global conformations, which are then clustered (**Fig. 10d**).[63,116] The gapped reads (and the assembled DGs) are then mapped onto the clustered conformations. This approach outputs complete global conformations, however, it also suffers from several caveats. The structures may be over-folded, and not all duplexes in such models are necessarily supported by experimental constraints. The folding step may not be feasible for long cellular and viral RNAs, many of which measure tens to hundreds of thousands of nucleotides. As a result, folding is often performed in windows of limited lengths (e.g., 300-1000 nts), which excludes long-range structures. In the future, integration of these two approaches is needed to increase the efficiency and accuracy of 2D structure modeling. To quickly validate and rank the biological relevance of these secondary structure models, multiple sequence alignments and identification of conserved and covaried base pairs are often performed, and compared to experimental constraints.[60,114]

3D structures.

Computational modeling of 3D RNA structures is a rapidly progressing field, and many different approaches have been proposed, such as molecular dynamics, fragment assembly, and deep learning (**Fig. 10e**).[14] In addition to the primary sequence and secondary structure models, experimental constraints are typically incorporated into the modeling as pseudo-energy terms: predicted conformations are penalized where the constraint is not satisfied. Compared to 2D modeling, the basic rules in 3D structure modeling have not been thoroughly studied,[117] and therefore, the modeling results are typically confounded by multiple sources of error, both computational and experimental. Despite that experimental constraints can be obtained for RNAs of any length, modeling of large RNAs remain impractical due to the high computational cost.

5. Biological applications: critical considerations and recent examples.

Crosslink-ligation methods have a wide range of applications in RNA biology and medicine. The earliest applications of these chemical tools were critical in establishing the structural mechanisms of pre-mRNA splicing and rRNA processing. Here we discuss critical considerations in applying these methods, and review some of the most interesting examples in recent years, focusing on lncRNAs, ncRNA networks, genetic disorders and RNA viruses, which will help researchers to properly apply them to their biological problems.

5.1. Structure-guided studies of RNA functions: choosing proper methods and important biological targets.

Recent development of sequencing-based strategies allowed deeper interrogation of transcriptome-wide RNA folding principles, lower abundance RNAs and low frequency alternative/dynamics conformations. While general properties of RNA folding can be extracted from transcriptome-wide measurements, rigorous functional studies of the newly

discovered RNA structures and interactions have not caught up with the rapidly growing large lists of new data from high-throughput sequencing. The choice of chemical crosslinkers should be based on the needs of specific biological questions. For example, detection of base pairing mediated structures and interactions require psoralens, whereas general spatial proximity can be studied using other crosslinkers (**Fig. 9a**). Even for chemical crosslinkers that detects general spatial proximity, proteinase K treatment is necessary to exclude protein-mediated conformations (**Fig. 9a**). Multiple different types of crosslinkers may be needed to obtain more details regarding complex RNP complexes, such as 2D structures, 3D structures and RNA-protein interactions.[118,119]

To pursue discovery-driven studies of RNA structures, several important aspects need to be considered even if the sequencing experiment was designed to test a specific hypothesis as opposed to an unbiased discovery. First, analysis of the relative coverage and statistical significance is needed to predict the biological significance of the newly discovered structures or interactions, as described above (**Fig. 9g**).[110,111] More abundant and reproducible structures are more likely to contribute to functions. Second, validation by various other methods, especially structure conservation/covariation increases the confidence in the functional significance.[120] Third, direct connection of the structural elements to functional sequence motifs (e.g., in splicing regulation), protein binding sites, RNA modification sites and processing patterns, etc. indicates potential functions in these aspects. Last, linkage of the structures to human diseases is a strong motivation for further mechanistic analysis. These principles have been employed in several recent studies. After discovery of potentially important RNA structures, care should be taken in studies of their function, since functions for one specific DNA and its RNA products can be encoded on multiple levels, such as DNA sequence, RNA sequences, RNA-protein, or RNA-RNA interactions, as well as RNA structures. To demonstrate that functional consequences are indeed due to structure formation, multiple types of mutation and compensatory rescue tests are needed.

5.2. Solving in vivo RNA structures and interactions to understand RNA functions: recent examples.

Recent applications of the crosslink-ligation methods have been focused on several directions, including RNA virus genome structures and host-virus interactions,[63,121–123] noncoding RNA interactions[108] and lncRNA functions.[60,113,124]

Single-stranded RNA viruses are a major target of the crosslink-ligation methods, because of the strong dependence of the viruses on their genome structures and RNA-mediated host-virus interactions, and the therapeutic potential of targeting viral RNA genomes. Recent studies have reported genome structures and host-virus interactions for *Flaviviruses*, such as Zika and Dengue, *Coronaviruses*, such as SARS-CoV-2, Influenza, and *Picornaviruses*, such as EV-D68.[57,63,114,121,123] These studies confirmed many previously predicted structures, revealed new structural elements and new miRNA and snoRNA-mediated host-virus interactions that affect virus fitness.

The transcriptomes of most organisms contain a large number of short noncoding RNAs that regulate other RNAs through base pairing. Applications of the crosslink-ligation methods have led to the discovery of interaction networks for a variety of species and include bacterial sRNAs, eukaryotic miRNAs and snoRNAs.[57,63,108,125,126] These studies dramatically expanded the known targets of important ncRNA regulators. Given the low abundance of many ncRNAs and their targets, it is expected that future deeper sequencing, coupled with RNA enrichment methods (**Fig. 7**), will continue to reveal new targets. Application of these methods to other recently characterized ncRNAs, such as tRNA fragments and rRNA fragments, will help decipher their functions, most of which remain unknown up to now.

LncRNAs are a diverse class of regulatory RNAs with broad roles in gene expression control and other cellular processes. Crosslink-ligation methods have led to new insights into their mechanisms of action, including the discovery of modular domains that organize their multiple functions and spatial separation of these functions.[60,113,124,127] Such studies provide a proof-of-principle for future analysis of other large RNAs, including mRNA UTRs and introns, whose primary functions are noncoding and thus may use similar mechanisms.

6. Conclusions and perspectives.

In the past 10-20 years, extensive technological improvements of 1D chemical probing methods have led to their widespread applications in various biological systems. Despite the rapid progress in the field of RNA crosslinking in recent years, there remain several technical challenges in the study of RNA structures and interactions in living cells. Crosslinking-based methods are difficult to implement, due to the inefficiencies at multiple steps, and complexities of computational analysis, creating a major barrier for its wide adoption. Given that most of these methods were only recently developed, evaluation of the computational tools has not been performed. More efficient tools will likely be needed to tackle challenging biological problems, including RNA structural dynamics and heterogeneity in space and time, such as embryonic development and disease, large RNAs and their complexes with other RNAs and proteins, and high resolution structure modeling. At the same time, systematic side-by-side comparison of different methods will help clarify their strengths and weaknesses, increase the efficiency of these methods and make it possible for more labs to implement them.[57] Here we provide some perspectives on further technology development in this field.

6.1. New RNA crosslinking chemistry: faster, less bias, more precise and more efficient.

Even though a wide variety of physical and chemical crosslinkers are available now, there are still several unsolved problems, including the crosslinking efficiency, side reactions, bias, and reaction kinetics. In vitro tests have shown that SHARC reagents are the most efficient, but they are not reactive towards tightly packed structures in cells.[40] In other words, they are biased towards single stranded and un-constrained nucleotides. All current crosslinkers, except direct

UV, are relatively slow and typically require at least 10 mins to achieve sufficient efficiency for subsequent sequencing experiments. Structural dynamics in cellular RNAs can span 18 orders of magnitude, down to picoseconds.[128] While the ultrafast dynamics can only be studied using physical methods in vitro, some of the structural transition dynamics that range from seconds to minutes are potentially tractable using faster chemical crosslinkers. Several new chemical tools have been reported recently to improve the temporal resolution in 1D chemical probing and one-sided crosslinking, such as nicotinoyl azide and cyanovinyl carbazoles, offering interesting new ideas for developing new RNA crosslinkers.[129] Achieving both fast kinetics and reversibility in the same bifunctional crosslinker is still challenging.

