Chapter 12

Genome-Wide Profi ling of RNA–Protein Interactions Using CLIP-Seq

Cheryl Stork and Sika Zheng

Abstract

UV crosslinking immunoprecipitation (CLIP) is an increasingly popular technique to study protein–RNA interactions in tissues and cells. Whole cells or tissues are ultraviolet irradiated to generate a covalent bond between RNA and proteins that are in close contact. After partial RNase digestion, antibodies specific to an RNA binding protein (RBP) or a protein-epitope tag is then used to immunoprecipitate the protein– RNA complexes. After stringent washing and gel separation the RBP–RNA complex is excised. The RBP is protease digested to allow purification of the bound RNA. Reverse transcription of the RNA followed by high-throughput sequencing of the cDNA library is now often used to identify protein bound RNA on a genome-wide scale. UV irradiation can result in cDNA truncations and/or mutations at the crosslink sites, which complicates the alignment of the sequencing library to the reference genome and the identification of the crosslinking sites. Meanwhile, one or more amino acids of a crosslinked RBP can remain attached to its bound RNA due to incomplete digestion of the protein. As a result, reverse transcriptase may not read through the crosslink sites, and produce cDNA ending at the crosslinked nucleotide. This is harnessed by one variant of CLIP methods to identify crosslinking sites at a nucleotide resolution. This method, individual nucleotide resolution CLIP (iCLIP) circularizes cDNA to capture the truncated cDNA and also increases the efficiency of ligating sequencing adapters to the library. Here, we describe the detailed procedure of iCLIP.

Key words UV crosslinking immunoprecipitation, RNA binding proteins, Immunoprecipitation, RBP–RNA complex, iCLIP

1 Introduction

RNA–protein interaction plays an important role in posttranscriptional regulation $[1]$. Currently, different methods study RNA–RBP interactions with varying resolutions. Electrophoretic mobility shift assay (EMSA) is a simple and sensitive technique to identify RNA–RBP interaction based on the observation that RNA–RBP complexes migrate slower in a native gel than the RNA alone. EMSA has many limitations including (1) possible complex disassembly and mis-assembly during electrophoresis, (2) poor sequence resolution to determine interacting nucleotides,

Ren-Jang Lin (ed.), *RNA-Protein Complexes and Interactions: Methods and Protocols*, Methods in Molecular Biology, vol. 1421, DOI 10.1007/978-1-4939-3591-8_12, © Springer Science+Business Media New York 2016

and (3) loss of in vivo cellular context $[2]$. Traditional RNA binding protein immunoprecipitation (RIP) allows for quantifiable identification of RNA targets bound to an RBP in cellular extracts. Immunoprecipitation of endogenous RNA–protein complexes using antibodies targeting the RBP is followed by purification and profiling of the bound RNA by microarray or high-throughput sequencing. When additional RBPs are co-immunoprecipitated with the assayed RBP (e.g. as part of a larger RNP complex), RIP is unable to distinguish indirectly bound RNA $\lceil 3 \rceil$. Promiscuous protein–RNA binding in cell lysates is also a significant source of false positives.

To circumvent the limitations of the previously developed methods, crosslinking immunonoprecipitation (CLIP) was developed. Crosslinking RNA–protein complexes using UV irradiation combined with immunoprecipitation has allowed us to map the RNA sequence bound to a RNA binding protein in vivo with high resolution and specificity $[4-9]$. RNA–RBP complexes are crosslinked using UV irradiation, partially digested with RNase, and purified by immunoprecipitation; the RNA is then extracted after protease digestion, ligated to an adapter, reverse-transcribed to cDNA, and sequenced. Kirk Jensen and Jenej Ule first applied CLIP to the study of RNA binding protein Nova in brain in 2003 [4]. They were able to identify 340 new RNA tags that were crosslinked to Noval, 18 of which flanked alternative exons suggesting a role of Nova in alternative splicing $[4]$. Various methods of CLIP have since been developed to increase the resolution of RNA–protein binding to the near nucleotide level allowing for the identification of not only specific binding sites but specific motifs or patterns of RNA–protein binding $[6]$.

