



REVIEW

Genome-wide Mapping of Cellular Protein–RNA Interactions Enabled by Chemical Crosslinking



Xiaoyu Li ^{1,#}, Jinghui Song ^{1,#}, Chengqi Yi ^{1,2,3,*}

¹ State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China

² Synthetic and Functional Biomolecules Center, Beijing 100871, China

³ Peking-Tsinghua Center for Life Sciences, Beijing 100871, China

Received 9 March 2014; revised 24 March 2014; accepted 25 March 2014

Available online 18 April 2014

Handled by Jun Yu

KEYWORDS

Protein–RNA interactions;
High-throughput
sequencing;
Crosslinking;
RNA-binding proteins;
Aza-IP;
miCLIP

Abstract RNA–protein interactions influence many biological processes. Identifying the binding sites of RNA-binding proteins (RBPs) remains one of the most fundamental and important challenges to the studies of such interactions. Capturing RNA and RBPs via chemical crosslinking allows stringent purification procedures that significantly remove the non-specific RNA and protein interactions. Two major types of chemical crosslinking strategies have been developed to date, *i.e.*, UV-enabled crosslinking and enzymatic mechanism-based covalent capture. In this review, we compare such strategies and their current applications, with an emphasis on the technologies themselves rather than the biology that has been revealed. We hope such methods could benefit broader audience and also urge for the development of new methods to study RNA–RBP interactions.

Introduction

RNAs undergo multiple RBP-mediated processing and regulatory steps to exert their biological functions. To understand RNA processing and regulation, great efforts have been made

to study protein–RNA interactions in the cellular context. Among many strategies that have been developed, mapping RNA–protein interactions on a genomic scale (thus often coupled with high-throughput analyses) represents one great example. In the initial attempts to identify RNAs that are bound by specific RBPs, RNA immunoprecipitation was combined with microarray analysis (RIP-ChIP) [1,2]. However, these methods are limited to stable ribonucleoprotein particles (RNPs), which are prone to contamination of indirect or non-physiological interactions [3]. In order to achieve high specificity (and also resolution), methods that allow covalent capture of RNA–protein interactions *in vivo* have been developed. In the following sections, we will describe these novel methods for mapping the binding sites of RBPs on a genome-wide scale.

* Corresponding author.

E-mail: chengqi.yi@pku.edu.cn (Yi C).

Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.



Production and hosting by Elsevier

Panoramic views of protein–RNA interactions enabled by UV crosslinking

CLIP and HITS-CLIP

UV cross-linking and immunoprecipitation (CLIP) is a milestone technology invented by the Darnell Laboratory in 2003 (Figure 1) [4]. Back in the 1980s, it was reported that UV light induces covalent crosslinks between proteins and RNAs, without causing crosslinks between proteins [5,6]. CLIP makes use of short wave UV irradiation at 254 nm to induce the formation of covalent crosslinks only at sites of direct contact between proteins and RNAs, in the context of whole tissues, organisms or individual cell types. After UV irradiation, cells are lysed and the cross-linked RNA–RBP is first treated with RNase to trim the RNA size to about 60–100 nucleotides in length. Since the protein–RNA complex is covalently linked, stringent conditions can be applied during immunoprecipitation to purify the protein–RNA complexes. Such immunoprecipitation procedure is generally enabled with antibodies against the proteins of interest or protein epitope tags. Subsequently, the protein–RNA complexes are separated on SDS–polyacrylamide gel electrophoresis (SDS–PAGE), and transferred to nitrocellulose membrane followed by proteinase K treatment. Such step also helps to remove contamination caused by the antibody or during immunoprecipitation step. The RNA molecules recovered are then 5' and 3' ligated, allowing PCR amplification of the target RNA from RNPs of interest.

This original CLIP protocol was first applied to identify Nova RNA targets in the brain by the Darnell group [4]. Nova is a neuron-specific RBP, which regulates neuronal RNA splicing. In 2005, Darnell and colleagues optimized the original CLIP protocol and further improved the specificity of the method. In the current CLIP protocol, the 3' RNA linker ligation step is done on-bead [7]. There are three advantages conferred by this modification. First, free 3'-linkers can be removed during SDS–PAGE, so that it can prevent the ligation of 5'-linker and 3'-linker, self-ligation, circularization of target RNA and the ligation of bacterial rRNA that may come from the commercial RNA ligase or proteinase K. Therefore, signal-to-noise ratio can be greatly improved. Second, there is no need to separate free RNAs on urea–PAGE, which is a necessary step to separate linker–linker ligation from previous CLIP protocol. Third, the usage of high concentration of RNA linkers can be avoided, which alleviates the aforementioned self-ligation problem as well. In 2008, a new modification to CLIP, high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP), was adopted to harness the power of next-generation sequencing (NGS). Thus, HITS-CLIP has made it highly feasible to obtain genome-wide protein–RNA interactions. HITS-CLIP was first applied to study Nova–RNA interactions and uncovered the 3' end RNA processing rules in the brain [8].

