Methods 48 (2009) 287–293

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

CLIP: Construction of cDNA libraries for high-throughput sequencing from RNAs cross-linked to proteins *in vivo*

Zhen Wang [a], James Tollervey [a], Michael Briese [a], Daniel Turner [b], Jernej Ule [a,*]

[a] MRC-Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK
[b] Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK

ARTICLE INFO

Article history:
Accepted 26 February 2009
Available online 9 March 2009

Keywords:
CLIP
UV cross-linking
Immunoprecipitation
RNA–protein binding
High-throughput sequencing

ABSTRACT

UV cross-linking and immunoprecipitation assay (CLIP) can identify direct interaction sites between RNA-binding proteins and RNAs *in vivo*, and has been used to study several proteins in tissues and cell cultures. The main challenge of the method is to specifically amplify the low amount of isolated RNA. The current protocol is optimised for efficient RNA purification and ligation of barcoded RNA adapters. High-throughput sequencing of the multiplexed cDNA library allows for a comprehensive coverage of the target sequences.

1. Introduction

RNA-binding proteins (RBPs) play important roles in every step of gene expression, with individual proteins often binding to hundreds of different RNA targets. In many cases the manner in which RBPs interact with individual RNAs is well understood, but little is known about the full spectrum of their RNA targets. In addition to identifying the RNA targets of an RBP, the information on the position of protein–RNA interaction is also required in order to fully understand the function of the protein.

UV cross-linking was first used in combination with immunoprecipitation and SDS–PAGE analysis of the protein–RNA complex to study the alterations in an IRE1–RNA complex in the mammalian unfolded protein response [1]. This study showed that the amount of RNA associated with IRE1 as revealed by end-labelling with T4 poly-nucleotide kinase (PNK) was greater in IRE1-containing complexes isolated from stressed cells. A method employing immunoprecipitation and SDS–PAGE to purify the RNA cross-linked to a specific protein, followed by isolation of the RNA with proteinase digestion and amplification using RNA adapter ligation, as was initially developed for cloning of small interfering RNAs [2], was termed UV cross-linking and immunoprecipitation (CLIP) [3,4]. CLIP method was instrumental in understanding the RNA map, or the ability of the splicing factor Nova to both enhance and inhibit inclusion of alternative exons and poly-A sites in a position-dependent manner [3,5,6]. It has also implicated hnRNP A1 protein in regulation of pre-miRNA processing [7], identified new nuclear and cytoplasmic targets of the SR protein SF2/ASF [8,9] and defined the RNA map for the activity of FOX2 [10]. These studies show that CLIP can be used to define the RNA interactions of functionally diverse RBPs, allowing for a better understanding of these interactions in different biological processes.

The original CLIP method involved the immunoprecipitation and SDS–PAGE purification of the RNA targets of a specific protein of interest, followed by their sequential ligation to 5′ and 3′ RNA adapters [3,4]. The potential disadvantages of this method were loss of RNAs due to self-ligation, contamination with bacterial RNAs from the ligase and other recombinant enzymes, as well as formation of RNA adapter concatamer RNA products. To overcome this, an improved CLIP method allowed for the ligation of 3′ adapter to the RNA while it is still bound to the protein on beads during immunoprecipitation [11]. Here however, we describe a protocol that performs 5′ adapter ligation on beads, thus increasing overall ligation efficiency, as ligation of the 5′ adapter is more efficient than that of the 3′ adapter ([12] and unpublished observation of the authors) (Fig. 1). In addition, this protocol allows for monitoring of the ligation efficiency via the amount of incorporated radioactivity.

The CLIP method uses ribonuclease to partially digest the RNA and generate 40–80 nucleotide lengths of RNA with 5′ hydroxyl and 3′ phosphate groups. Neither of these end groups are compatible with ligation to RNA adapters, therefore, to be able to ligate the RNA to the 5′ adapter, the 5′ end must be phosphorylated by PNK. However, most commercial PNK enzymes also have a 3′ phosphatase activity, which generates RNAs with 5′ phosphate and 3′ hydroxyl groups that have a tendency to self-ligate. To avoid this, our current protocol uses a truncated form of PNK that lacks the 3′ phosphatase activity [13]. The truncated PNK generates RNA

* Corresponding author.
E-mail address: jule@mrc-lmb.cam.ac.uk (J. Ule).