CLIP experiments often have limited yield partly due to inefficient crosslinking using UV 254 nm. To improve RNA recovery, Thomas Tuschl's group developed photoactivatableribonucleoside- enhanced crosslinking and immunoprecipitation (PAR-CLIP). PAR-CLIP labels living cells with photoactivatable nucleosides such as 4-thiouridine (4-SU) and 6-thioguanosine (6-SG) that are incorporated into nascent RNA prior to irradiation with UV at 365 nm [8, [10\]](#page-14-0). When crosslinked, incorporation of 4-SU or 6-SG results in U to C and G to A mutations respectively, allowing researchers to use mutational analysis to identify crosslinked sites and filter out noise. An increase in crosslinking efficiency and resolution are the advantages of PAR-CLIP. However, PAR-CLIP relies on metabolic labeling which has inherent biases as well as difficulties for in vivo application.

A third variant of CLIP or iCLIP, developed by Ule's group, allows identification of crosslinking sites at the near nucleotide level (Fig. [1\)](#page-2-0) \lceil 11]. Unlike other CLIP methods that ligate adapters to RNA before reverse transcription, iCLIP introduces a circularization step after reverse transcription to ligate the cleavable 5' end

 Fig. 1 Schematic representation of iCLIP. RNA binding protein (RBP) and RNA are covalently bound in vivo using UV radiation (step 1). Potential RBP–RNA complexes are then purified together (steps 2–5). L3 linker adapter ligation to the 3' end allows for sequence-specific priming of reverse transcription and the 5' is radioactively labeled (steps 6 and 7). RBP–RNA complexes are purified from free RNA using SDS-PAGE and wet membrane transfer (step 8). Complexes are recovered from the membrane using Proteinase K (step 9). Reverse transcription truncates at the remaining polypeptide and introduces two cleavable adapter regions and a barcode (step 10). Size selection using Urea-PAGE removes the RT primer prior to circularization of the cDNA (steps 11–13). Linearization generates templates for PCR amplification (steps 14 and 15). High-throughput sequencing generates reads where the barcode sequences are immediately before the last nucleotide of the input cDNA (step 16). This nucleotide is one position upstream of the crosslinked nucleotide, allowing for identification of the RBP–RNA binding site with high resolution

adapter to the 3' end of the cDNA $[12]$. Because reverse transcriptase often stops at the crosslink sites and does not necessarily extend to the 3' adapter that is used by other CLIP methods, circularization allows iCLIP to capture these truncated cDNA. The cleavable 5' end adapter includes random barcodes to eliminate PCR artifacts. iCLIP was successfully used to globally identify RNA that were bound to HNRPNPC, TDP-43, FUS, and TIA1/TIAL1 $[12–15]$.

Despite the power of CLIP, cautions should be taken in analyzing any CLIP result in particular due to the sequence- dependent UV crosslinking biases. RBPs are multifunctional. To understand their complete functionality, other genomic methods are needed in conjunction with CLIP to globally identify their functionally relevant RNA targets.

2 Materials

12. T4 RNA Ligase I (NEB).

AGCGTCGTGgatcCTGAACCGC

3 Methods

3.1 UV-C Crosslinking

- 1. After culturing adherent cells in 10 cm plate, remove medium, add 6 mL ice-cold PBS and place on ice.
- 2. Remove lid and irradiate cells with 150 mJ/cm^2 (needs to be optimized for different proteins) at 254 nm in a Stratalinker 2400 (*see* **Note 5**).
- 3. **(Optional)** Have non-irradiated cells as a negative control of the crosslinking.
- 4. Scrape off the cells with cell lifters and transfer the cell suspension into three 2 mL microtubes. Centrifuge at $800 \times g$ at 4° C for 1 min to pellet cells, and then remove supernatant.
- 5. Snap-freeze the cell pellets on dry ice and *store at* −80°C *until further use*.