Control experiments are key components in every technique, and CLIP or HITS-CLIP is no exception. There are two essential controls in CLIP or HITS-CLIP. First, to verify that UV light has indeed caused covalent crosslinks between proteins and RNA, a non-UV treated sample should be used for immunoprecipitation, in which no RNA should be detected. Second, the antibody that recognizes the protein of interest must be specific. Therefore, it is very necessary to

run a control without antibody, or a control with knockout cell or tissue, and do a mock IP experiment.

Apart from Nova, the RNA targets of several RBPs have been successfully identified by CLIP or HITS-CLIP in various biological systems. For example, with the help of the CLIP method, hnRNP A1, a nucleo-cytoplasmic shuttling protein, is shown to be required for processing of miR-18a [9] and SF2/ASF, a prototype member of the SR protein family, was reported to regulate processing of specific mRNAs with the help of CLIP [10]. CLIP was also used to study the function of the RNA-binding protein Rrm4, discovering that Rrm4 may transport RNAs from the nucleus to cell poles [11]. Additionally, HITS-CLIP was also applied to study other RBPs, such as Argonaute in mouse brain [12], polypyrimidine tract-binding protein (PTB, also known as hnRNP I) in HeLa cells [13], FOX2 in human embryonic stem cells [14] and so on.

Nonetheless, HITS-CLIP and CLIP methods discussed above cannot achieve individual-nucleotide resolution of the binding site. In 2011, it was reported that the cross-linked nucleotide could be deleted or mutated during reverse transcription [15,16]. Therefore, through the analysis of deletion or mutation in sequencing reads, HITS-CLIP can identify exact crosslink sites and achieve individual-nucleotide resolution of the binding site (Figure 1) [15,16].

In summary, CLIP and HITS-CLIP are very useful techniques to study protein–RNA interactions, which can be applied to different biological samples such as bacteria, fungi, yeast, *Caenorhabditis elegans*, mammalian tissue including brain, and culture cells including human embryonic stem cells and HeLa cells.

PAR-CLIP

It was estimated that in CLIP or HITS-CLIP, the maximal crosslinking efficiency ranges from 1% to 5% with purified protein and radiolabeled RNA [17]. To improve the crosslinking efficiency, photoactive nucleoside analogues, 4-thiouridine (4-SU) and 6-thioguanosine (6-SG), and hence photoactivatable ribonucleoside-enhanced CLIP (PAR-CLIP) were used by Tuschl and colleagues [18]. In PAR-CLIP, 4-SU and 6-SG are added to the growth medium, which are then taken up by cells and eventually incorporated into newly-synthesized RNA molecules without obvious toxicity. The formation of covalent crosslinks between proteins and RNAs is performed under UV irradiation at 365 nm, instead of UV 254 nm used in CLIP. The following steps are similar to CLIP protocol, including RNase treatment, immunoprecipitation, recovery of RNA fragment, reverse transcription and sequencing.

It is noteworthy that the incorporated 4-SU can lead to T to C transition in the sequenced cDNA. Therefore, it is possible to identify crosslink sites at individual-nucleotide resolution by analyzing the mutations in cDNA sequences (Figure 1). The accuracy of PAR-CLIP has been verified by identifying the RNA targets of several RBPs in recent years. At the very beginning, Thomas Tuschl and colleagues selected several intensely-studied RBPs and microRNA-containing RNP complexes (miRNPs), including pumilio homologue 2 (PUM2), quaking (QKI), insulin-like growth factor 2 mRNA-binding proteins 1–3 (IGF2BP1–3), the Argonaute proteins (AGO) and trinucleotide repeat-containing proteins