1046-2023/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2009.02.021

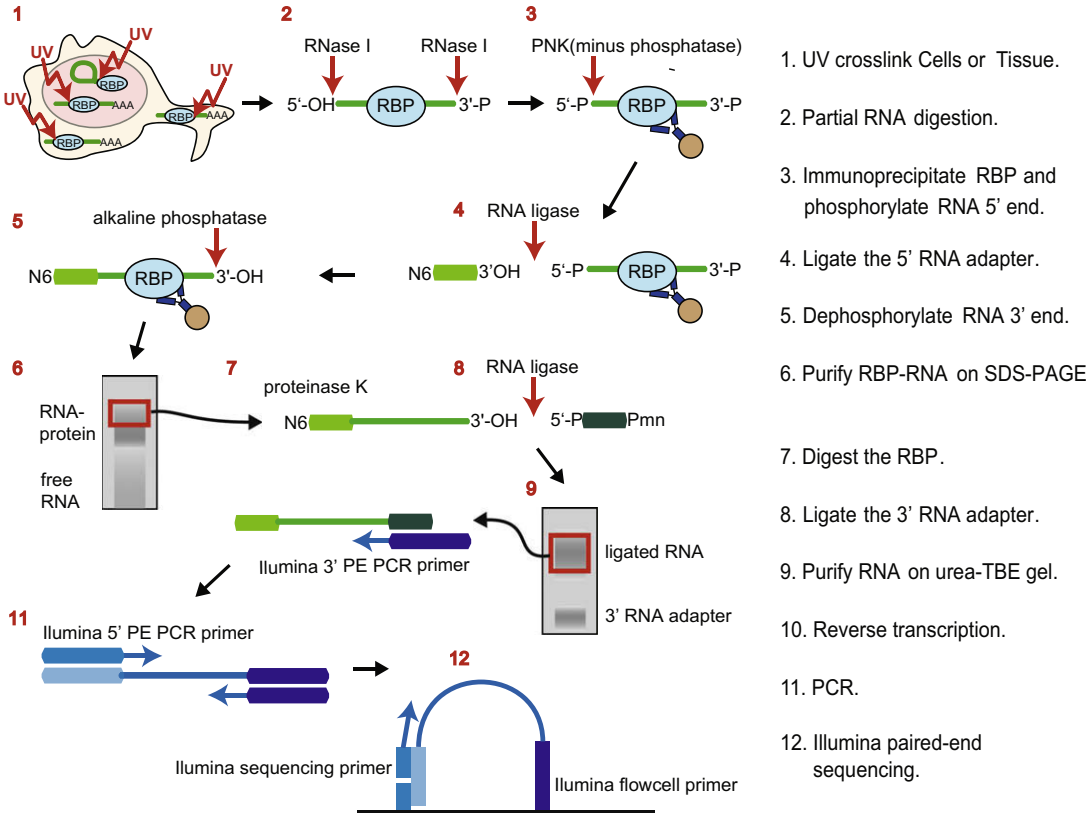

Fig. 1. A flowchart of the CLIP method. Dissociated cells or tissues are UV cross-linked on ice, leading to formation of a covalent bond between protein and RNA. After lysis, the RNA is partially digested, and protein–RNA complex is purified by immunoprecipitation. The RNA is radioactively labelled with 5′ ^{32}P-γ-ATP and ligated to 5′ adapter. The 5′ adapter has a 5′-Amino (6-carbon) linker (N6) to prevent it from being phosphorylated by contaminating PNK. The protein–RNA complex is separated by SDS–PAGE and transferred to nitrocellulose membrane. The membrane is exposed to X-ray film and the region of the membrane corresponding to the protein–RNA complex is excised. Protein is digested by proteinase K and 3′ adapter is ligated to free RNA. The 3′ adapter has a 3′ puromycin and 5′ phosphate to allow its ligation to the CLIP RNA while preventing self-ligation. The RNA is gel purified and amplified by RT-PCR with primers complimentary to the RNA adapters. The 5′ end of the primers contains the sequence required for hybridisation to the Illumina Genome Analyser flowcell. The 3′ end of the primers is identical to the Illumina Genome Analyser sequencing primers.

with phosphate residues at both ends which cannot self-ligate, but can efficiently ligate to the 5′ RNA adapter.

The CLIP cDNA library was initially sequenced by cloning and Sanger sequencing [3,11], whereas recently the high-throughput sequencing methods of 454 FLX systems and Illumina Genome Analyser have been used for a more comprehensive coverage of the library [6,9,14]. The cDNA prepared by the current protocol was successfully sequenced using the Illumina Genome Analyser, but it could be used with any other sequencing method. The RNA adapters used are complementary to the primers used for paired-end sequencing via Illumina Genome Analyser. The adapters also contain two additional nucleotides, which we term 'barcodes', that in principle allow for multiplexing of over hundred experiments in a single lane of the flowcell. Using the Illumina GA2 technology, each lane in the flowcell generates approximately 7 million pairs of sequences. Thirty-six nucleotide long sequences are generated from the 5′ and 3′ end of the sequence. CLIP sequences are 40–70 nucleotides long, therefore most Illumina reads cover all of the CLIP sequence. We usually identify between 10,000 and 100,000 diverse genomic loci in each CLIP experiment, therefore in our hands multiplexing of up to ten experiments per lane can generate sufficient number of diverse sequences per experiment.

2. Material and reagents

(1) UV cross-linking: cold PBS, Stratalinker (we use model 2400), 10 cm tissue culture dishes.

(2) Partial RNA digestion and immunoprecipitation:
- (a) Lysis buffer: 50 mM Tris–HCl, pH 7.4; 100 mM NaCl; 1mM $MgCl_2$; 0.1 mM $CaCl_2$; 1% NP-40; 0.5% sodium deoxycholate; 0.1% SDS; protease inhibitor and ANTI-RNase added fresh.
- (b) High-salt wash: 50 mM Tris–HCl, pH 7.4; 1 M NaCl;1 mM EDTA; 1% NP-40; 0.5% sodium deoxycholate; 0.1% SDS.
- (c) PNK wash: 20 mM Tris–HCl, pH 7.4; 10 mM $MgCl_2$; 0.2% Tween-20.

Protein A Dynabeads (Dynal, 100.02), protease inhibitor cocktail (Calbiochem, 535140), ANTI-RNase (Ambion, AM2692), RNase I (Ambion, AM2295), Turbo DNase (Ambion, AM2239), 10× PNK buffer and PNK (3′ phosphatase minus) (NEB, M0236L), P^{32}-γ-ATP, 10 mM ATP, Magnetic stand (Invitrogen) and Thermomixer R (Eppendorf).