3.2 4-Thiouridine Labeling and UV-A Crosslinking (Alternative to UV-C)

- 1. Add 50 μL of 4SU (stock concentration 100 mM 4SU) to a 10 cm plate of cells culture in 10 mL DMEM supplemented with FBS, PenStrep to get a final concentration of $500 \mu M$ 4SU.
- 2. Alternatively add 10 μL of 4-thiouridine (stock concentration: 100 mM) to get a final concentration of 100 μM 4SU (*see* Note **6**).
- 3. Incubate the cells with 4-thiouridine for 60 min at 500 μ M or 8 h at 100 μM: afterward check the viability of your cells. Have cells that are not incubated with 4SU as a negative control.
- 4. Aspirate the medium and add 6 mL of ice-cold PBS to cells growing in a 10 cm plate. Remove lid and place on ice.
- 5. Irradiate twice at once with 2×400 mJ/cm² in a Stratalinker 2400 with 365 nm bulbs.
- 6. Scrape off the cells with cell lifters.
- 7. Transfer the cell suspension into three 2 mL microtubes. Centrifuge at $800 \times g$ at 4° C for 1 min to pellet cells, and then remove the supernatant.
- 8. Snap-freeze the cell pellets on dry ice and *store at* −80°C *until further use*.

1. Add 100 μL of protein G dynabeads per experiment to a fresh microtube. *3.3 Bead Preparation for Immuno-*

- 2. Wash beads twice with 900 μL lysis buffer (without protease inhibitor).
- 3. Resuspend the beads in 100 μ L lysis buffer with 2–10 μ g antibody per experiment (*see* **Note 7**).
- 4. Rotate tubes for 30–60 min at room temperature.
- 5. Wash 3 times with 900 μL lysis buffer and leave in the last wash until ready to proceed with immunoprecipitation.
- 1. Resuspend the cell pellet in 1 mL lysis buffer (with 1:100 Protease Inhibitor Cocktail Set III) and transfer to a 1.5 mL microtube. Prepare enough lysate for two experimental samples and appropriate controls (*see* **Notes 8** and **9**).
	- 2. Sonication of samples (Note: Optional).
	- 3. *Option 1*: Sonicate the sample on ice using a probe sonicator. The probe should be approximately 0.5 cm from the bottom of

precipitation

3.4 Cell Lysis and Partial RNA Digestion

the tube and not touching the tube sides in order to avoid foaming. Sonicate twice with 10 s bursts at 5 dB. Clean the probe by sonicating H_2O before and after sample treatment.

- 4. *Option 2*: Use Bioruptor plus for five cycles with alternating 30 s on/30 s off at low intensity.
- 5. Prepare two dilutions of RNase I in PBS. 1:50 for High RNase I and 1:250 for low RNAseI. This will need to be carefully optimized (*see* **Note 10**).
- 6. Add 10 μL of low or high RNAse I dilution and 2 μL Turbo DNase to the cell lysate and immediately place the samples at 37 °C for 3 min shaking at 1100 rpm. Immediately afterward transfer to ice for > 3 min.
- 7. Centrifuge for 10 min at $22,000 \times g$ and at 4 °C to clear the lysate. Carefully collect the supernatant.
- 8. *Optional*: load 500 μL of the lysate onto a Proteus Clarification Mini Spin Column. Centrifuge for 1 min at $22,000 \times g$ at 4 ^oC. Transfer flow-through to a new tube and place on ice. Repeat for the rest of the lysate and combine the fractions.

1. Add the lysate to the beads and rotate for 1 h or overnight at 4 °C. *3.5 Immuno-*

- 2. Place on magnet and discard the supernatant and wash three times with high-salt wash (rotate the third wash for at least 1 min at 4 °C) (*see* **Note 11**).
	- 3. Wash twice with PNK Buffer and then leave in 1 mL PNK Buffer and proceed to 3' end dephosphorylation step.
- 1. Discard the supernatant and resuspend the beads in 20 μL of the following mix: $4 \mu L$ 5× PNK pH 6.5 Buffer, 0.5 μL PNK, 0.5 μL RNasin, and 15 μL H_2O . *3.6 RNA 3' End Dephosphorylation*
	- 2. Incubate for 20 min at 37 °C in a thermomixer at 1100 rpm.
	- 3. Wash with PNK Buffer.