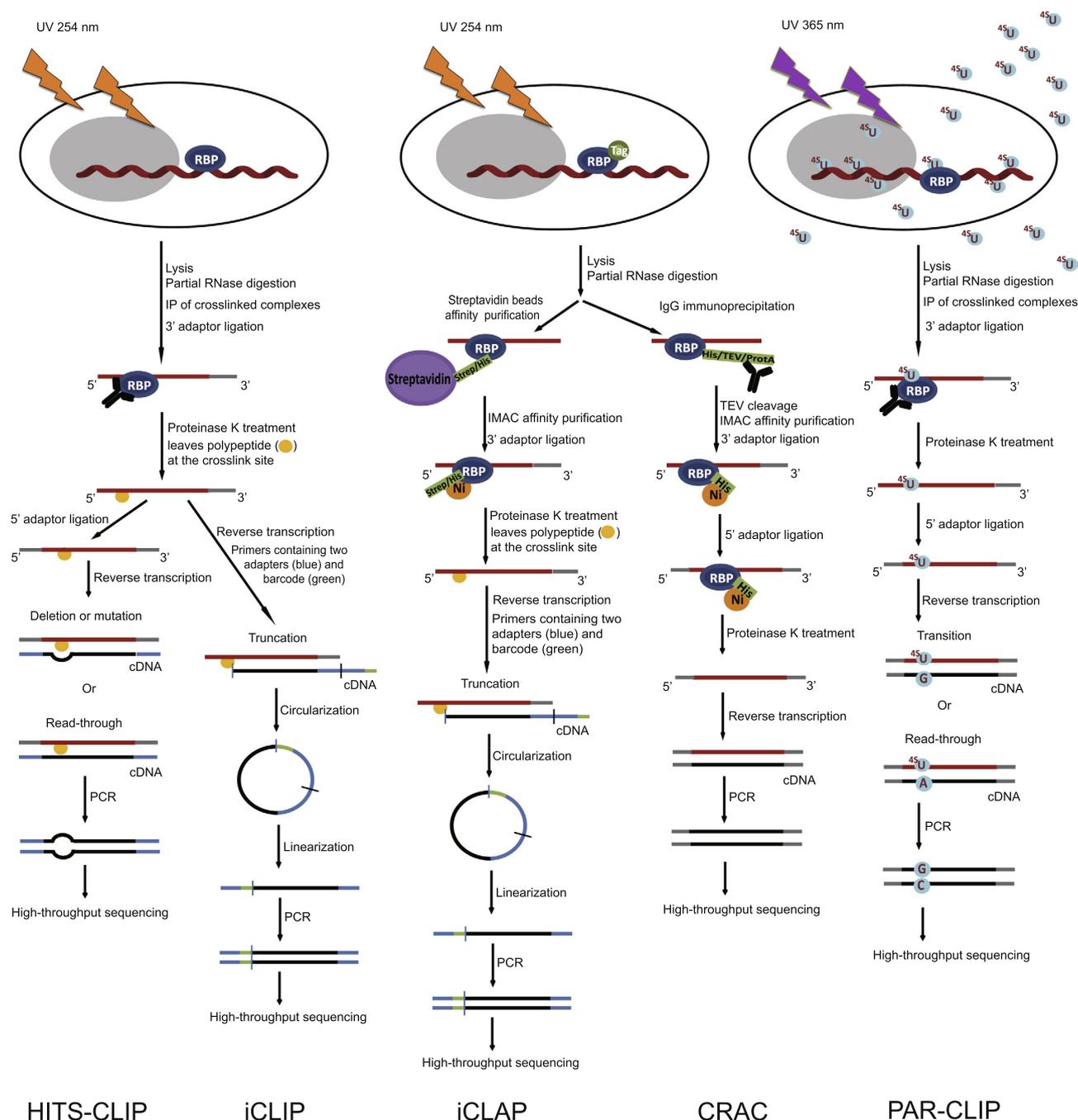


Figure 1 Outline of HITS-CLIP, PAR-CLIP and several variants, iCLIP, iCLAP and CRAC

High-throughput sequencing CLIP (HITS-CLIP) and individual-nucleotide resolution CLIP (iCLIP) are in the left panels; individual-nucleotide resolution crosslinking affinity purification (iCLAP) and crosslinking and cDNA analysis (CRAC) are in the middle panels; photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) is in the right panel. PAR-CLIP uses thioribonucleosides and UV at 365 nm to form the complex of RNA and RNA-binding protein (RBP), while the other four methods utilize UV at 254 nm. Isolation of RNA–RBP complexes is achieved either by immunoprecipitation (IP) (PAR-CLIP, HITS-CLIP and iCLIP) or by double affinity purification (iCLAP and CRAC). iCLAP and CRAC use immobilized metal ion affinity chromatography (IMAC) under denaturing conditions as a secondary purification. To achieve individual-nucleotide resolution, HITS-CLIP utilizes deletion or mutation during reverse transcription, iCLIP and iCLAP take advantage of truncated cDNAs, and PAR-CLIP makes use of thymidine (T) to cytidine (C) transition in cDNA. TEV, tobacco etch virus; ProtA, *Staphylococcus aureus* protein A.