Another important shortcoming that potentially can be addressed with improved chemistry is the existence of monoadducts that complicate analysis. Recent advances in the field of self-immolative linkers[130,131] may provide inspiration for chemical crosslinkers that display significantly fewer monoadducts.

6.2. Long-read sequencing to capture whole-transcript structures.

Current ligation and template switching based methods can only capture small structure units, such as duplexes, even though they are not constrained by sequence length. Simultaneous crosslinking and ligation have enabled the discovery of combinations of the duplexes, including consecutive stemloops, pseudoknots and triplexes (**Fig. 9d**).[108] Given the large number of possible conformations for each RNA species, it is currently impossible to stitch together the structural units into complete models for whole transcripts. Long-read sequencing has been used to capture potentially full-length or longer conformations from foot-printing based chemical probing methods.[119] Dramatically improved crosslinking and ligation efficiency may produce longer RNA fragment combinations, which together with long-read sequencing, can link co-occurring and distinguish mutually exclusive conformations.

6.3. Single-cell analysis of RNA structure heterogeneity.

Alternative RNA structures are pervasive in cells, and likely contribute to cell-type-specific regulation of gene expression. Indeed, such examples have been demonstrated in RNA-structure-dependent alternative splicing of cell adhesion molecules that determine neuronal connectivity, such as insect *Dscam* and mammalian *Neurexins*.[132] Even though splicing-regulatory intronic structures have been discovered, little is known about their cell type specificity and whether they correlate with the splicing outcome in individual cells. Development of single cell structure mapping methods will be necessary to link the structural conformations to the functional outcomes. Despite the success of single cell RNA-seq, structure analysis is more challenging. First, the low sequencing coverage in individual cells make it hard to get sufficient reads to build structure models. Targeted enrichment of RNA species will be necessary to obtain sufficient data. Second, incorporation of the crosslink-ligation protocol into single cell sequencing workflows is not trivial. Some of the critical steps in the crosslink-ligation methods, such as enrichment of crosslinked fragments (2D gels and biotinylated crosslinkers) have been only implemented in bulk samples.

6.4. Computational integration of various structure mapping methods.

The crosslink-ligation data provide constraints for global organization of RNA structures, but lack sufficient resolution on the details. Even though the exo trimming and template-switching methods can pinpoint some of the crosslinking sites (PARIS-exo, SHARC-exo and SHAPE-JuMP),[40,41] not every base pair can be identified with high confidence. On the other hand, 1D chemical probing data, such as from SHAPE and DMS-seq, provides detailed information on the structure/interaction constraints on individual nucleotides, despite the lack of confidence in their discovery of complex and long-range structures. New computational tools that carefully evaluates data from various experimental methods are needed to effectively take advantage of each method and avoid artifacts. With new computational modeling algorithms, 1D accessibility, 2D base pairing and 3D spatial proximity constraints should enable better modeling of RNA structures.

6.5. High-resolution and high-level structure modeling.

The long-term goal of developing crosslink-ligation and other experimental tools is to build high resolution and comprehensive structure models for RNA, which will hopefully provide new insights into their functions in vivo. While many computational tools have been developed to build 2D and 3D structures, such tools are often developed based on energy terms and parameters derived from a small number of in vitro models. Furthermore, RNA structural dynamics is not only determined by its primary sequence, but also cellular environments, such as proteins, other RNAs, ions, metabolites, temperature, and pH, most of which cannot be incorporated into the modeling studies.

Earlier studies have made use of in vitro constraints from correlated chemical probing to improve computational modeling.[133] Now, the crosslinking-derived in vivo constraints offer a unique opportunity to model 2D/3D RNA that represent true in vivo conformations. On the experimental side, distance measurements are not always accurate. Exo trimming is effective but is can be problematic due to monoadducts blocking the exonuclease. The template switching strategy in SHAPE-Jump is another approach but suffers from high noise, due to the low efficiency and randomness of template switching.[41] Even though chemical crosslinking indicates a probable distance between two nucleotides, the range of the distance distribution can be quite large due to the natural dynamics of the RNA in cells. Furthermore, very few RNA molecules exist in their naked form; most RNAs are bound by a wide variety of proteins. Ignoring the proteins in the modeling is bound to introduce errors. On the computational side, current implementation of such approaches are still rudimentary, and only produce very simple models.[40,41] New methods that

consider the in vivo flexibility and dynamics during modeling are needed to produce ensembles of structure conformations that are more biologically relevant. Conventional 3D modeling tools are extremely computationally expensive, currently only able to handle RNAs within 200-300 nucleotides. Faster new algorithms are needed for the vast majority of the large RNA molecules in cells, which often span hundreds to tens of thousands of nucleotides.

6.6. Discovery of in vivo RNA structures and interactions for targeted therapeutics.
Normal and abnormal functions of RNA molecules have been implicated in many human diseases, such as nucleotide repeat disorders, RNA virus infections, and splicing mutations, which account for at least a third of all disease mutations. The past decade has witnessed a revolution in the field of RNA therapeutics, with the approval of multiple small molecule and antisense oligo drugs that target a wide variety of RNA sequence and structural elements with critical roles in human diseases.[134] Many more RNA-targeting therapeutics are either in pre-clinical studies or clinical trials. Structural studies have played critical roles in some of these RNA therapeutics and we expect that the increasing availability of RNA crosslinking experiments will further assist towards this goal in the future. For example, the discovery of intronic sequence and structure elements in the SMN2 gene led to the development of antisense oligo drugs to treat spinal muscular atrophy, one of the most prevalent pediatric genetic disorders.[135]

In summary, we hope that our critical review of chemical crosslinking methods, including the chemistry, enzymology, computational analysis, and biological applications, will spur further improvements of these methods and applications in chemistry, biology and therapeutics. Improved structural information will bring conventional sequence motif-based studies of RNA regulation to higher dimensions in the future. Better structure models obtained with enhanced chemical crosslinkers will provide essential guidance for the development of RNA-based and RNA-targeting drugs.

Competing interests
W.A.V. and Z.L. are named inventors on a patent application describing the SHARC technology.

Author Contributions
W.A.V. and Z.L. contributed equally.

Acknowledgements
This work has received funding from the European Research Council under the European Union's Horizon Europe research and innovation programme under grant agreement number 101041938 (RIBOCHEM) to W.A.V.. Z.L. is supported by startup funds from the University of Southern California, School of Pharmacy, the Pathway to Independence Award from NHGRI (R00HG009662), NIGMS (R35GM143068), NIA (R21AG075665). USC Research Center for Liver Disease (P30DK48522), Illumina and USC Keck Genomics Platform Core Lab Partnership Program, the Norris Comprehensive Cancer Center (P30CA014089), USC ALPD and Cirrhosis Research Center (P50AA011999) and USC Center for Advanced Research Computing.