precipitation

- 4. Wash with high-salt wash (rotate wash in the cold room for at least 1 min).
- 5. Wash twice with PNK Buffer.
- 1. Carefully remove the supernatant and resuspend the beads in 20 μL of the following mix: $8 \mu L H_2O$, $5 \mu L 4 \times$ ligation buffer, 1 μL RNA ligase, 0.5 μL RNasin, 1.5 μL pre-adenylated adapter L3-App $(20 \mu M)$, and $4 \mu L$ PEG400. *3.7 L3 Adapter Ligation*
	- 2. Incubate overnight at 16 °C in thermomixer at 1100 rpm.
	- 3. Add 500 μL PNK Buffer.
	- 4. Wash twice with high-salt buffer, rotating each wash in the cold room for 5 min, and twice with PNK Buffer.

- 7. Precipitate by adding 0.75 μL glycoblue and 40 μL 3 M sodium acetate pH 5.5. Then mix and add 1 mL 100 % ethanol, mix again and place at *−*20 °C overnight.
- 8. Centrifuge at 15,000 rpm at 4 °C for 20 min. Remove the supernatant and wash the pellet with 0.9 mL 80 % ethanol and spin for 5 min (*see* **Note 15**).
- 9. Resuspend pellet in 5 μ L H₂O and transfer to a PCR tube.
- 1. Add 1 μ L primer Rt#clip (0.5 pmol/ μ L) and 1 μ L dNTP mix (10 mM) to the resuspended pellet. For each experiment or replicate, use a different Rclip primer containing individual barcode sequences (*see* **Note 16**).
	- 2. Incubate at 70 °C for 5 min. Cool to 25 °C. Hold at 25 °C until RT mix is added.
	- 3. Add RT mix to the resuspended pellet: $7 \mu L H_2O$, $4 \mu L 5 \times$ first strand buffer, 1 μ L 0.1 M DTT, 0.5 μ L RNasin, and 0.5 μ L Superscript III.
	- 4. Run RT Program: 25 °C for 5 min, 43 °C for 20 min, 50 °C for 40 min, 80 °C for 5 min, 4 °C hold.
	- 5. Add 1.65 μL 1 M NaOH and incubate at 98 °C for 20 min. Then add 20 μL 1 M HEPES-NaOH pH 7.3 to eliminate radioactivity from strongly labeled samples and to prevent RNA from interfering with subsequent reactions.
	- 6. Add 350 μL TE buffer, 0.75 μL glycoblue, and 40 μL 3 M sodium acetate pH 5.5. Mix and then add 1 mL 100 % ethanol. Mix again and precipitate at *−*20 °C overnight.
- 1. Centrifuge for 15 min at 15,000 rpm at 4 °C. Remove the supernatant and wash the pellet with 500 μL 80 % EtOH. Centrifuge again, remove supernatant, and resuspend the pellet in 6 μ L H₂O. *3.12 Gel Purifi cation of cDNA*
	- 2. Add 6 μ L 2 \times TBE-urea loading buffer to the cDNA. *Recommended*: add 6 μL loading buffer to DNA size marker.
	- 3. Heat samples to 80 °C for 5 min immediately before loading. *Optional*: leave one lane free between each sample to avoid cross-contamination.
	- 4. Prepare 800 mL of $1 \times$ TBE running buffer and fill the upper chamber with 200 mL and the lower chamber with 600 mL. Use a P1000 pipette to flush remaining urea out of the wells before loading 12 μL of each sample. Load the marker into the last lane.
	- 5. Run 6 % TBE-urea gel for 40 min at 180 V until the lower dye (dark blue) is close to the bottom.
	- 6. Cut off the last line containing the size marker and incubate it gently shaking for 10 min in 20 mL TBE buffer with 2 μL SYBR green II. Wash once with TBE and visualize by UV

3.11 Reverse Transcription

transillumination with 100 % scaling and use as a mask to guide cDNA excision from the rest of the gel.