6 (TNRC6A–C) in HEK293 cells, to identify the binding sites [18]. One year later, transcriptome-wide binding sites of the RBP defective germline development protein 1 (GLD-1) in *C. elegans* [19], RNA targets of the RNA-binding

protein HuR [16,20,21], the binding targets of the sequence-specific RBP complex Nrd1 and Nab3 in yeast [22] and the binding motif of PAP-associated domain-containing 5 (PAPD5) [23] were also identified using PAR-CLIP.

In addition, PAR-CLIP is particularly suitable for studies using cultured cells, because of its high uptake efficiency of nucleoside analogues (incorporation efficiency is up to 4% for 4-SU in relative to uridine [18]). In short, the PAR-CLIP method has higher cross-linking efficiency and can also achieve individual-nucleotide resolution of the binding sites.

CLIP-derived methods

In the case of the protein Nova, 85% of cDNAs are truncated, because reverse transcriptase stops at the UV-induced cross-link site, in which peptides may not be removed by proteinase K [24]. However, every coin has two sides. Individual-nucleotide resolution CLIP (iCLIP), a variation of CLIP, was developed by the Ule laboratory, which made use of this experimental observation to identify crosslinked sites [25]. In iCLIP protocol, co-immunoprecipitated RNA is ligated to a 3' adaptor, followed by proteinase K treatment, resulting in a covalently-bound peptide on the RNA. Using primers containing two cleavable adaptor regions and one random barcode, cDNA will be truncated at crosslink site during reverse transcription. Subsequently, cDNA is circularized, linearized, amplified and sequenced. Thus, it is easy to find out the residue next to the crosslink site, which ideally should be the first nucleotide after the barcoded sequence (Figure 1). iCLIP was first applied to study the function of hnRNP in splicing [25]. Subsequently, this technique was used to identify the RNA targets of the T-cell-restricted intracellular antigen-1 (TIA1), TIA-like 1 (TIAL1) [26] and TAR DNA-binding protein 43 (TDP-43) [27].

Crosslinking and affinity purification (iCLAP), a method to purify streptavidin/histidine (Strep/His) double-tagged RBP using stringent affinity purification instead of immunoprecipitation, is a variation of iCLIP (Figure 1). In the case of TIA1 and TIAL1, the TIA1 antibody has a slight cross-reactivity during affinity purification to TIAL1 [26]. Therefore, it is necessary to use RBP with a tag at the N- or C-terminus to co-purify the RNP complex, under the conditions of poor antibodies. In the iCLAP technique, Jernej Ule and colleagues, who invented the iCLIP method, used magnetic streptavidin beads in the first purification step followed by cobalt beads under denaturing conditions as a secondary purification step [26]. Subsequently, cDNAs are amplified and sequenced in the similar way as using iCLIP.

Another variation of CLIP is the crosslinking and cDNA analysis (CRAC), where RBPs are tagged with C-terminal $6 \times$ His, tobacco etch virus (TEV) protease site and protein A tags (Figure 1). Similar to iCLIP, immunoglobulin G (IgG) beads are used in the first purification step. The RNA-protein complexes were subsequently treated with TEV protease, followed by immobilized metal-ion affinity chromatography (IMAC) as a secondary purification. Using the CRAC method, the Tollervey laboratory identified the binding sites of the probable pre-mRNA-splicing factor ATP-dependent RNA helicase Prp43 [28] and the small nucleolar RNPs (snoRNPs) Nop1, Nop56, Nop58 and rRNA processing 9 (Rrp9) [29] in yeast.

In summary, as CLIP variants, iCLIP, iCLAP and CRAC are similar to HITS-CLIP and PAR-CLIP in essence. All methods can achieve single nucleotide resolution and produce data with high precision.

Crosslinking to identify RNA-bound proteome

RBPs play diverse roles in many biological processes and influence the RNA metabolism [30]. Many RBPs have canonical RNA binding domains, including the RNA recognition motif (RRM), heterogeneous nuclear RNP K-homology domain (KH), zinc finger (Znf), etc. [31]. Based on such observation, the mammalian genome has been predicted to encode about 600 RBPs [32]. However, researchers also have reported that there are quite a few RBPs that do not contain such canonical RNA-binding domains [23,33]. Therefore, mere computational predictions will likely miss these non-canonical RBPs.