References
(1) Statello, L.; Guo, C.-J.; Chen, L.-L.; Huarte, M. Gene Regulation by Long Non-Coding RNAs and Its Biological Functions. *Nat Rev Mol Cell Biol* **2021**, *22* (2), 96–118. https://doi.org/10.1038/s41580-020-00315-9.
(2) Cech, T. R.; Steitz, J. A. The Noncoding RNA Revolution - Trashing Old Rules to Forge New Ones. *Cell* **2014**, *157* (1), 77–94. https://doi.org/10.1016/j.cell.2014.03.008.
(3) Vicens, Q.; Kieft, J. S. Thoughts on How to Think (and Talk) about RNA Structure. *Proceedings of the National Academy of Sciences* **2022**, *119* (17), e2112677119. https://doi.org/10.1073/pnas.2112677119.
(4) Yao, R.-W.; Wang, Y.; Chen, L.-L. Cellular Functions of Long Noncoding RNAs. *Nat Cell Biol* **2019**, *21* (5), 542–551. https://doi.org/10.1038/s41556-019-0311-8.
(5) Childs-Disney, J. L.; Yang, X.; Gibaut, Q. M. R.; Tong, Y.; Batey, R. T.; Disney, M. D. Targeting RNA Structures with Small Molecules. *Nat Rev Drug Discov* **2022**, *21* (10), 736–762. https://doi.org/10.1038/s41573-022-00521-4.
(6) Falese, J. P.; Donlic, A.; Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles towards the Clinic. *Chem. Soc. Rev.* **2021**, *50* (4), 2224–2243. https://doi.org/10.1039/D0CS01261K.
(7) M. Meyer, S.; C. Williams, C.; Akahori, Y.; Tanaka, T.; Aikawa, H.; Tong, Y.; L. Childs-Disney, J.; D. Disney, M. Small Molecule Recognition of Disease-Relevant RNA Structures. *Chemical Society Reviews* **2020**, *49* (19), 7167–7199. https://doi.org/10.1039/D0CS00560F.
(8) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nature Reviews Drug Discovery* **2018**, *17*, 547.
(9) Crooke, S. T.; Liang, X.-H.; Baker, B. F.; Crooke, R. M. Antisense Technology: A Review. *Journal of Biological Chemistry* **2021**, *296*. https://doi.org/10.1016/j.jbc.2021.100416.
(10) Crooke, S. T.; Witztum, J. L.; Bennett, C. F.; Baker, B. F. RNA-Targeted Therapeutics. *Cell Metab* **2018**, *27* (4), 714–739. https://doi.org/10.1016/j.cmet.2018.03.004.
(11) Lu, Z.; Chang, H. Y. Decoding the RNA Structurome. *Curr Opin Struct Biol* **2016**, *36*, 142–148. https://doi.org/10.1016/j.sbi.2016.01.007.

(12) Lu, Z.; Chang, H. Y. *The RNA Base-Pairing Problem and Base-Pairing Solutions*; Cold Spring Harb Perspect Biol, 2018; Vol. 10. https://doi.org/10.1101/cshperspect.a034926.
(13) Spitale, R. C.; Incarnato, D. Probing the Dynamic RNA Structurome and Its Functions. *Nat Rev Genet* **2022**, 1–19. https://doi.org/10.1038/s41576-022-00546-w.
(14) Zhang, J.; Fei, Y.; Sun, L.; Zhang, Q. C. Advances and Opportunities in RNA Structure Experimental Determination and Computational Modeling. *Nat Methods* **2022**, *19* (10), 1193–1207. https://doi.org/10.1038/s41592-022-01623-y.
(15) Peattie, D. A.; Gilbert, W. Chemical Probes for Higher-Order Structure in RNA. *Proceedings of the National Academy of Sciences* **1980**, *77* (8), 4679–4682. https://doi.org/10.1073/pnas.77.8.4679.
(16) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. RNA SHAPE Chemistry Reveals Nonhierarchical Interactions Dominate Equilibrium Structural Transitions in TRNAAsp Transcripts. *Journal of the American Chemical Society* **2005**, *127* (13), 4659–4667. https://doi.org/10.1021/ja0436749.
(17) Spitale, R. C.; Crisalli, P.; Flynn, R. A.; Torre, E. A.; Kool, E. T.; Chang, H. Y. RNA SHAPE Analysis in Living Cells. *Nature chemical biology* **2013**, *9*, 18–20. https://doi.org/10.1038/nchembio.1131.
(18) Marinus, T.; Fessler, A. B.; Ogle, C. A.; Incarnato, D. A Novel SHAPE Reagent Enables the Analysis of RNA Structure in Living Cells with Unprecedented Accuracy. *Nucleic Acids Research* **2021**, *49* (6), e34. https://doi.org/10.1093/nar/gkaa1255.
(19) Morandi, E.; van Hemert, M. J.; Incarnato, D. SHAPE-Guided RNA Structure Homology Search and Motif Discovery. *Nat Commun* **2022**, *13* (1), 1722. https://doi.org/10.1038/s41467-022-29398-y.
(20) Mustoe, A. M.; Busan, S.; Rice, G. M.; Hajdin, C. E.; Peterson, B. K.; Ruda, V. M.; Kubica, N.; Nutiu, R.; Baryza, J. L.; Weeks, K. M. Pervasive Regulatory Functions of MRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* **2018**, *173* (1), 181-195.e18. https://doi.org/10.1016/j.cell.2018.02.034.
(21) Spitale, R. C.; Flynn, R. A.; Zhang, Q. C.; Crisalli, P.; Lee, B.; Jung, J.-W.; Kuchelmeister, H. Y.; Batista, P. J.; Torre, E. A.; Kool, E. T.; Chang, H. Y. Structural Imprints in Vivo Decode RNA Regulatory Mechanisms. *Nature* **2015**, *519* (7544), 486–490. https://doi.org/10.1038/nature14263.
(22) Tomezsko, P. J.; Corbin, V. D. A.; Gupta, P.; Swaminathan, H.; Glasgow, M.; Persad, S.; Edwards, M. D.; Mcintosh, L.; Papenfuss, A. T.; Emery, A.; Swanstrom, R.; Zang, T.; Lan, T. C. T.; Bieniasz, P.; Kuritzkes, D. R.; Tsibris, A.; Rouskin, S. Determination of RNA Structural Diversity and Its Role in HIV-1 RNA Splicing. *Nature* **2020**, *582* (7812), 438–442. https://doi.org/10.1038/s41586-020-2253-5.
(23) Morandi, E.; Manfredonia, I.; Simon, L. M.; Anselmi, F.; van Hemert, M. J.; Oliviero, S.; Incarnato, D. Genome-Scale Deconvolution of RNA Structure Ensembles. *Nat Methods* **2021**, *18* (3), 249–252. https://doi.org/10.1038/s41592-021-01075-w.
(24) Olson, S. W.; Turner, A.-M. W.; Arney, J. W.; Saleem, I.; Weidmann, C. A.; Margolis, D. M.; Weeks, K. M.; Mustoe, A. M. Discovery of a Large-Scale, Cell-State-Responsive Allosteric Switch in the 7SK RNA Using DANCE-MaP. *Molecular Cell* **2022**, *82* (9), 1708-1723.e10. https://doi.org/10.1016/j.molcel.2022.02.009.
(25) Hearst, J. E. PSORALEN PHOTOCHEMISTRY. *Annu. Rev. Biophys. Bioeng.* **1981**, *10* (1), 69–86. https://doi.org/10.1146/annurev.bb.10.060181.000441.
(26) Cimino, G. D.; Shi, Y. B.; Hearst, J. E. Wavelength Dependence for the Photoreversal of a Psoralen-DNA Crosslink. *Biochemistry* **1986**, *25* (10), 3013–3020. https://doi.org/10.1021/bi00358a042.
(27) Cimino, G. D.; Gamper, H. B.; Isaacs, S. T.; Hearst, J. E. . Psoralens as Photoactive Probes of Nucleic Acid Structure and Function: Organic Chemistry, Photochemistry, and Biochemistry. *Annual Review of Biochemistry* **1985**, *54* (1), 1151–1193. https://doi.org/10.1146/annurev.bi.54.070185.005443.
(28) Hearst, J. E. Photochemistry of the Psoralens. *Chem. Res. Toxicol.* **1989**, *2* (2), 69–75. https://doi.org/10.1021/tx00008a001.
(29) Harris, M. E.; Christian, E. L. RNA Crosslinking Methods. *Methods Enzymol* **2009**, *468*, 127–146. https://doi.org/10.1016/S0076-6879(09)68007-1.
(30) Calvet, J. P.; Pederson, T. Base-Pairing Interactions between Small Nuclear RNAs and Nuclear RNA Precursors as Revealed by Psoralen Cross-Linking in Vivo. *Cell* **1981**, *26* (3), 363–370. https://doi.org/10.1016/0092-8674(81)90205-1.
(31) Rimoldi, O. J.; Raghu, B.; Nag, M. K.; Eliceiri, G. L. Three New Small Nucleolar RNAs That Are Psoralen Cross-Linked in Vivo to Unique Regions of Pre-RRNA. *Mol Cell Biol* **1993**, *13* (7), 4382–4390. https://doi.org/10.1128/mcb.13.7.4382-4390.1993.
(32) Fayet-Lebaron, E.; Atzorn, V.; Henry, Y.; Kiss, T. 18S RRNA Processing Requires Base Pairings of SnR30 H/ACA SnoRNA to Eukaryote-Specific 18S Sequences. *EMBO J* **2009**, *28* (9), 1260–1270. https://doi.org/10.1038/emboj.2009.79.
(33) Kudla, G.; Wan, Y.; Helwak, A. RNA Conformation Capture by Proximity Ligation. *Annu Rev Genomics Hum Genet* **2020**, *21*, 81–100. https://doi.org/10.1146/annurev-genom-120219-073756.
(34) Wang, D.; Ye, R.; Cai, Z.; Xue, Y. Emerging Roles of RNA-RNA Interactions in Transcriptional Regulation. *Wiley Interdiscip Rev RNA* **2022**, *13* (5), 1712. https://doi.org/10.1002/wrna.1712.
(35) Singh, S.; Shyamal, S.; Panda, A. C. Detecting RNA-RNA interactome. *Wiley Interdiscip Rev RNA* **2022**, 1715. https://doi.org/10.1002/wrna.1715.