- 7. Together the full L3-App and primer sequence accounts for 52 nt of the cDNA. Cut three bands: at 70–80 nt, 80–100 nt, and 100–150 nt.
- 8. Add 400 μL TE to each gel piece and crush gel piece with a 1 mL syringe plunger.
- 9. Incubate shaking for 1 h at 1100 rpm at $37 \degree C$, place on dry ice for 2 min, and place back for 1 h at 1100 rpm at 37 °C. Transfer the liquid portion to Costar SpinX column, into which you placed two 1 cm glass pre-filters.
- 10. Centrifuge at full speed for 1 min. collect the solution and add it together with 400 μL DNA phenol/chloroform to a 2 mL Phase Lock Gel Heavy Tube.
- 11. Incubate for 5 min shaking at 1100 rpm at 30 °C. Separate the phases by spinning 5 min at full speed at RT.
- 12. Transfer the aqueous layer into a new tube. Centrifuge again for 1 min and transfer into a new tube.
- 13. Add 1 μL glycoblue and 40 μL 3 M sodium acetate, pH 5.5, mix. Then add 1 mL 100 % ethanol. Mix again and precipitate overnight at *−*20 °C.
- 1. Centrifuge the precipitate to pellet the cDNA. Wash pellet with 500 μL 80 % ethanol. Resuspend pellet in 8 μL ligation mix $(6.5 \mu L H_2O, 0.8 \mu L 10 \times$ CircLigase buffer II, 0.4 μL 50 mM $MnCl₂$, 0.3 µL CircLigase II) in PCR tubes.
- 2. Transfer PCR tube and incubate at 60 °C for 1 h.
- 3. Add 30 μL annealing mix: 6 μL H₂O, 3 μL Fast digest buffer, 1 μL 10 μM Cut_oligo. Anneal the oligonucleotide using program 95 °C for 2 min, successive 20 s cycles starting at 95 °C and decreasing 1 °C down to 25 °C, hold at 25 °C.
- 4. Add 2 μLl BamHI and incubate for 30 min at 37 °C, then incubate for 5 min at 80 °C.
- 5. Add 350 μL TE, 0.75 μL glycoblue, and 40 μL 3 M sodium acetate, pH 5.5, and mix. Then add 1 mL 100 % EtOH. Mix again and precipitate at *−*20 °C overnight.
- 1. Centrifuge precipitate and wash with 500 μL 80 % EtOH. Resuspend in 21 μ L of H₂O.
	- 2. Prepare PCR mix: 1 μL cDNA, 0.25 μL primer, 5 μL Accuprime Supermix I, and $3.75 \mu L H_2O$.
	- 3. Run PCR reaction: 94 °C for 2 min, [94 °C 15 s, 65 °C 30 s, 68 °C 30 s] for 25–35 cycles, 68 °C for 3 min, 25 °C hold.
	- 4. Mix 8 μ L PCR product with 2 μ L 5× TBE Loading Buffer and load on a 6 % TBE gel and stain with SYBR green I (*see* **Note 16**).

3.13 Ligation of Primer to 5' End of cDNA

3.14 PCR Amplifi cation

- 1. Lysis Buffer without protease inhibitors can be made in advance. However, Protease Inhibitor Cocktail III or ANTI-RNase should be added on the day of use.
- 2. High concentrations of DTT can lower IP efficiency, so lower DTT was used to increase recovery of protein complexes.
- 3. Crosslinking with 4-thiouridine allows crosslinking with UV-A light as seen in photoactivatable-ribonucleoside-enhanced CLIP (PAR-CLIP). Addition is optional and used to enhance crosslinking of certain proteins. 4-Thiouridine is light- sensitive and cells should be placed back in the incubator immediately after addition.
- 4. Phenol/chloroform is used at a pH of 6.7 to reduce DNA– RNA hybrids in the phenol phase.
- 5. Crosslinking depends on the protein that you are working with and method used and should be optimized. However, increasing crosslinking dose could distort library preparation by damaging RNA.
- 6. Photoreactive nucleosides are toxic to cells. You should determine optimal duration of incubation prior to use.
- 7. The amount of antibody to add depends on quality and purity and will need to be optimized prior to use.
- 8. Measure protein concentration using Nanodrop 2000c or a Bradford assay. Normalize samples to lowest concentration. 2 mg/mL protein is recommended.
- 9. Recommended controls include absence of RBP in cells, absence of crosslinking, or absence of antibody during IP.
- 10. RNase I has been shown to deactivate after prolonged incubation with 0.1 % SDS in lysis buffer.
- 11. Save 15 μL from supernatant to monitor depletion efficiency.
- 12. SDS-PAGE gels changes it's pH during its run which can result in alkaline hydrolysis of RNA. The NuPAGE system keeps the pH at around 7 when used with the appropriate buffer.
- 13. To determine the specificity of the protein–RNA you will need to check the high RNase I condition. The radioactive band should be \sim 5 kDa above the MW. The negative controls should have no bands. The band in low RNase I should be diffuse. RNA is usually ~20 kDa. Cutting 15–80 kDa above the expected MW of the protein is recommended.
- 14. Try to not disturb the pellet. If you disturb the pellet, repeat spin. Leave on bench for 3 min with cap open to dry.
- 15. Use distinct Rtclip primers (Rt1clip–Rt16clip) on the controls and experimental samples. Each Rtclip contains a 4 nt barcode sequences that differs by two nucleotides to ensure that mutations cannot convert one barcode to another. Barcodes enable the user to remove PCR artifacts and control for crosscontamination during multiplexing.
- 16. It is recommended that cDNA work be done on a designated bench. Avoid taking cDNA from PCR to an area where iCLIP RNA was used.