Two types of experiments can be performed in order to identify an RBP. The first one takes advantage of protein arrays, which were spotted with tagged recombinant proteins. In fact, two groups have identified about 200 RBPs in yeast via protein arrays [34,35]. The other one utilizes RNA pull-down methods, in which RNA of interest is immobilized and then used to pull out potential RBPs from (very often) cell lysates via affinity. Combining the use of an RNA tag and high-resolution quantitative mass spectrometry (MS), interacting partners for RNA motifs of general interest can be detected [36]. For instance, by hybridizing with affinity-tagged oligodT, mRNA-binding proteome can be studied [37]. Nonetheless, these methods cannot discriminate direct RNA-protein interactions from indirect ones.

In 2012, two groups independently reported a new approach to study the mRNA interactome. In the first step of their methods, they use UV light to covalently crosslink RNA and RBPs. The UV light can only crosslink protein bound to RNA directly but does not affect protein-protein interaction. The next step involves affinity purification of mRNAs using oligodT, instead of trimming down the RNAs as performed in CLIP-type experiments. In this way, mRNA-binding proteins that are covalently attached to mRNA molecules will also be pulled down at the same time. By high-resolution quantitative MS, mRNA interactome can be identified in a high throughput manner. These studies have identified about 800 RBPs, including about 300 RBPs that are not annotated previously [38,39].

Cellular protein-RNA interactions enabled by mechanistic crosslinking

More than 100 post-transcriptional RNA modifications have been identified in cellular RNAs until now [40]. Yet, identifying the modifying enzymes along with their physiological sites of modification remains a challenging task [41].

Among these RNA modifications, RNA cytosine methylation, occurring at the C5 position (m^5C), has been detected in tRNAs, rRNAs and mRNAs [42,43]. There are six families of m^5C RNA methyltransferase (RMT), but only two of them, NOP2/Sun RMT family member 2 (NSUN2) and DNA (cytosine-5-)-methyltransferase 2 (DNMT2), have been identified in higher eukaryotes [42]. Both were known previously to work on specific tRNA molecules [44-46]; yet the full spectrum of targets remains an open question for both enzymes. To identify the target sites of these RMT, two groups separately developed two new crosslinking methods to capture RNA-modifying enzymes that are linked directly to their targets. Both methods involve the formation of covalent RNP

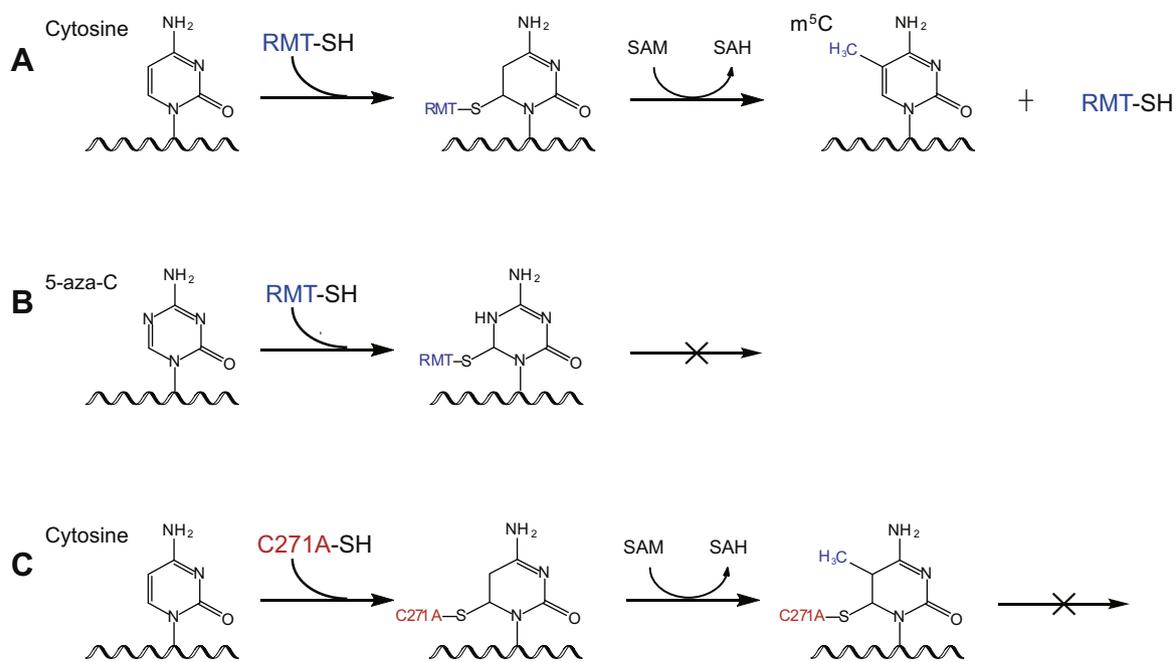


Figure 2 Mechanistic crosslinking of RNA-modifying enzymes and target RNA molecules

A. Reaction mechanism of a typical m^5C -RNA methyltransferase (RMT, in blue). B. 5-aza-cytidine (5-aza-C) traps RMT (in blue) with its target RNA molecules. C. Cysteine mutation (C271A, in red) in the active site of NSUN2 prevents release of the methyltransferase after methylation reaction, thereby enabling crosslinking of protein–RNA complex. SAM, S-adenosyl methionine; SAH, S-adenosyl homocysteine.

complexes, yet they are based on enzymatic mechanism instead of using UV-crosslinking [47,48].