(36) Fei, N.; Sauter, B.; Gillingham, D. The PKa of Brønsted Acids Controls Their Reactivity with Diazo Compounds. *Chem. Commun.* **2016**, *52* (47), 7501–7504. https://doi.org/10.1039/C6CC03561B.
(37) Shirakami, N.; Higashi, S. L.; Kawaki, Y.; Kitamura, Y.; Shibata, A.; Ikeda, M. Construction of a Reduction-Responsive Oligonucleotide via a Post-Modification Approach Utilizing 4-Nitrophenyl Diazomethane. *Polym J* **2021**, *53* (6), 741–746. https://doi.org/10.1038/s41428-021-00464-4.
(38) Ando, H.; Furuta, T.; Tsien, R. Y.; Okamoto, H. Photo-Mediated Gene Activation Using Caged RNA/DNA in Zebrafish Embryos. *Nat Genet* **2001**, *28* (4), 317–325. https://doi.org/10.1038/ng583.
(39) Velema, W. A.; Park, H. S.; Kadina, A.; Orbai, L.; Kool, E. T. Trapping Transient RNA Complexes by Chemically Reversible Acylation. *Angewandte Chemie International Edition* **2020**, *59* (49), 22017–22022. https://doi.org/10.1002/anie.202010861.
(40) Damme, R.; Li, K.; Zhang, M.; Bai, J.; Lee, W. H.; Yesselman, J. D.; Lu, Z.; Velema, W. A. Chemical reversible crosslinking enables measurement of RNA 3D distances and alternative conformations in cells. *Nat Commun* **2022**, *13* (1), 911. https://doi.org/10.1038/s41467-022-28602-3.
(41) Christy, T. W.; Giannetti, C. A.; Houlihan, G.; Smola, M. J.; Rice, G. M.; Wang, J.; Dokholyan, N. V.; Laederach, A.; Holliger, P.; Weeks, K. M. Direct Mapping of Higher-Order RNA Interactions by SHAPE-JuMP. *Biochemistry* **2021**, *60* (25), 1971–1982. https://doi.org/10.1021/acs.biochem.1c00270.
(42) Jash, B.; Kool, E. T. Conjugation of RNA via 2′-OH Acylation: Mechanisms Determining Nucleotide Reactivity. *Chem. Commun.* **2022**, *58* (22), 3693–3696. https://doi.org/10.1039/D2CC00660J.
(43) Velema, W. A.; Kool, E. T. The Chemistry and Applications of RNA 2′-OH Acylation. *Nat Rev Chem* **2020**, *4* (1), 22–37. https://doi.org/10.1038/s41570-019-0147-6.
(44) Nakamura, S.; Ishino, K.; Fujimoto, K. Photochemical RNA Editing of C to U by Using Ultrafast Reversible RNA Photo-Crosslinking in DNA/RNA Duplexes. *Chembiochem* **2020**, *21* (21), 3067–3070. https://doi.org/10.1002/cbic.202000269.
(45) Karmakar, S.; Harcourt, E. M.; Hewings, D. S.; Scherer, F.; Lovejoy, A. F.; Kurtz, D. M.; Ehrenschwender, T.; Barandun, L. J.; Roost, C.; Alizadeh, A. A.; Kool, E. T. Organocatalytic Removal of Formaldehyde Adducts from RNA and DNA Bases. *Nature Chem* **2015**, *7* (9), 752–758. https://doi.org/10.1038/nchem.2307.
(46) Polavarapu, A.; Stillabower, J. A.; Stubblefield, S. G. W.; Taylor, W. M.; Baik, M.-H. The Mechanism of Guanine Alkylation by Nitrogen Mustards: A Computational Study. *J. Org. Chem.* **2012**, *77* (14), 5914–5921. https://doi.org/10.1021/jo300351g.
(47) Baik, M.-H.; Friesner, R. A.; Lippard, S. J. Theoretical Study of Cisplatin Binding to Purine Bases: Why Does Cisplatin Prefer Guanine over Adenine? *J. Am. Chem. Soc.* **2003**, *125* (46), 14082–14092. https://doi.org/10.1021/ja036960d.
(48) Krokhotin, A.; Mustoe, A. M.; Weeks, K. M.; Dokholyan, N. V. Direct Identification of Base-Paired RNA Nucleotides by Correlated Chemical Probing. *RNA* **2017**, *23* (1), 6–13. https://doi.org/10.1261/rna.058586.116.
(49) Zehnder, J.; Van Atta, R.; Jones, C.; Sussman, H.; Wood, M. Cross-Linking Hybridization Assay for Direct Detection of Factor V Leiden Mutation. *Clinical Chemistry* **1997**, *43* (9), 1703–1708. https://doi.org/10.1093/clinchem/43.9.1703.
(50) Favre, A.; Saintomé, C.; Fourrey, J.-L.; Clivio, P.; Laugâa, P. Thionucleobases as Intrinsic Photoaffinity Probes of Nucleic Acid Structure and Nucleic Acid-Protein Interactions. *Journal of Photochemistry and Photobiology B: Biology* **1998**, *42* (2), 109–124. https://doi.org/10.1016/S1011-1344(97)00116-4.
(51) Qiu, Z.; Lu, L.; Jian, X.; He, C. A Diazirine-Based Nucleoside Analogue for Efficient DNA Interstrand Photocross-Linking. *J. Am. Chem. Soc.* **2008**, *130* (44), 14398–14399. https://doi.org/10.1021/ja805445j.
(52) Chapman, E. G.; DeRose, V. J. Site-Specific Platinum(II) Cross-Linking in a Ribozyme Active Site. *J. Am. Chem. Soc.* **2012**, *134* (1), 256–262. https://doi.org/10.1021/ja206455t.
(53) Elskens, J.; Madder, A. Crosslinker-Modified Nucleic Acid Probes for Improved Target Identification and Biomarker Detection. *RSC Chem. Biol.* **2021**, *2* (2), 410–422. https://doi.org/10.1039/D0CB00236D.
(54) Shim, S.-C.; Jeon, Y. H.; Kim, D.; Han, G.; Yoo, D. J. PHOTOCHEMISTRY AND PHOTOBIOLOGY OF PSORALENS. *Journal of Photoscience* **1995**, *2* (1), 37–45.
(55) Ariza-Mateos, A.; Prieto-Vega, S.; Díaz-Toledano, R.; Birk, A.; Szeto, H.; Mena, I.; Berzal-Herranz, A.; Gómez, J. RNA Self-Cleavage Activated by Ultraviolet Light-Induced Oxidation. *Nucleic Acids Research* **2012**, *40* (4), 1748–1766. https://doi.org/10.1093/nar/gkr822.
(56) Yeung, A. T.; Dinehart, W. J.; Jones, B. K. Alkali Reversal of Psoralen Cross-Link for the Targeted Delivery of Psoralen Monoadduct Lesion. *Biochemistry* **1988**, *27* (17), 6332–6338. https://doi.org/10.1021/bi00417a020.
(57) Zhang, M.; Li, K.; Bai, J.; Velema, W. A.; Yu, C.; van Damme, R.; Lee, W. H.; Corpuz, M. L.; Chen, J.-F.; Lu, Z. Optimized Photochemistry Enables Efficient Analysis of Dynamic RNA Structuromes and Interactomes in Genetic and Infectious Diseases. *Nat Commun* **2021**, *12* (1), 2344. https://doi.org/10.1038/s41467-021-22552-y.
(58) Sutherland, B. M.; Sutherland, J. C. Mechanisms of Inhibition of Pyrimidine Dimer Formation in Deoxyribonucleic Acid by Acridine Dyes. *Biophysical Journal* **1969**, *9* (3), 292–302. https://doi.org/10.1016/S0006-3495(69)86387-3.