Acknowledgements

This work was supported by the NIH grant R00 MH 096807.

References

- 1. Modic M, Ule J, Sibley CR (2013) CLIPing the brain: studies of protein–RNA interactions important for neurodegenerative disorders. Mol Cell Neurosci 56:429–435
- 2. Hellman LM, Fried MG (2007) Electrophoretic Mobility Shift Assay (EMSA) for detecting protein-nucleic acid interactions. Nat Protoc 2:1849–1861
- 3. Keene JD, Komisarow JM, Friedersdorf MB (2006) RIP-Chip: the isolation and identification of mRNAs, microRNAs, and protein components of ribonucleoprotein complexes from cell extracts. Nat Protoc 1:302–307
- 4. Ule J, Jensen KB, Ruggiu M et al (2003) Clip identifies nova-regulated RNA networks in the brain. Science 302:1212–1215
- 5. Darnell RB (2010) *HITS-CLIP*: panoramic views of protein–RNA regulation in living cells. Wiley Interdiscip Rev RNA 1: 266–286
- 6. Licatalosi DD, Aldo M, Fak JJ et al (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456:464–470
- 7. Milek M, Wyler E, Landthaler M (2012) Transcriptome-wide analysis of protein– RNA interactions using high-throughput sequencing. Semin Cell Dev Biol 23:206–212
- 8. Ascano M, Hafner M, Cekan P et al (2012) Identification of RNA–protein networks using PAR-CLIP. Wiley Interdiscip Rev RNA 3:159–177
- 9. Huppertz I, Attig J, D'Ambrogio A et al (2014) iCLIP: protein–RNA interactions at nucleotide resolution. Methods 65:274–287
- 10. Hafner M, Lanthaler M, Burger L et al (2010) Transcriptome-wide identification of RNA-

binding protein and microRNA target sites by PAR-CLIP. Cell 141:129–141

- 11. Konig J, Zarnack K, Rot G et al (2011) iCLIPtranscriptome- wide mapping of protein–RNA interactions with individual nucleotide resolution. J Vis Exp. doi[: 10.3891/2638](http://dx.doi.org/10.3891/2638)
- 12. Konig J, Rot G, Curk T et al (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 17:909–915
- 13. Tollervey JR, Curk T, Rogelj B et al (2011) Characterizing the RNA targets and positiondependent splicing regulation by TDP-43. Nat Neurosci 14:452–458
- 14. Rogelj B, Easton LE, Gireesh KB et al (2012) Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. Sci Rep. doi: [10.1038/srep00603](http://dx.doi.org/10.1038/srep00603)
- 15. Wang Z, Kayikci M, Briese M et al (2010) iCLIP predicts the dual splicing effects of TIA-RNA interactions. PLoS Biol. doi[: 10.1371/](http://dx.doi.org/10.1371/journal.pbio.1000530) [journal.pbio.1000530](http://dx.doi.org/10.1371/journal.pbio.1000530)