(3) Linker ligation and dephosphorylation: Tween-20, 10× RNA ligase buffer and RNA ligase (Takara, 2050A), 50% PEG 6000 (Sigma, 81260), 20× BSA (Takara, 2050A), 10× FastAP buffer and FastAP (Fermentas, EF0652), 0.5% Tween-20.

(4) SDS–PAGE and nitrocellulose transfer: Novex NuPAGE 4–12% Bis–Tris gel (Invitrogen, NP0321), 20× MOPS Novex NuPAGE running buffer (Invitrogen, NP0001), nitrocellulose membrane (Whatman, Protran), Novex wet transfer apparatus (Invitrogen), 20× Transfer buffer (Invitrogen, NP0006-1), methanol.

(5) RNA isolation:

(a) PK buffer: 100 mM Tris–HCl, pH 7.4; 50 mM NaCl; 10 mM EDTA.

(b) PK + urea buffer: 100 mM Tris–HCl, pH 7.4; 50 mM NaCl; 10 mM EDTA, 7 M urea.
Proteinase K (Roche, 1373196), RNA phenol/chloroform (Ambion, 9722), glycoblue (Ambion, 9510), 3 M sodium acetate, pH 5.5 (Ambion, AM9740), 1:1 ethanol:isopropanol, 70% ethanol.
(6) RNA size purification:
(a) Eight percentage denature polyacrylamide gel mix: 8% 19:1 Acrylamide:Bis-acrylamide (Severn Biotech, 20-2400-10); 7 M urea; 10× TBE; Rhinoide Polyacrylamide Gel Strengthener (Molecular probes, R33400).
(b) Two times loading buffer: 95% formamide, deionised (Sigma, F9037); 5% 100 mM EDTA, pH 8; Bromophenol blue + Xylene cyanol (Sigma, B3269).
(c) RNA elution buffer: 1 M sodium acetate, pH 5.5 (Ambion, AM9740); 1 mM EDTA.
Ten percentage APS (Sigma, A3678), TEMED (Sigma, T9281), 1× TBE, phiX markers (Promega, E3511), GelBond PAG Film (Lonza, 54711), 1 ml syringe plunger, Costar SpinX column (Corning, 8161), 1 cm glass pre-filter (Whatman, 1823010), Glycoblue (Ambion, 9510), 1:1 ethanol:isopropanol, 70% ethanol, DNase and 10× buffer (Ambion, AM1906).
(7) RT-PCR: Qiagen OneStep RT-PCR kit (Qiagen, 210210), ExoSAP-IT (GE Healthcare, US78200), G-25 column (GE Healthcare, 27-5325-01).
(8) RNA adapters: If using high-throughput sequencing methods, such as Illumina Genome Analyser, then the sequence of the 5′ RNA adapter should be identical to the last 20 nucleotides of the 5′ sequencing primer, and the sequence of 3′ RNA adapter should be reverse complementary to the last 20 nucleotides of the 3′ sequencing primer. To allow multiplexing during sequencing, each adapter could contain two additional nucleotides at the end, and several of each 5′ and 3′ adapters can be used in different combinations. The 5′ adapter should contain a 5′-Amino (6-carbon) linker, and the 3′ adapter should contain 5′ phosphate and 3′ puromycin. We order the RNA adapters from Dharmacon.
(9) Primer sequences: The 5′ amplification primer should contain sequence of the 5′ RNA adapter without the last two barcode nucleotides. The 3′ amplification primer should be reverse complementary to the sequence of the 3′ RNA adapter without the last two barcode nucleotides. At their 5′ end, the amplification primers should contain the sequence necessary for anchoring to the sequencing flowcell. This approach has been tested with Illumina Genome Analyser, but should be applicable to any other high-throughput sequencing approach.

2.1. Comment

The ANTI-RNase inhibitor used during the lysis step inhibits RNAse A, the predominant RNAse in mammalian tissues, but does not inhibit RNAse I. This allows for more standardised RNAse conditions to be applied to diverse biologic source materials.

In general, we recommend using polyclonal antibodies for immunoprecipitation. The most important factor for efficient immunoprecipitation is high avidity, which most often requires recognition of multiple epitopes on the protein. Therefore, mouse monoclonal antibodies are rarely efficient, unless several monoclonal antibodies against different epitopes are mixed. In this case, or in case of mouse polyclonal antibodies, we recommend using Dynabeads® M-450 Sheep anti-Mouse IgG instead of the protein A beads, which can be used with rabbit polyclonal antibodies.

Both RNA adapters contain a modification that blocks self-ligation or formation of linker concatemers. We use amino linker (N6) on the 5′ end of the 5′ adapter, and puromycin on the 3′ end of the 3′ adapter (Fig. 1), but other modifications can be used instead of these.

3. Protocol

3.1. UV cross-linking of tissue/cells

For mouse tissue:

(a) Harvest tissue and cover with cold PBS.
(b) Add 10 cell volumes of cold PBS and partially dissociate the tissue by triturating using a 5 or 10 ml pipette.
(c) Add a 200 μl pipette tip to the end of 5 or 10 ml pipette, and further dissociate the tissue by passing through the tips several times.
(d) Transfer to 10 cm tissue culture dish and place on a tray with ice.
(e) With lid removed, irradiate suspension 3 times on ice for 100 mJ/cm^2 in Stratalinker. Mix suspension between each irradiation.
(f) Split cell suspension into 2 ml tubes, and centrifuge at maximum speed in a table top centrifuge for 3 min.
(g) Remove supernatant, snap freeze pellets on dry ice, and store at −80 °C until further use.