(59) Calvet, J. P.; Pederson, T. Heterogeneous Nuclear RNA Double-Stranded Regions Probed in Living HeLa Cells by Crosslinking with the Psoralen Derivative Aminomethyltrioxsalen. *Proceedings of the National Academy of Sciences* **1979**, *76* (2), 755–759. https://doi.org/10.1073/pnas.76.2.755.
(60) Lu, Z.; Zhang, Q. C.; Lee, B.; Flynn, R. A.; Smith, M. A.; Robinson, J. T.; Davidovich, C.; Gooding, A. R.; Goodrich, K. J.; Mattick, J. S.; Mesirov, J. P.; Cech, T. R.; Chang, H. Y. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell* **2016**, *165* (5), 1267–1279. https://doi.org/10.1016/j.cell.2016.04.028.
(61) Sharma, E.; Sterne-Weiler, T.; O'Hanlon, D.; Blencowe, B. J. Global Mapping of Human RNA-RNA Interactions. *Molecular Cell* **2016**, *62* (4), 618–626. https://doi.org/10.1016/j.molcel.2016.04.030.
(62) Pradère, U.; Brunschweiger, A.; Gebert, L. F. R.; Lucic, M.; Roos, M.; Hall, J. Chemical Synthesis of Mono- and Bis-Labeled Pre-MicroRNAs. *Angewandte Chemie International Edition* **2013**, *52* (46), 12028–12032. https://doi.org/10.1002/anie.201304986.
(63) Ziv, O.; Gabryelska, M. M.; Lun, A. T. L.; Gebert, L. F. R.; Sheu-Gruttadauria, J.; Meredith, L. W.; Liu, Z.-Y.; Kwok, C. K.; Qin, C.-F.; MacRae, I. J.; Goodfellow, I.; Marioni, J. C.; Kudla, G.; Miska, E. A. COMRADES Determines in Vivo RNA Structures and Interactions. *Nat Methods* **2018**, *15* (10), 785–788. https://doi.org/10.1038/s41592-018-0121-0.
(64) Tong, Y.; Gibaut, Q. M. R.; Rouse, W.; Childs-Disney, J. L.; Suresh, B. M.; Abegg, D.; Choudhary, S.; Akahori, Y.; Adibekian, A.; Moss, W. N.; Disney, M. D. Transcriptome-Wide Mapping of Small-Molecule RNA-Binding Sites in Cells Informs an Isoform-Specific Degrader of QSOX1 MRNA. *J. Am. Chem. Soc.* **2022**, *144* (26), 11620–11625. https://doi.org/10.1021/jacs.2c01929.
(65) Suresh, B. M.; Li, W.; Zhang, P.; Wang, K. W.; Yildirim, I.; Parker, C. G.; Disney, M. D. A General Fragment-Based Approach to Identify and Optimize Bioactive Ligands Targeting RNA. *Proceedings of the National Academy of Sciences* **2020**, *117* (52), 33197 LP – 33203. https://doi.org/10.1073/pnas.2012217117.
(66) Crielaard, S.; Maassen, R.; Vosman, T.; Rempkens, I.; Velema, W. A. Affinity-Based Profiling of the Flavin Mononucleotide Riboswitch. *J. Am. Chem. Soc.* **2022**, *144* (23), 10462–10470. https://doi.org/10.1021/jacs.2c02685.
(67) Baigude, H.; Ahsanullah; Li, Z.; Zhou, Y.; Rana, T. M. MiR-TRAP: A Benchtop Chemical Biology Strategy to Identify MicroRNA Targets. *Angewandte Chemie International Edition* **2012**, *51* (24), 5880–5883. https://doi.org/10.1002/anie.201201512.
(68) Aw, J. G. A.; Shen, Y.; Wilm, A.; Sun, M.; Lim, X. N.; Boon, K.-L.; Tapsin, S.; Chan, Y.-S.; Tan, C.-P.; Sim, A. Y. L.; Zhang, T.; Susanto, T. T.; Fu, Z.; Nagarajan, N.; Wan, Y. In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Molecular Cell* **2016**, *62* (4), 603–617. https://doi.org/10.1016/j.molcel.2016.04.028.
(69) Wielenberg, K.; Wang, M.; Yang, M.; Ozer, A.; Lis, J. T.; Lin, H. An Improved 4′-Aminomethyltrioxsalen-Based Nucleic Acid Crosslinker for Biotinylation of Double-Stranded DNA or RNA. *RSC Adv.* **2020**, *10* (65), 39870–39874. https://doi.org/10.1039/D0RA07437C.
(70) BLUM, F. Der Formaldehyde Als Hartungsmittle. *Microscopia Acta* **1893**, *10*, 314–315.
(71) Migneault, I.; Dartiguenave, C.; Bertrand, M. J.; Waldron, K. C. Glutaraldehyde: Behavior in Aqueous Solution, Reaction with Proteins, and Application to Enzyme Crosslinking. *BioTechniques* **2004**, *37* (5), 790–802. https://doi.org/10.2144/04375RV01.
(72) Knutson, S. D.; Sanford, A. A.; Swenson, C. S.; Korn, M. M.; Manuel, B. A.; Heemstra, J. M. Thermoreversible Control of Nucleic Acid Structure and Function with Glyoxal Caging. *J. Am. Chem. Soc.* **2020**, *142* (41), 17766–17781. https://doi.org/10.1021/jacs.0c08996.
(73) Leone, D.-L.; Pohl, R.; Hubálek, M.; Kadeřábková, M.; Krömer, M.; Sýkorová, V.; Hocek, M. Glyoxal-Linked Nucleotides and DNA for Bioconjugations and Crosslinking with Arginine-Containing Peptides and Proteins. *Chemistry – A European Journal* **2022**, *28* (14), e202104208. https://doi.org/10.1002/chem.202104208.
(74) Huang, H.; Hopkins, P. B. DNA Interstrand Cross-Linking by Formaldehyde: Nucleotide Sequence Preference and Covalent Structure of the Predominant Cross-Link Formed in Synthetic Oligonucleotides. *J. Am. Chem. Soc.* **1993**, *115* (21), 9402–9408. https://doi.org/10.1021/ja00074a005.
(75) Yi, Q.; Yang, R.; Shi, J.; Zeng, N.; Liang, D.; Sha, S.; Chang, Q. Effect of Preservation Time of Formalin-Fixed Paraffin-Embedded Tissues on Extractable DNA and RNA Quantity. *J Int Med Res* **2020**, *48* (6), 0300060520931259. https://doi.org/10.1177/0300060520931259.
(76) Engreitz, J. M.; Sirokman, K.; McDonel, P.; Shishkin, A. A.; Surka, C.; Russell, P.; Grossman, S. R.; Chow, A. Y.; Guttman, M.; Lander, E. S. RNA-RNA Interactions Enable Specific Targeting of Noncoding RNAs to Nascent Pre-MRNAs and Chromatin Sites. *Cell* **2014**, *159* (1), 188–199. https://doi.org/10.1016/j.cell.2014.08.018.
(77) Ulmer, E.; Meinke, M.; Ross, A.; Fink, G.; Brimacombe, R. Chemical Cross-Linking of Protein to RNA within Intact Ribosomal Subunits from Escherichia Coli. *Molec. Gen. Genet.* **1978**, *160* (2), 183–193. https://doi.org/10.1007/BF00267480.
(78) Zwieb, C.; Ross, A.; Rinke, J.; Meinke, M.; Brimacombe, R. Evidence for RNA-RNA Cross-Link Formation in Escherichia Coli Ribosomes. *Nucleic Acids Research* **1978**, *5* (8), 2705–2720. https://doi.org/10.1093/nar/5.8.2705.