It has been found that m^5C -RMT could form a covalent enzyme–substrate intermediate during the methylation reactions [42]. The covalent bond is formed, at the early stage of methylation process, between a cysteine residue of the m^5C -RMT and the C6 atom of cytosine of the RNA target. Then the RNA methyltransferase transfers a methyl group from cofactor S-adenosyl methionine (SAM) to the C5 of the target cytosine (Figure 2). The enzyme is finally released by β -elimination. Interestingly, studies have shown that if the cytosine of the RNA targets is replaced by a cytosine analogue, 5-azacytidine (5-aza-C), the methyltransferase cannot be released from the RNA target sites and thus remains covalently attached to its targets (Figure 2). Based on this mechanism, Khoddami and Cairns developed a mechanistic crosslinking method called Aza-IP. They incorporated 5-aza-C into nascent RNA by feeding cells with 5-aza-C; they then immunoprecipitated the m^5C -RMT, together with its target RNAs. Finally, the target sites can be identified by deep sequencing and computational analysis [48].

Besides the use of cytosine analogues to trap the m^5C -RMT, mutation to the methyltransferase can also stabilize the covalently-linked protein–RNA during the catalytic process. For instance, Cys321 of NSUN2 can form the covalent bond with the cytosine of the target RNA sites. Another cysteine at position 271 (Cys271) of NSUN2 is required for the release of the methylated RNA. If Cys271 is mutated to alanine, the methyltransferase cannot be released from the RNA target sites (Figure 2). By utilizing such enzyme properties, Hussain and his colleagues successfully crosslinked the RNA targets of NSUN2; the remaining strategy for

sequencing the RNA resembles those of iCLIP, therefore they termed the method methylation iCLIP (miCLIP) [47].

Towards more RNA-modifying enzymes

RNA molecules are intensively modified; yet the interactions between RNA targets and RNA-modifying enzymes are often transient. We discussed mechanistic crosslinking in the section above to study interactions between RNA targets and m^5C -RMT; both “Aza-IP” and “miCLIP” rely on the detailed knowledge of the enzymatic mechanism and the usage of either a mechanism-based inhibitor, 5-aza-C, or mutant protein with altered enzymatic action. However, successful examples towards other RNA modifications and their related enzymes are still scarce at the moment.

We will use N^6 -methyladenosine (m^6A) RNA modification as an example to formulate the challenges and opportunities in identifying, on the genomic level, protein–RNA interactions for RNA-modifying enzymes. m^6A is the most abundant endogenous modification for mRNA in eukaryotes [49]; and the recent discoveries of two novel m^6A demethylases have gained renewed interest for this long-known modification [50,51]. The fat mass and obesity-associated protein (FTO), which was first identified through genome-wide association studies to be linked to fat mass and obesity [52], shows efficient demethylation activity towards m^6A [50]. Alpha-ketoglutarate-dependent dioxygenase alkB homologue 5 (ALKBH5), another homologue protein of FTO, was later shown to be another m^6A demethylase, which impacts RNA metabolism and mouse fertility [51]. More recently, three labs

independently reported the identification of m⁶A-methyltransferases [53–55]. In addition, reader proteins of m⁶A have been suggested [56]. For example, the YTH domain family member 2 (YTHDF2) protein, one of the reader proteins, selectively recognizes m⁶A and targets the bound mRNA to decay sites, thereby affecting the translation status and lifetime of mRNA [57].

Although the genome-wide characterization of m⁶A modification in mRNA has been reported [56,58], many open questions remain for the writer, reader and eraser proteins of the m⁶A modification. PAR-CLIP experiments have been used to investigate the protein–RNA interactions for the writer and reader proteins [53,57]; yet the detailed interaction profiles for the eraser proteins FTO and ALKBH5 are still lacking. Both FTO and ALKBH5 are Fe²⁺-dependent dioxygenases; they also use the putative “base-flipping” mechanism to gain access to the target m⁶A base for demethylation reactions [59]. While existing crosslinking methods might still not be able to capture the transient interactions for these eraser proteins, new strategies that take into account their unique enzymatic mechanism could be useful in stabilizing protein–RNA interactions.