For cell culture:

(a) For adherent cells, grow cells on 10 cm tissue culture dishes until confluent.
(b) Remove media, add 6 ml of cold PBS and place in Stratalinker on ice with the lid off. Irradiate once for 150 mJ/cm^2.
(c) For cells grown in suspension, place the Petri dish in Stratalinker on ice without the lid. Irradiate once for 150 mJ/cm^2.
(d) Scrape off the cells and split into three 2 ml tubes. Centrifuge at maximum speed in a table top centrifuge for 3 min.
(e) Remove supernatant, snap freeze pellets on dry ice, and store at −80 °C until further use.

3.1.1. Comment

We generally use 20 mg of tissue, or if using cell culture, one third of a 80% confluent 10 cm dish (5–10 mg of cell pellet) per experiment. It is not necessary to disrupt the tissue into single cells for tissue cross-linking, as UV light can penetrate a few cell layers. The amount of UV light used might need to be optimised for some proteins, since each RNA-binding domain cross-links with different efficiency, depending on the availability of aromatic amino acids. For a preliminary experiment, try a few different energy levels (100, 200 and 400 mJ/cm^2), and then use the least amount of energy that gives >70% of the maximum signal. For adherent cell cultures, a single cross-linking with UV light is sufficient as they grow as a monolayer.

3.2. Partial RNA digestion and immunoprecipitation

Bead preparation:

(a) For each experiment use 100 μl of protein A Dynabeads. Resuspend the beads and transfer 100 μl to a non-sticky 1.5 ml tube.
(b) Wash beads 3 times with lysis buffer.
(c) Resuspend in 200 μl of lysis buffer and add 1–5 μg of antibody depending on the efficiency of specific antibodies.
(d) Rotate at 4 °C for 30–60 min.
(e) Wash 3 times with lysis buffer and leave beads in last wash until you are ready to add the lysates.

Partial RNA digestion and immunoprecipitation:

(a) Resuspend each pellet of cross-linked material in 1 ml lysis buffer. Add 10 μl protease inhibitor and 1 μl ANTI-RNase.
(b) Sonicate on ice until cells or tissues are dissociated. Avoid foaming of the lysates.
(c) Optional: make a dilution of RNase I at 1:50 in lysis buffer (high-RNase).
(d) Make a dilution of RNase I at 1:500 in lysis buffer (low-RNase).
(e) Add 10 μl RNase dilution and 5 μl Turbo DNase per 1 ml of lysates. Incubate at 37 °C for 3 min.
(f) Centrifuge at maximum speed in a table top centrifuge for 3 min.
(g) Carefully collect the supernatant and add it to the antibody-conjugated beads.
(h) Rotate at 4 °C for 1 h or overnight.
(i) Discard the supernatant. Wash beads 2 times with high-salt buffer and 2 times with PNK buffer.
(j) Prepare 80 μl 1× PNK/Tween from 10× NEB stock with a 0.1% final concentration of Tween-20.
(k) Dilute PNK (3′ phosphatase minus) 1:10 in PNK/Tween.

5′ phosphorylation:

(a) Wash beads once with 50 μl of PNK/Tween. Leave on ice and remove the buffer just before adding the PNK mix.
(b) Add 8 μl of hot PNK mix to the beads:
- 5.5 μl PNK/Tween
- 1.5 μl ^{32}P-γ-ATP
- 1 μl diluted PNK (3′ phosphatase minus)

(c) Incubate in a Thermomixer at 37 °C for 3 min at 1000 rpm.
(d) Add 20 μl of cold PNK mix:
- 18 μl PNK/Tween
- 2 μl 10 mM ATP

(e) Incubate in Thermomixer R at 37 °C for another 3 min, 1000 rpm.
(f) Place on magnetic stand and collect the supernatant as radioactive waste.
(g) Cool the Thermomixer to 16 °C. Wash beads 2 times with 0.7 ml high-salt buffer and collect the supernatant as radioactive waste (as well as all remaining supernatants up to the SDS–PAGE gel step).
(h) Wash 2 times with 0.9 ml PNK buffer.
(i) Make 30 μl of 1× Ligation buffer (from 10× stock) with 0.1% Tween-20. Wash with 30 μl Ligation buffer/Tween.

3.2.1. Comment

The current protocol uses RNase I, as this RNase has no nucleotide bias. This avoids any sequence biases that could be introduced by other RNases. The protocol also uses Turbo DNase instead of standard DNase I, since Turbo DNase is active in conditions of up to 200 mM NaCl. Use of PNK (3′ phosphatase minus) generates RNA with 5′ phosphate and 3′ phosphate groups, which can ligate to the 5′ RNA adapter, but cannot self-ligate.

When adding the final PNK/Tween and Ligation buffer/Tween buffers, it is important to dissociate the beads properly by pipetting up and down. Beads tend to stick together after washing with PNK buffer. If doing the preliminary experiments, or for the high-RNase samples, then skip the cold PNK mix step and the RNA adapter ligation step and go directly to SDS–PAGE step.