(79) Lindahl, T.; Andersson, A. Rate of Chain Breakage at Apurinic Sites in Double-Stranded Deoxyribonucleic Acid. *Biochemistry* **1972**, *11* (19), 3618–3623. https://doi.org/10.1021/bi00769a019.
(80) K. Watts, J.; Katolik, A.; Viladoms, J.; J. Damha, M. Studies on the Hydrolytic Stability of 2′-Fluoroarabinonucleic Acid (2′F-ANA). *Organic & Biomolecular Chemistry* **2009**, *7* (9), 1904–1910. https://doi.org/10.1039/B900443B.
(81) Gates, K. S.; Nooner, T.; Dutta, S. Biologically Relevant Chemical Reactions of N7-Alkylguanine Residues in DNA. *Chem. Res. Toxicol.* **2004**, *17* (7), 839–856. https://doi.org/10.1021/tx049965c.
(82) Jobst, K. A.; Klenov, A.; Neller, K. C. M.; Hudak, K. A. Effect of Depurination on Cellular and Viral RNA. In *Modified Nucleic Acids in Biology and Medicine*; Jurga, S., Erdmann (Deceased), V. A., Barciszewski, J., Eds.; RNA Technologies; Springer International Publishing: Cham, 2016; pp 273–297. https://doi.org/10.1007/978-3-319-34175-0_12.
(83) Liu, Y.; Rodriguez, Y.; Ross, R. L.; Zhao, R.; Watts, J. A.; Grunseich, C.; Bruzel, A.; Li, D.; Burdick, J. T.; Prasad, R.; Crouch, R. J.; Limbach, P. A.; Wilson, S. H.; Cheung, V. G. RNA Abasic Sites in Yeast and Human Cells. *Proceedings of the National Academy of Sciences* **2020**, *117* (34), 20689–20695. https://doi.org/10.1073/pnas.2011511117.
(84) Datta, B.; Weiner, A. M. Cross-Linking of U2 SnRNA Using Nitrogen Mustard. Evidence for Higher Order Structure. *Journal of Biological Chemistry* **1992**, *267* (7), 4497–4502. https://doi.org/10.1016/S0021-9258(18)42860-8.
(85) Zhang, Q.; Yu, E. T.; Kellersberger, K. A.; Crosland, E.; Fabris, D. Toward Building a Database of Bifunctional Probes for the MS3D Investigation of Nucleic Acids Structures. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (11), 1570–1581. https://doi.org/10.1016/j.jasms.2006.06.002.
(86) Zuravka, I.; Roesmann, R.; Sosic, A.; Wende, W.; Pingoud, A.; Gatto, B.; Göttlich, R. Synthesis and DNA Cleavage Activity of Bis-3-Chloropiperidines as Alkylating Agents. *ChemMedChem* **2014**, *9* (9), 2178–2185. https://doi.org/10.1002/cmdc.201400034.
(87) Gltterman, C. O.; Rlckes, E. L.; Wolf, D. E.; Madas, J.; Zimmerman, S. B.; Stoudt, T. H.; Demny, T. C. THE HUMAN TUMOR-EGG HOST SYSTEM IV. DISCOVERY OF A NEW ANTI-TUMOR AGENT, COMPOUND 593 A. *J. Antibiot.* **1970**, *23* (6), 305–310. https://doi.org/10.7164/antibiotics.23.305.
(88) Sosic, A.; Göttlich, R.; Fabris, D.; Gatto, B. B-CePs as Cross-Linking Probes for the Investigation of RNA Higher-Order Structure. *Nucleic Acids Research* **2021**, *49* (12), 6660–6672. https://doi.org/10.1093/nar/gkab468.
(89) Kurinomaru, T.; Kojima, N.; Kurita, R. An Alkylating Immobilization Linker for Immunochemical Epigenetic Assessment. *Chem. Commun.* **2017**, *53* (59), 8308–8311. https://doi.org/10.1039/C7CC02883K.
(90) Kojima, N.; Suda, T.; Kurinomaru, T.; Kurita, R. Immobilization of DNA with Nitrogen Mustard–Biotin Conjugate for Global Epigenetic Analysis. *Analytica Chimica Acta* **2018**, *1043*, 107–114. https://doi.org/10.1016/j.aca.2018.09.008.
(91) Knorre, D. G.; Pustoshilova, N. M.; Teplova, M.; Shamovskiĭ, G. G. The production of transfer RNA acetylated by 2'-oxy groups. *Biokhimiia* **1965**, *30* (6), 1218–1224.
(92) Habibian, M.; Velema, W. A.; Kietrys, A. M.; Onishi, Y.; Kool, E. T. Polyacetate and Polycarbonate RNA: Acylating Reagents and Properties. *Org Lett* **2019**, *21* (14), 5413–5416. https://doi.org/10.1021/acs.orglett.9b01526.
(93) Lu, Z.; Gong, J.; Zhang, Q. C. P. A. R. I. S. Psoralen Analysis of RNA Interactions and Structures with High Throughput and Resolution. *Methods Mol Biol* **2018**, *1649*, 59–84. https://doi.org/10.1007/978-1-4939-7213-5_4.
(94) Sugimoto, Y.; Vigilante, A.; Darbo, E.; Zirra, A.; Militti, C.; D'Ambrogio, A.; Luscombe, N. M.; Ule, J. HiCLIP Reveals the in Vivo Atlas of MRNA Secondary Structures Recognized by Staufen 1. *Nature* **2015**, *519* (7544), 491–494. https://doi.org/10.1038/nature14280.
(95) Kudla, G.; Granneman, S.; Hahn, D.; Beggs, J. D.; Tollervey, D. Cross-Linking, Ligation, and Sequencing of Hybrids Reveals RNA-RNA Interactions in Yeast. *Proc Natl Acad Sci U S A* **2011**, *108* (24), 10010–10015. https://doi.org/10.1073/pnas.1017386108.
(96) Fazal, F. M.; Han, S.; Parker, K. R.; Kaewsapsak, P.; Xu, J.; Boettiger, A. N.; Chang, H. Y.; Ting, A. Y. Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell* **2019**, *178* (2), 473-490 426. https://doi.org/10.1016/j.cell.2019.05.027.
(97) Cai, Z.; Cao, C.; Ji, L.; Ye, R.; Wang, D.; Xia, C.; Wang, S.; Du, Z.; Hu, N.; Yu, X. RIC-Seq for Global in Situ Profiling of RNA-RNA Spatial Interactions. *Nature* **2020**, *582* (7812), 432–437. https://doi.org/10.1038/s41586-020-2249-1.
(98) Kundu, L. M.; Linne, U.; Marahiel, M.; Carell, T. RNA Is More UV Resistant than DNA: The Formation of UV-Induced DNA Lesions Is Strongly Sequence and Conformation Dependent. *Chemistry – A European Journal* **2004**, *10* (22), 5697–5705. https://doi.org/10.1002/chem.200305731.
(99) Wurtmann, E. J.; Wolin, S. L. RNA under Attack: Cellular Handling of RNA Damage. *Critical Reviews in Biochemistry and Molecular Biology* **2009**, *44* (1), 34–49. https://doi.org/10.1080/10409230802594043.
(100) Kladwang, W.; Hum, J.; Das, R. Ultraviolet Shadowing of RNA Can Cause Significant Chemical Damage in Seconds. *Sci Rep* **2012**, *2* (1), 517. https://doi.org/10.1038/srep00517.

