RNA::Protein Crosslinking and Immunoprecipitation *in vivo* (CLIP)

Branislav Kusenda and Andrea Barta

Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohrgasse 9/3, A-1030 Wien, Austria

Address correspondence to:
Andrea Barta or Branislav Kusenda,
Max F. Perutz Laboratories, Medical University of Vienna,
Dr. Bohrgasse 9/3, A-1030 Wien, Austria ;
email: andrea.barta@meduniwien.ac.at; branislav.kusenda@meduniwien.ac.at

1. Abstract

The functions of RNA binding proteins can only be elucidate if their RNA targets are known. “Crosslinking and Immunoprecipitation" (CLIP) of proteins and nucleic acids allows the identification of interacting RNAs in the context of an intact cell which is important for genome-wide reconstructions of genetic circuits. Here we present a modified experimental procedure which includes *in vivo* crosslinking with formaldehyde leading to the formation of a reversible bond between protein and interacting RNAs. This is followed by partial digestion of RNAs, stringent immunoprecipitation of the complex and ligation of RNA adapters. The formaldehyde crosslinks are then reversed by heat treatment and the released RNAs reverse transcribed, amplified and cloned into a vector for sequencing. Alternatively the amplified cDNAs are directly sequenced using any next generation sequencing technology. Using an Rct1 knock out strain of *S. pombe* which is complemented by an HA-tagged Rct1 gene, we have optimized the CLIP analysis for *S.pombe*. If successfully established, the time requirement is about two weeks, followed by sequencing.

Keywords: CLIP, RNA-Protein interaction, Crosslinking, massive parallel sequencing,

2. Introduction

Many methods to study protein-RNA interaction *in vitro* are currently available [1]. If the RNA binding preference of a protein is known, *in vitro* methods like gel shift assay or nuclease protection mapping are relatively quick, easy and produce robust results. However, the environment in a test tube might not at all times reflect the situation in a living cell so false positive or negative results might be produced. If the RNA binding sequence of your favorite protein is not known, an *in vitro* selection method, like SELEX (systematic evolution of ligands by exponential enrichment) could be employed [2]. These experiments use randomized artificial RNA-aptamer libraries to select strong binders for a given target. This allows the isolation of high affinity binding RNAs (RNA aptamer) to a given protein which could potentially be used in a clinical setting to inhibit the function of the protein. However, it has to be kept in mind that high affinity binding motifs might not occur *in vivo.* Many RNA binding proteins, in particular the ones regulating alternative splicing, are dynamic molecules who are binding to their RNA targets in the course of splicing but often have to dissociate again to allow splicing to proceed. Consequently, their binding affinities and hence sequence motifs might be different in the living cell. This RNA binding proteins might have diverse RNA targets which could vary depending on cell type and developmental stage. It was therefore essential to develop an *in vivo* method performed in the context of intact cells or tissue to find RNA targets for a particular protein [3,4]. Crosslinking and immunoprecipitation (CLIP) is a technically complicated method and produces results with high complexity, hence careful data analysis and interpretation is necessary [4]. Whatever *in vivo* or *in vitro* method is used in the first place, it is essential to verify outcomes with a second method to produce conclusive results.

Crosslinking is a powerful approach for RNA-protein interaction studies, because it prevents artificial associations of non target RNAs and RNA binding proteins in cell extracts. In addition, it also allows more stringent washing steps after immunoprecipitation. UV crosslinking was successfully used in mammalian cell culture and provided valuable information about binding sites [5]; however the efficiency of UV crosslinking is not very high. Furthermore UV may damage RNA and not all RNA-binding proteins might be cross-linked to their targets by UV. In contrast, formaldehyde crosslinking can be very efficient and forms reversible crosslinks between proteins and between protein and target RNA [6]. To avoid extensive non-specific crosslinking, formaldehyde treatment should be short and with low concentration. It is necessary to optimize the concentration and time of treatment with formaldehyde for your RNA-binding protein of interest. We used formaldehyde crosslinking for S. pombe cells, as UV crosslinking did not produce satisfying results in our hands.

The protein we used for our investigation was spRct1 which was described as "RNA Recognition Motive Containing Cyclophilin Regulating Transcription" as it is suggested to possess possible functions in activities connecting transcription and pre-mRNA processing [7]. spRct1 is a highly conserved multidomain protein consisting of a peptidyl-prolyl *cis–trans* isomerase (PPIase) domain, followed by an RNA recognition motif (RRM), and a C-terminal domain enriched in charged amino acids. Cyclophilins possess peptidyl–prolyl cis–trans isomerase (PPIase) activity which catalyzes cis–trans isomerization of peptide bonds preceding prolines. spRct1 has the dual ability to bind to specific RNA sequences and to interact with proteins. It was previously shown that spRct1 interacts with the C-terminal domain (CTD) of RNA polymerase II and regulates its phosphorylation, thereby influencing transcription as well as RNA processing [7]. This multidomain cyclophilin is conserved from yeast to humans; however, its RRM shows evolutionarily the highest conservation of the protein domains [8]. As this sequence conservation suggests an important function for the RRM we are aiming at defining RNA sequences which bind to this protein. As it might bind to many different targets we developed a CLIP method to be used in *S. pombe*. Time requirement is about two weeks followed by a sequencing method, whereby any high-throughput sequencing technology is recommended.