3.3. 5′ adapter ligation and dephosphorylation

(a) Remove the 1× Ligation buffer/Tween.
(b) Resuspend the beads in 5 μl L5 adapter (20 μM). Use different L5 adapters (containing different nucleotide barcodes) for different experiments.
(c) Add 10 μl of the following mix:
- 7.5 μl 50% PEG6000
- 1.5 μl 10× RNA Ligation buffer
- 0.75 μl BSA
- 0.25 μl RNA ligase

(d) Incubate in cooled Thermomixer R at 16 °C for 2 h or overnight with intermittent shaking.
(e) Wash once with 0.9 ml PNK buffer.
(f) Add 30 μl of phosphatase mix:
- 3 μl 10× FastAP buffer
- 6 μl 0.5% Tween-20
- 20 μl dH_2O
- 1 μl FastAP

(g) Incubate in Thermomixer R at 37 °C for 20 min, 1000 rpm.
(h) Wash 2 times with 0.9 ml PNK buffer.
(i) Add 25 μl of 1× NuPAGE loading buffer (dilute from 4×) to the beads, incubate in Thermomixer R at 70 °C for 10 min, 1000 rpm.
(j) Place on magnetic stand and collect eluate.

3.3.1. Comment

PEG 6000 is used in the ligation mixture because we found it increases ligation efficiency dramatically (up to a hundred fold). If performing several experiments with the same antibody, then different 5′ adapter can be used for each experiment, and after ligation to the adapter, the experiments can be mixed, loaded on one single lane of the gel, amplified and sequenced in the same lane of the flowcell. The barcoding later allows identification of sequences specific for different experiments.

3.4. SDS–PAGE electrophoresis and nitrocellulose transfer

(a) Load a 9 or 10 well Novex NuPAGE 4–12% Bis–Tris gel. Use 500 ml 1× MOPS running buffer.
(b) Run gels at 200 V, and afterwards, transfer to nitrocellulose membrane using Novex wet transfer apparatus.
(c) Rinse the membrane in 1× PBS, and gently blot dry. Wrap in plastic wrap and expose to autoradiogram.

3.4.1. Comment

The use of Novex NuPAGE gels is critical. A pour-your-own SDS–PAGE gel can have a pH up to 9.5 during the run, which will lead to alkaline hydrolysis of the RNA. The Novex NuPAGE buffer system is close to pH7. During nitrocellulose transfer, most of the free RNA will pass through the membrane, whereas the protein–RNA complex would be transferred to the membrane. The membrane will be 10–100 times less radioactive than the samples loaded on the gel.

Analysis of the control experiments (see point 6) is crucial in determining how to excise the protein–RNA complex. The high- and low-RNase and the negative control without the protein in starting material allow determination of which band is the specific protein–RNA complex, whether the band representing bound RNA shifts up following low-RNase treatment, and how close to the specific band are any other non-specific bands. If the specific band does not shift up with low-RNase, then do not use any RNase in future experiments, and for cloning excise the whole radioactive band to purify the short RNAs.

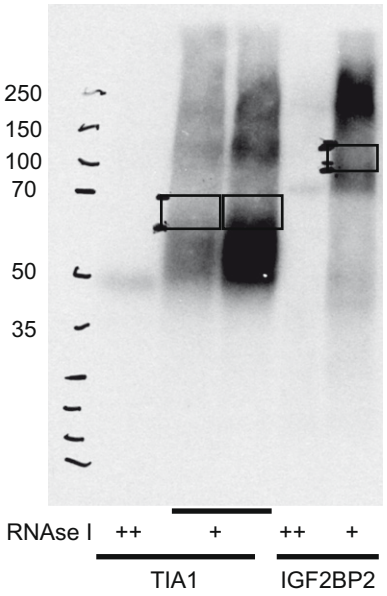

Fig. 2. Analysis of CLIP protein–RNA complexes separated by SDS–PAGE. Extract from embryonic day 18 mouse brain was prepared with high (++) or low (+) RNase I conditions. Two RBP antibodies were used for immunoprecipitation from both conditions. Anti-TIA1 antibody (Proteintech, 12133-2-AP) gave a band at around 45 kDa with high-RNase, and the radioactively labelled protein–RNA complexes shifted up with low-RNase. To prepare the cDNA library, the protein–RNA complexes of 65–75 kDa were isolated from the membrane. Immunoprecipitation with anti-IGF2BP2 antibody (Proteintech, 11601-1-AP) gave a band at around 75 kDa with high-RNase, and the protein–RNA complexes of 95–110 kDa were isolated from the membrane of the low-RNase experiment.

If the band shifts up with low-RNase, then use low-RNase as the condition for CLIP experiment. Based on the control experiments, calculate the distance of the specific band from the closest contaminating band. If your protein band migrates less than 10 kDa away from other contaminating protein band, then it will be difficult to isolate specific RNA targets—in this case, it is worth spending more effort in optimising immunoprecipitation (increasing the stringency by raising the amount of salt and detergents during the washes). If the contaminating bands are between 10 and 30 kDa away from the specific band, then to isolate specific RNA, a thin band should be excised at 20–25 kDa above the expected MW of the protein of interest, and the specific RNA should be further purified on the RNA gel (see Section 3.6.1). If no contaminating bands are present on the gel, or if they are far from the specific band, then a wider band can be excised from the gel, from 20 up to 50 kDa above the MW of the protein of interest (Fig. 2). Bands less than 20 kDa above the MW of the protein of interest should not be cut, since the insert RNA will be short and less informative. RNAs isolated from a protein–RNA complex that migrates 20 kDa above the MW of the protein have an average size of 65 nucleotides [11]. Taking away 22 nucleotides coming from the RNA adapter, the sequence of interest would in this case span 43 nucleotides, which is the minimum length we find useful for unique mapping to the genome.