(101) Knörlein, A.; Sarnowski, C.; Vries, T. de; Stoltz, M.; Götze, M.; Aebersold, R.; Allain, F.; Leitner, A.; Hall, J. Structural Requirements for Photo-Induced RNA-Protein Cross-Linking. **2021**. https://doi.org/10.26434/chemrxiv-2021-05zhj.
(102) Urdaneta, E. C.; Beckmann, B. M. Fast and Unbiased Purification of RNA-Protein Complexes after UV Cross-Linking. *Methods* **2020**, *178*, 72–82. https://doi.org/10.1016/j.ymeth.2019.09.013.
(103) Groehler, A. I.; Villalta, P. W.; Campbell, C.; Tretyakova, N. Covalent DNA–Protein Cross-Linking by Phosphoramide Mustard and Nornitrogen Mustard in Human Cells. *Chem. Res. Toxicol.* **2016**, *29* (2), 190–202. https://doi.org/10.1021/acs.chemrestox.5b00430.
(104) Siegfried, N. A.; Busan, S.; Rice, G. M.; Nelson, J. A.; Weeks, K. M. RNA Motif Discovery by SHAPE and Mutational Profiling (SHAPE-MaP. *Nat Methods* **2014**, *11* (9), 959–965. https://doi.org/10.1038/nmeth.3029.
(105) Ozsolak, F.; Platt, A. R.; Jones, D. R.; Reifenberger, J. G.; Sass, L. E.; McInerney, P.; Thompson, J. F.; Bowers, J.; Jarosz, M.; Milos, P. M. Direct RNA sequencing. *Nature* **2009**, *461* (7265), 814–818. https://doi.org/10.1038/nature08390.
(106) Sastry, S. S.; Spielmann, H. P.; Hoang, Q. S.; Phillips, A. M.; Sancar, A.; Hearst, J. E. Laser-Induced Protein-DNA Cross-Links via Psoralen Furanside Monoadducts. *Biochemistry* **1993**, *32* (21), 5526–5538.
(107) Travis, A. J.; Moody, J.; Helwak, A.; Tollervey, D.; Kudla, G. Hyb: A Bioinformatics Pipeline for the Analysis of CLASH (Crosslinking, Ligation and Sequencing of Hybrids) Data. *Methods* **2014**, *65* (3), 263–273. https://doi.org/10.1016/j.ymeth.2013.10.015.
(108) Zhang, M.; Hwang, I. T.; Li, K.; Bai, J.; Chen, J. F.; Weisman, T.; Zou, J. Y.; Lu, Z. Classification and Clustering of RNA Crosslink-Ligation Data Reveal Complex Structures and Homodimers. *Genome Res* **2022**. https://doi.org/10.1101/gr.275979.121.
(109) Gabryelska, M. M.; Badrock, A. P.; Lau, J. Y.; O'Keefe, R. T.; Crow, Y. J.; Kudla, G. Global Mapping of RNA Homodimers in Living Cells. *Genome Res* **2022**, *32* (5), 956–967. https://doi.org/10.1101/gr.275900.121.
(110) Schäfer, R. A.; Voß, B. RNAnue: efficient data analysis for RNA-RNA interactomics. *Nucleic Acids Res* **2021**, *49* (10), 5493–5501. https://doi.org/10.1093/nar/gkab340.
(111) Zhong, C.; Zhang, S. Accurate and Efficient Mapping of the Cross-Linked MicroRNA-MRNA Duplex Reads, 2019, *18*, 11–19. https://doi.org/10.1016/j.isci.2019.05.038.
(112) Robinson, J. T.; Thorvaldsdottir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P. Integrative genomics viewer. *Nat Biotechnol* **2011**, *29* (1), 24–26. https://doi.org/10.1038/nbt.1754.
(113) Lu, Z.; Guo, J. K.; Wei, Y.; Dou, D. R.; Zarnegar, B.; Ma, Q.; Li, R.; Zhao, Y.; Liu, F.; Choudhry, H. Structural Modularity of the XIST Ribonucleoprotein Complex. *Nat Commun* **2020**, *11* (1), 6163. https://doi.org/10.1038/s41467-020-20040-3.
(114) Li, P.; Wei, Y.; Mei, M.; Tang, L.; Sun, L.; Huang, W.; Zhou, J.; Zou, C.; Zhang, S.; Qin, C. F. Integrative Analysis of Zika Virus Genome RNA Structure Reveals Critical Determinants of Viral Infectivity. *Cell Host Microbe* **2018**, *24* (6), 875-886 875. https://doi.org/10.1016/j.chom.2018.10.011.
(115) Metkar, M.; Ozadam, H.; Lajoie, B. R.; Imakaev, M.; Mirny, L. A.; Dekker, J.; Moore, M. J. Higher-Order Organization Principles of Pre-Translational MRNPs. *Mol Cell* **2018**, *72* (4), 715-726 713. https://doi.org/10.1016/j.molcel.2018.09.012.
(116) Zhou, J.; Li, P.; Zeng, W.; Ma, W.; Lu, Z.; Jiang, R.; Zhang, Q. C.; Jiang, T. IRIS: A Method for Predicting in Vivo RNA Secondary Structures Using PARIS Data. *Quantitative Biology* **2020**, *8* (4), 369–381.
(117) Denny, S. K.; Bisaria, N.; Yesselman, J. D.; Das, R.; Herschlag, D.; Greenleaf, W. J. High-Throughput Investigation of Diverse Junction Elements in RNA Tertiary Folding. *Cell* **2018**, *174* (2), 377-390 320. https://doi.org/10.1016/j.cell.2018.05.038.
(118) Tivon, Y.; Falcone, G.; Deiters, A. Protein Labeling and Crosslinking by Covalent Aptamers. *Angewandte Chemie* **2021**, *133* (29), 16035–16040. https://doi.org/10.1002/ange.202101174.
(119) Aw, J. G. A.; Lim, S. W.; Wang, J. X.; Lambert, F. R. P.; Tan, W. T.; Shen, Y.; Zhang, Y.; Kaewsapsak, P.; Li, C.; Ng, S. B.; Yang, M.; Zhu, P.; Cheema, J.; Bloomer, R.; Mikulski, P.; Liu, Q.; Zhang, Y.; Dean, C.; Ding, Y.; Stephenson, W.; Razaghi, R.; Busan, S.; Weeks, K. M.; Timp, W.; Smibert, P. Determination of Isoform-Specific RNA Structure with Nanopore Long Reads. *Nat Biotechnol* **2021**, *39* (3), 336–346. https://doi.org/10.1038/s41587-020-0712-z.
(120) Rivas, E.; Clements, J.; Eddy, S. R. A Statistical Test for Conserved RNA Structure Shows Lack of Evidence for Structure in LncRNAs. *Nat Methods* **2016**.
(121) Ziv, O.; Price, J.; Shalamova, L.; Kamenova, T.; Goodfellow, I.; Weber, F.; Miska, E. A. The Short- and Long-Range RNA-RNA Interactome of SARS-CoV-2. *Mol Cell* **2020**, *80* (6), 1067-1077 1065. https://doi.org/10.1016/j.molcel.2020.11.004.
(122) Dadonaite, B.; Gilbertson, B.; Knight, M. L.; Trifkovic, S.; Rockman, S.; Laederach, A.; Brown, L. E.; Fodor, E.; Bauer, D. L. V. The Structure of the Influenza A Virus Genome. *Nat Microbiol* **2019**, *4* (11), 1781–1789. https://doi.org/10.1038/s41564-019-0513-7.
(123) Yang, S. L.; DeFalco, L.; Anderson, D. E.; Zhang, Y.; Aw, J. G. A.; Lim, S. Y.; Lim, X. N.; Tan, K. Y.; Zhang, T.; Chawla, T. Comprehensive Mapping of SARS-CoV-2 Interactions in Vivo Reveals Functional Virus-Host Interactions. *Nat Commun* **2021**, *12* (1), 5113. https://doi.org/10.1038/s41467-021-25357-1.