Conclusion and outlook

Formation of covalent bond between protein and RNA offers an important tool to the identification of RNA–protein interaction, as covalently-crosslinked complex can bear stringent purification to remove nonspecifically-bound RNA and protein. While UV crosslink-based methods are preferentially used for RBPs, mechanism-based crosslinking methods will be particularly useful for RNA-modifying enzymes, since these proteins only stay transiently with their RNA targets. In this review, we summarize technical aspects of several UV-enabled crosslinking strategies, including HITS-CLIP, PAR-CLIP and several derived methods with individual nucleotide-resolution. While these methods have allowed many successful studies of RNA–protein interactions, challenges still exist for more robust assays. We hope our comparisons of the existing methods could really stimulate the development of new crosslinking methods with improved crosslinking efficiency. Ultimately, more high-resolution maps of RNA–protein interactions can be revealed to understand their biological roles in greater depth and detail.

Competing interests

The authors declare no competing financial interests.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 31270838) and the National Basic Research Foundation of China (Grant No. 2014CB964900).

References

- [1] Trifillis P, Day N, Kiledjian M. Finding the right RNA: identification of cellular mRNA substrates for RNA-binding proteins. *RNA* 1999;5:1071–82.
- [2] Brooks SA, Rigby WFC. Characterization of the mRNA ligands bound by the RNA binding protein hnRNP A2 utilizing a novel *in vivo* technique. *Nucleic Acids Res* 2000;28:e49.
- [3] Mili S, Steitz JA. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 2004;10:1692–4.
- [4] Ule J, Jensen KB, Ruggiu M, Mele A, Darnell RB. CLIP identifies Nova-regulated RNA networks in the brain. *Science* 2003;302:1212–5.
- [5] Wagenmakers AJ, Reinders RJ, van Venrooij WJ. Cross-linking of mRNA to proteins by irradiation of intact cells with ultraviolet light. *Eur J Biochem* 1980;112:323–30.
- [6] Pinol-Roma S, Adam SA, Choi YD, Dreyfuss G. Ultraviolet-induced cross-linking of RNA to proteins *in vivo*. *Methods Enzymol* 1989;180:410–8.
- [7] Ule J, Jensen K, Mele A, Darnell RB. CLIP: a method for identifying protein–RNA interaction sites in living cells. *Methods* 2005;37:376–86.
- [8] Licatalosi DD, Mele A, Fak JJ, Ule J, Kayicki M, Chi SW, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 2008;456:464–9.
- [9] Guil S, Cáceres JF. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* 2007;14:591–6.
- [10] Sanford JR, Coutinho P, Hackett JA, Wang X, Ranahan W, Cáceres JF. Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. *PLoS One* 2008;3:e3369.
- [11] Becht P, König J, Feldbrügge M. The RNA-binding protein Rrm4 is essential for polarity in *Ustilago maydis* and shuttles along microtubules. *J Cell Sci* 2006;119:4964–73.
- [12] Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. *Nature* 2009;460:479–86.
- [13] Xue Y, Zhou Y, Wu T, Zhu T, Ji X, Kwon YS, et al. Genome-wide analysis of PTB–RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol Cell* 2009;36:996–1006.
- [14] Yeo GW, Coufal NG, Liang TY, Peng GE, Fu XD, Gage FH. An RNA code for the FOX2 splicing regulator revealed by mapping RNA–protein interactions in stem cells. *Nat Struct Mol Biol* 2009;16:130–7.
- [15] Zhang C, Darnell RB. Mapping *in vivo* protein–RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat Biotechnol* 2011;29:607–14.
- [16] Kishore S, Jaskiewicz L, Burger L, Hausser J, Khorshid M, Zavolan M. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat Methods* 2011;8:559–64.
- [17] Darnell RB. HITS-CLIP: panoramic views of protein–RNA regulation in living cells. *Wiley Interdiscip Rev RNA* 2010;1:266–86.
- [18] Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 2010;141:129–41.
- [19] Jungkamp AC, Stoeckius M, Mecnas D, Grün D, Mastrobuoni G, Kempa S, et al. *In vivo* and transcriptome-wide identification of RNA binding protein target sites. *Mol Cell* 2011;44:828–40.
- [20] Mukherjee N, Corcoran DL, Nusbaum JD, Reid DW, Georgiev S, Hafner M, et al. Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol Cell* 2011;43:327–39.
- [21] Lebedeva S, Jens M, Theil K, Schwanhäusser B, Selbach M, Landthaler M, et al. Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol Cell* 2011;43:340–52.
- [22] Jamonnak N, Creamer TJ, Darby MM, Schaughency P, Wheelan SJ, Corden JL. Yeast Nrd1, Nab3, and Sen1 transcriptome-wide

- binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA* 2011;17:2011–25.
- [23] Rammelt C, Bilen B, Zavolan M, Keller W. PAPD5, a non-canonical poly(A) polymerase with an unusual RNA-binding motif. *RNA* 2011;17:1737–46.
- [24] Sugimoto Y, König J, Hussain S, Zupan B, Curk T, Frye M, et al. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein–RNA interactions. *Genome Biol* 2012;13:1–13.
- [25] König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, et al. ICLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* 2010;17:909–15.
- [26] Wang Z, Kayikci M, Briese M, Zarnack K, Luscombe NM, Rot G, et al. ICLIP predicts the dual splicing effects of TIA-RNA interactions. *PLoS Biol* 2010;8:e1000530.
- [27] Tollervey JR, Curk T, Rogelj B, Briese M, Cereda M, Kayikci M, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* 2011;14:452–8.
- [28] Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D. Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol Cell* 2009;36:583–92.
- [29] Granneman S, Kudla G, Petfalski E, Tollervey D. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A* 2009;106:9613–8.
- [30] Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 2008;582:1977–86.
- [31] Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* 2007;8:479–90.
- [32] de Lima Morais DA, Fang H, Rackham OJ, Wilson D, Pethica R, Chothia C, et al. SUPERFAMILY 1.75 including a domain-centric gene ontology method. *Nucleic Acids Res* 2011;39:D427–34.
- [33] Lee I, Hong W. RAP – a putative RNA-binding domain. *Trends Biochem Sci* 2004;29:567–70.
- [34] Tsvetanova NG, Klass DM, Salzman J, Brown PO. Proteome-wide search reveals unexpected RNA-binding proteins in *Saccharomyces cerevisiae*. *PLoS One* 2010;5:e12671.
- [35] Scherrer T, Mittal N, Janga SC, Gerber AP. A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. *PLoS One* 2010;5:e15499.
- [36] Butter F, Scheibe M, Mörl M, Mann M. Unbiased RNA-protein interaction screen by quantitative proteomics. *Proc Natl Acad Sci U S A* 2009;106:10626–31.
- [37] Lindberg U, Sundquist B. Isolation of messenger ribonucleoproteins from mammalian cells. *J Mol Biol* 1974;86:451–68.
- [38] Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 2012;149:1393–406.
- [39] Baltz AG, Munschauer M, Schwanhauser B, Vasile A, Murakawa Y, Schueler M, et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell* 2012;46:674–90.
- [40] Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, et al. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res* 2011;39:D195–201.
- [41] He C. Grand challenge commentary: RNA epigenetics? *Nat Chem Biol* 2010;6:863–5.
- [42] Motorin Y, Lyko F, Helm M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res* 2010;38:1415–30.
- [43] Squires JE, Patel HR, Nusch M, Sibbritt T, Humphreys DT, Parker BJ, et al. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res* 2012;40:5023–33.
- [44] Jurkowski TP, Meusburger M, Phalke S, Helm M, Nellen W, Reuter G, et al. Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *RNA* 2008;14:1663–70.
- [45] Schaefer M, Pollex T, Hanna K, Tuorto F, Meusburger M, Helm M, et al. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev* 2010;24:1590–5.
- [46] Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, et al. RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. *Nat Struct Mol Biol* 2012;19:900–5.
- [47] Hussain S, Sajini AA, Blanco S, Dietmann S, Lombard P, Sugimoto Y, et al. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep* 2013;4:255–61.
- [48] Khoddami V, Cairns BR. Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nat Biotechnol* 2013;31:458–64.
- [49] Grosjean H. Fine-tuning of RNA functions by modification and editing. XXIV. New York: Springer; 2005, p. 442.
- [50] Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N⁶-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol* 2011;7:885–7.
- [51] Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* 2013;49:18–29.
- [52] Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;316:889–94.
- [53] Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N⁶-adenosine methylation. *Nat Chem Biol* 2014;10:93–5.
- [54] Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N⁶-methyladenosine methyltransferase. *Cell Res* 2014;24:177–89.
- [55] Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, Zhao JC. N⁶-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol* 2014;16:191–8.
- [56] Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* 2012;485:201–6.
- [57] Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N⁶-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 2014;505:117–20.
- [58] Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 2012;149:1635–46.
- [59] Yi C, He C. DNA repair by reversal of DNA damage. *Cold Spring Harb Perspect Biol* 2013;5:a012575.