3.5. RNA isolation

(a) Analyse the exposure, and use the high-RNase sample to determine the specificity of the protein–RNA complexes (Fig. 2). See the comment above to decide where to cut out a band corresponding to the protein–RNA complex. The ideal size of the protein–RNA complexes of interest in low-RNase experiment is at around 20 kDa above the expected molecular weight of your protein.
(b) Use a clean scalpel blade and put the piece of membrane into a 1.5 ml tube.
(c) Make a 2 mg/ml proteinase K solution in PK buffer. Incubate at 37 °C for 5 min.
(d) Add 200 μl of proteinase K solution to each tube of isolated membrane and incubate at 37 °C for 20 min.
(e) Transfer the solution into a fresh 1.5 ml tube.
(f) Add 200 μl of PK/7M urea buffer, wash the membranes by pipetting up and down, and join with the proteinase K solution.
(g) Incubate the joint solution at 60 °C for 20 min.
(h) Let the solution to cool to 37 °C and add 540 μl RNA phenol/chloroform. Incubate in Thermomixer R at 37 °C for 5 min at 1100 rpm.
(i) Centrifuge for 3 min at maximum speed in a table top centrifuge at room temperature.
(j) Take 300 μl of the aqueous phase and transfer to a fresh 1.5 ml tube.
(k) Add 1 μl of glycoblue and 50 μl of 3 M sodium acetate, pH 5.5. Mix well. Add 1 ml of 1:1 ethanol:isopropanol, mix well and precipitate at −20 °C for 20 min or overnight.

3.5.1. Comment

RNA phenol can also be prepared by equilibrating pure phenol with 0.15 M sodium acetate, pH 5.5; Chloroform is 49:1 with isoamyl alcohol. Glycoblue is necessary to precipitate the small amount of RNA, and the blue colour makes it easier to visualise the pellet. Potassium acetate, pH 5.5 can also be used instead of sodium acetate, pH 5.5.

3.6. 3′ adapter ligation and size purification of ligated RNA

(a) Centrifuge the precipitated RNA at maximum speed in a table top centrifuge at 4 °C for 10 min. Wash the RNA pellet twice with 0.5 ml 70% Ethanol. Dry the pellet with open lid for 2 min.
(b) Resuspend the beads in 2.6 μl L3 adapter (50 μM). Use different L3 adapters for different experiments. Transfer to PCR tubes.
(c) Add 5.4 μl of the following mix:
  (a) 4 μl 50% PEG6000
  (b) 0.8 μl 10× Ligation buffer
  (c) 0.4 μl BSA
  (d) 0.2 μl RNA ligase
(d) Incubate in PCR machine at 16 °C for 2 h or overnight.
(e) Add 10 μl 1× TBE and 18 μl of 2× formamide loading buffer.
(f) Heat in PCR machine at 90 °C for 2 min, then 20 °C for 2 min.
(g) Pour an 8% denature urea polyacrylamide gel by mixing 20 ml gel mix with 200 μl 10% APS and 10 μl TEMED.
(h) Preheat the gel by running for 20 min at 450 V.
(i) Load samples on the gel together with radioactively labelled phiX markers.
(j) Run at 450 V until the lower dye runs to the bottom.
(k) Support the gel on GelBond PAG Film, wrap in Saran wrap and autorad the gel at −70 °C. Most cases exposure overnight is necessary.

3.6.1. Comments

A control ligation with no RNA input (just water) can be set up the same time as L3 ligation, and this can be continued throughout the rest of the protocol. When excising from the gel (see next step), excise at the same size as other reactions, even though there will be no radioactivity. Sometimes, if the amount of RNA is extremely low, one may not be able to detect any radioactivity. In this case,

just excise from appropriate size of the RNA on the gel based on the position of radioactive molecular markers.

In cases where contaminating bands were present on the membrane after transfer from the SDS–PAGE gel, the RNA gel can serve to further purify the specific RNA, since the non-specific RNA will migrate at a different size from the specific RNA. If the band was cut between 20 and 25 kDa above the MW of the protein of interest, then the isolated RNA was of 65–80 nucleotides in length. After ligation of the 3′ RNA adapter, the resulting RNA is of 85–100 nucleotides in length, and this is the region of the RNA gel where the RNA should be excised. However, if no contaminating bands were present on the SDS–PAGE gel, and a wider band was excised (20–50 kDa), then a wider RNA band of 85–165 nucleotides can be excised from the RNA gel. Use the PhiX labelled MW markers as a guideline for RNA size determination.