(124) Lin, Y.; Schmidt, B. F.; Bruchez, M. P.; McManus, C. J. Structural Analyses of NEAT1 LncRNAs Suggest Long-Range RNA Interactions That May Contribute to Paraspeckle Architecture. *Nucleic Acids Res* **2018**, *46* (7), 3742–3752. https://doi.org/10.1093/nar/gky046.
(125) Helwak, A.; Kudla, G.; Dudnakova, T.; Tollervey, D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **2013**, *153* (3), 654–665. https://doi.org/10.1016/j.cell.2013.03.043.
(126) Iosub, I. A.; Nues, R. W.; McKellar, S. W.; Nieken, K. J.; Marchioretto, M.; Sy, B.; Tree, J. J.; Viero, G.; Granneman, S. Hfq CLASH Uncovers SRNA-Target Interaction Networks Linked to Nutrient Availability Adaptation. *Elife* **2020**, 9. https://doi.org/10.7554/eLife.54655.
(127) Lu, Z.; Carter, A. C.; Chang, H. Y. Mechanistic Insights in X-Chromosome Inactivation. *Philos Trans R Soc Lond B Biol Sci* **2017**, *372*. https://doi.org/10.1098/rstb.2016.0356.
(128) Dethoff, E. A.; Chugh, J.; Mustoe, A. M.; Al-Hashimi, H. M. Functional Complexity and Regulation through RNA Dynamics. *Nature* **2012**, *482* (7385), 322–330. https://doi.org/10.1038/nature10885.
(129) Feng, C.; Chan, D.; Joseph, J.; Muuronen, M.; Coldren, W. H.; Dai, N.; Corrêa, I. R.; Furche, F.; Hadad, C. M.; Spitale, R. C.; Yoshimura, Y.; Ohtake, T.; Okada, H.; Fujimoto, K. Light-Activated Chemical Probing of Nucleobase Solvent Accessibility inside Cells. *Nat Chem Biol* **2018**, *14* (3), 276–283. https://doi.org/10.1038/nchembio.2548.
(130) Rose, D. A.; Treacy, J. W.; Yang, Z. J.; Ko, J. H.; Houk, K. N.; Maynard, H. D. Self-Immolative Hydroxybenzylamine Linkers for Traceless Protein Modification. *J. Am. Chem. Soc.* **2022**, *144* (13), 6050–6058. https://doi.org/10.1021/jacs.2c01136.
(131) Šimon, P.; Tichotová, M.; García Gallardo, M.; Procházková, E.; Baszczyňski, O. Phosphate-Based Self-Immolative Linkers for Tuneable Double Cargo Release. *Chemistry – A European Journal* **2021**, *27* (50), 12763–12775. https://doi.org/10.1002/chem.202101805.
(132) Graveley, B. R.; Treutlein, B.; Gokce, O.; Quake, S. R.; Südhof, T. C. Mutually Exclusive Splicing of the Insect Dscam Pre-MRNA Directed by Competing Intronic RNA Secondary Structures. *Cell* **2005**, *123* (1), 65–73. https://doi.org/10.1016/j.cell.2005.07.028.
(133) Homan, P. J.; Favorov, O. V.; Lavender, C. A.; Kursun, O.; Ge, X.; Busan, S.; Dokholyan, N. V.; Weeks, K. M.; Cheng, C. Y.; Chou, F. C.; Kladwang, W.; Tian, S.; Cordero, P.; Das, R. Single-Molecule Correlated Chemical Probing of RNA. *Proc Natl Acad Sci U S A* **2014**, *111* (38), 13858–13863. https://doi.org/10.1073/pnas.1407306111.
(134) Wang, F.; Zuroske, T.; Watts, J. K. RNA Therapeutics on the Rise. *Nat Rev Drug Discov* **2020**, *19* (7), 441–442. https://doi.org/10.1038/d41573-020-00078-0.
(135) Singh, N. N.; Howell, M. D.; Androphy, E. J.; Singh, R. N. How the Discovery of ISS-N1 Led to the First Medical Therapy for Spinal Muscular Atrophy. *Gene Ther* **2017**, *24* (9), 520–526. https://doi.org/10.1038/gt.2017.34.