3.7. RNA gel extraction and DNase treatment

(a) Cut out a band from the gel corresponding to RNA size of 85–110 nucleotides (see comments above) and place in a 1.5 ml tube.
(b) Add 300 μl of RNA elution buffer and crush with a 1 ml syringe plunger. Incubate in Thermomixer R at 37 °C for 2 h, 1100 rpm.
(c) Transfer the supernatant to a new 1.5 ml tube. Add another 100 μl of elution buffer to the gel, incubate at 37 °C for 30 min, 1100 rpm.
(d) Join the two supernatants.
(e) Place two 1 cm glass pre-filter onto a Costar SpinX column, add the supernatant and spin at maximum speed in a table top centrifuge for 30 s.
(f) Collect the flow through, and add to another Costar SpinX column with one 1 cm glass pre-filter. Spin at maximum speed in a table top centrifuge for 30 s.
(g) Collect the flow through, and add 1 μl glycoblue. Mix well before adding 1 ml of 1:1 ethanol:isopropanol.
(h) Precipitate at −20 °C for 20 min or overnight.
(i) Centrifuge at maximum speed in a table top centrifuge at 4 °C for 10 min. Wash the RNA pellet twice with 0.5 ml 70% ethanol. Dry the pellet with open lid for 2 min.
(j) Dissolve the RNA is 20 μl of the following mix:
- 18 μl water
- 2 μl 10× DNase buffer
- 0.4 μl DNase

(k) Incubate at 37 °C for 10 min. Then add 2 μl of DNase inactivation solution, incubate for 2 min on Thermomixer at 1100 rpm.
(l) Centrifuge at maximum speed in a table top centrifuge for 30 s, and collect supernatant into a fresh 1.5 ml tube.
(m) Add 350 μl RNA elution buffer and 1 μl glycoblue. Mix well before adding 1 ml 1:1 ethanol:isopropanol. Precipitate at −20 °C for 20 min or overnight.
(n) Centrifuge at maximum speed in a table top centrifuge at 4 °C for 10 min. Wash the RNA pellet twice with 0.5 ml 70% ethanol. Dry the pellet with open lid for 2 min.
(o) Resuspend the pellet in 16 μl of 0.7 μM 3′ PE Illumina PCR primer and proceed to RT-PCR.

3.7.1. Comment

When extracting RNAs from acrylamide gels, it is very important to note the size of the band being excised. If the band measures 1–2 cm, then add 450 μl of RNA elution buffer, and subsequently add 1.5 μl of glycoblue and 1.5 ml of 1:1 ethanol:isopropanol. If the band is over 2 cm in size, then add 700 μl of RNA elution buffer, then spit into two tubes for precipitation. It is very important to add the glass pre-filter to the column as residual acrylamide can interfere with RNA precipitation. If no pellet is visible when spinning after 20 min of precipitation, then overnight precipitation might be more efficient.

The DNase step ensures that no contaminating DNA from the samples is amplified. Especially when repeating CLIP experiments, the cDNAs and PCR products from previous experiments may be amplified as contaminants, thus it is very important to use a different barcode to distinguish different experiments.

3.8. RT-PCR

(a) RT: 95 °C, 2 min; 50 °C, 20 min[1]; 55 °C, 10 min; 4 °C, hold:
4 μl 5× Qiagen OneStep RT-PCR buffer
0.8 μl Qiagen OneStep RT-PCR Enzyme Mix
0.8 μl 10 mM dNTP Mix
(b) Add 1 μl 10 μM 5′ PE Illumina PCR primer to the mix.
(c) PCR: 95 °C, 15 min; 35 cycle of 94 °C, 15s; 65 °C, 10s; 72 °C, 10s; 72 °C, 3 min.
(d) Part of the PCR product can be analysed on a gel or on Agilent DNA chip and store the rest 4 °C.
(e) Mix the PCRs that will be sequenced together in one lane. Add 0.5 μl of ExoSAP-IT and incubate at 37 °C for 15 min, 80 °C for 15 min and 20 °C for 1 min.
(f) Pass through G-25 column to desalt and submit for Illumina high-throughput sequencing.

3.8.1. Comment

We use primers sequences compatible with both the paired-end (PE) Illumina high-throughput sequencing protocol. The 5′ end of the primers contains the sequence required for anchoring to the Illumina Genome Analyser flowcell. The 3′ end of the primers is identical to the Illumina Genome Analyser primers. The sequences amplified by these primers can either be sequenced from only a single end (SE) or both ends (PE). Current GA2 protocols can generate over 50 nucleotide sequences with SE sequencing, which is sufficient for the CLIP method—however, in some cases sequence from both ends is required, which can be achieved with PE sequencing.

The number of PCR cycles needs to be optimised for each experiment. It primarily depends on the amount of starting RNA, and that depends on the efficiency of the RNA–protein cross-linking by the UV light, and the amount of starting tissue or cells used. For initial experiments, 30, 35 and 40 cycles can be done, and the minimal number of cycles necessary to obtain a product is then used for further experiments. Due to the high amplification process, one needs to be very careful when dealing with the resulting PCR products, and perform all work with these in an area separate from where CLIP experiments are done. We use a PCR workstation for this purpose. The current bar coding system in the RNA adapters can serve to identify potential contaminant sequences from previous experiments.

4. Controls

The use of two types of negative controls and a positive control is recommended. These controls can be performed only during the initial experiments, where the RNA is radioactively labelled with PNK during immunoprecipitation and then the complex is loaded

[1] When the temperature reaches 50 °C, add 5.6 μl of the following mix into the tubes.

on the gel (that is, skipping ligation to 5′ adapter and dephosphorylation). The control signals allow determination of whether the purified protein is specific, and whether it cross-links to long or short RNA species.

4.1. First negative control

The first aspect to be tested with CLIP is the specificity of immunoprecipitation. Initially, this can be done by Western blot analysis of the tissue or cell extract before and after incubation with beads/antibody. If the immunoprecipitation is efficient, then a decrease in the amount of RNA-binding protein in the depleted extract should be detected on Western blot. After this is confirmed, a CLIP experiment should be set up with starting material with or without the protein to be purified. This can be tissue from knockout mice or knockdown cells. If this is not possible, then for the purpose of testing the antibody, a tissue or cell line that normally does not express the protein can be used. This is compared to a tissue that expresses the protein, or a cell line where the protein is overexpressed. The radioactive signal of the RNA–protein complex should be absent when using the material that lacks the protein.

4.2. Second negative control

We find that PNK often generates sharp bands that do not respond to different RNase conditions. We tried to isolate the bound nucleic acid (either RNA or DNA) from such bands, but were rarely able to detect it. The only time we did, presence of microRNA was found. In this case, the protein–microRNA complex does not respond to RNase concentration, and migrates as a sharp band on the gel. However, this does not apply to all the radioactive bands that do not respond to the varying RNase conditions. To determine the type of protein–RNA complex that was isolated, it is therefore important to compare conditions with high- and low-RNase. In the high-RNase condition, the protein–RNA complex should migrate as a distinct radioactive band at ±5 kDa above the MW of the protein. In the low-RNase condition, the protein–RNA complex should migrate as a more diffuse radioactivity at ±10–30 kDa above the MW of the protein if the protein binds mRNA or any other species of longer RNAs.

4.3. Positive control

As a positive control, a known RNA-binding protein with a good antibody for immunoprecipitation should be used. If working with mammalian tissues, we recommend the hnRNP C1/C2 (H-105, sc-15386, Santa Cruz Biotech) antibody. The hnRNP C1/C2 protein is abundant and expressed in all tissues we tested so far. It contains RRM domains, which cross-links most efficiently to RNA according to our experience. The RNA/hnRNP C complex migrates on the gel at 40–45 kDa at high-RNase, and 45–65 kDa at low-RNase condition, therefore its response to RNase is easy to visualise.

5. Conclusions

The CLIP method has been successfully used to study *in vivo* protein–RNA interactions in several studies [3,6,8,9], and the recent advance in high-throughput sequencing technology allows for a much more comprehensive and global analysis of protein–RNA interactions. Compared with the previous CLIP methods [11], we have changed the protocol such that the 5′ adapter ligation is performed on beads as the first ligation step. This allows more efficient ligation, and reduces the chances of self-ligation. The main challenge of CLIP was that the low amount of starting RNA could be lost during purification or outcompeted by non-specific RNA, adapter concatamers, or contaminating DNA. The current protocol addresses these problems, thus reducing the probability of RNA loss or contamination. It also benefits from being quicker than previous protocols. With it, one can also prepare samples for high-throughput sequencing, and the barcoding system allows for multiplexing of different samples from different experiments, as well as monitoring potential carryover contaminations between experiments. The simultaneous sequencing of several experiments not only allows fast generation of vast amounts of data, but also makes the sequencing cheaper compared with the traditional cloning and sequencing method.

Acknowledgments

We thank Dalia Daujotyte, Martina Halleger, Lira Mamanova, Julian König and Kathi Zarnack for their advice, and Melis Kayikci, Rekins Janky, Madan Babu, Tomaž Curk and Blaž Zupan for help in data analysis. This work was supported by Medical Research Council, UK and European Research Council.

References

[1] A. Bertolotti, D. Ron, J. Cell Sci. 114 (2001) 3207–3212.
[2] S.M. Elbashir, W. Lendeckel, T. Tuschl, Genes Dev. 15 (2001) 188–200.
[3] J. Ule, K.B. Jensen, M. Ruggiu, A. Mele, A. Ule, R.B. Darnell, Science 302 (2003) 1212–1215.
[4] K.B. Jensen, R.B. Darnell, Methods Mol. Biol. 488 (2008) 85–98.
[5] J. Ule, G. Stefani, A. Mele, M. Ruggiu, X. Wang, B. Taneri, T. Gaasterland, B.J. Blencowe, R.B. Darnell, Nature 444 (2006) 580–586.
[6] D.D. Licatalosi, A. Mele, J.J. Fak, J. Ule, M. Kayikci, S.W. Chi, T.A. Clark, A.C. Schweitzer, J.E. Blume, X. Wang, J.C. Darnell, R.B. Darnell, Nature 456 (2008) 464–469.
[7] S. Guil, J.F. Caceres, Nat. Struct. Mol. Biol. 14 (2007) 591–596.
[8] J.R. Sanford, P. Coutinho, J.A. Hackett, X. Wang, W. Ranahan, J.F. Caceres, PLoS One 3 (2008) e3369.
[9] J.R. Sanford, X. Wang, M. Mort, N. Vanduyn, D.N. Cooper, S.D. Mooney, H.J. Edenberg, Y. Liu, Genome Res. 19 (2009) 381–394.
[10] G.W. Yeo, N.G. Coufal, T.Y. Liang, G.E. Peng, X.D. Fu, F.H. Gage, Nat. Struct. Mol. Biol. 16 (2009) 130–137.
[11] J. Ule, K. Jensen, A. Mele, R.B. Darnell, Methods 37 (2005) 376–386.
[12] T.E. England, O.C. Uhlenbeck, Biochemistry 17 (1978) 2069–2076.
[13] L.K. Wang, S. Shuman, Nucleic Acids Res. 30 (2002) 1073–1080.
[14] G.W. Yeo, N.G. Coufal, T.Y. Liang, G.E. Peng, X.D. Fu, F.H. Gage, Nat. Struct. Mol. Biol. 16 (2009) 130–137.