3. Protocol

3.1 Materials

RQ Dnase (Promega: 9PIM610), **T1 RNAse** (Fermentas: EN0541), **T4 Polynucleotide Kinase** (T4PNK) (Fermentas: EK0031), **Calf Intestinal Alkaline Phosphatase** (NEB: M0290S), **T4 RNA ligase** (EL0021), **SuperScriptIII** (Invitrogen:18080-044), **Accuprime Pfx** (Invitrogen: 12344-024), **Taq Polymerase** (Finnzymes:F-501S), **RNasin** (Promega: 9PIN261), **Protease Inhibitor Cocktail Tablets** (Roche: 11 873 580 001), **Formaldehyde** (Merck: 1.01003.1000), **Glycine** (Sigma G6279), **Anti-HA agarose beads** (Sigma: E6779), **HA-peptide** (Sigma:I2149), **QIAEX II** (Qiagen: 20021), **TOPO TA cloning Kit** (with pCR2.1-TOPO) (Invitrogen: K4500-01), **Polypropylene Columns** (Qiagen: 34964).

PBS buffer:

137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH of 7.4.

RIPA buffer:

50 mM Tris–Cl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl, Protease Inhibitor Cocktail Tablets: one tablet for 10 ml buffer

Buffer A:

1xPBS (tissue culture grade; no Mg^{2+}, no Ca^{2+}), 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40

Buffer B:

5x PBS (tissue culture grade; no Mg^{2+}, no Ca^{2+}), 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40

Buffer C:

50 mM Tris–Cl, pH 7.4, 10 mM $MgCl_2$, 0.5% NP-40

Buffer D:

50 mM Tris–HCl, pH 7.4, 20 mM EGTA, 0.5% NP-40

DNase Reaction Buffer:

40 mM Tris-HCl (pH 8.0), 10 mM $MgSO_4$ and 1mM $CaCl_2$.

T1 RNase Reaction Buffer:

50 mM Tris-HCl (pH 7.5), 1 mM EDTA

Phophatase Reaction Buffer:

50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM Dithiothreitol, pH 7.9 at 25°C

T4 RNA ligase Reaction Buffer:

50 mM HEPES-NaOH (pH 8.0 at 25°C), 10 mM $MgCl_2$, 10 mM DTT.

T4 Polynucleotide Kinase Reaction Buffer:

50 mM Tris-HCl (pH 7.6 at 25°C), 10 mM $MgCl_2$, 5 mM DTT, 0,1 mM spermidine, 0,1 mM EDTA.

RNA linker sequences [3]:

L5: 5´-OH AGGGAGGACGAUGCGG 3´ -OH

L3: 5´P GUGUCAGUCACUUCCAGCGG 3´-puromycin

DNA primers [3]:

P5: 5´-AGGGAGGACGATGCGG-3´

P3: 5´-CCGCTGGAAGTGACTGACAC-3´

3.2 Growth conditions and cell harvest

1litre of *S.pombe* culture strain with the protein of interest (if no antibody for immunoprecipitation is available use e.g. HA-tagged proteins. (HA – hemagglutinin YPYDVPDYA).

Harvest cells before the end of log phase by centrifugation for 10 min at 4°C, 3,000g.

Wash pellet with ice cold PBS.

Add up to 10 ml with PBS at room temperature, resuspend pellet, transfer to a 50 ml Falcon (polypropylene) tube.

3.3 Crosslinking

Add 0.3 ml 37% formaldehyde.

Incubate for 10 min with shaking at room temperature.

Add 2.75 ml 1 M glycine.
Incubate for 5 min with shaking at room temperature, spin down at 4°C, 3,000g
Wash with ice cold PBS.
Freeze pellets in liquid nitrogen for 10 min defrost on ice (or store at -80°C if you want to continue later).

3.4 Sonication of the cells

Resuspend pellet in 10 ml ice cold RIPA,
Sonication 6x10 sec (settings 50/50 on Microson XL2007 ultrasonic homogenizer).
Between each cycles, the samples are kept in ice for at least 2 min.
Transfer to 2 ml Eppendorf tubes and remove insoluble material by centrifugation for 10 min, at 4°C, 14,000*g.* Combine supernatants into a 15 ml Falcon tube.

3.5 Immunoprecipitation

Add 50 -150 µl Anti-HA agarose beads equilibrated with RIPA to the supernatant.
Incubate for 120 min at 4°C on a turning wheel.
Transfer beads on a column or into a fresh tube.
Wash 2 times with each: Buffer A, Buffer B, Buffer C, 1x DNase Reaction Buffer.

3.6 DNase treatment

Add 500 µl of 1xDNase Reaction Buffer
Add 10 µl RQ DNase (1u/µl).
Incubate for 10 min at 37°C with constant mixing at 1000 rpm.
Wash 2 times with each: Buffer A, Buffer B, Buffer C, T1 RNase Reaction Buffer.

3.7 RNAse treatment:

Add 500 µl of 1xT1 RNase Reaction Buffer
Add 10 µl T1 RNase dilution (optimized for your conditions: high (1:100) or/and low (1:5000), concentration of stock 1000 u/µl)
Incubate for 10 min at 37°C with constant mixing at 1000 rpm.
Wash 2 times with each: Buffer A, Buffer B, Buffer C, Phophatase Reaction Buffer.

3.8 Dephosphorylation

Add 80µl dephosphorylation reaction:

69 µl water

8 µl of 10×dephosphorylation buffer

3 µl Calf Intestinal Alkaline Phosphatase (1 u/µl)

Incubate for 10 min at 37°C with mixing at 1000 rpm.

Wash 2 times with each: Buffer D, Buffer C, T4 RNA ligase Reaction Buffer.

3.9 L3 linker ligation

Add 80 – 300 µl L3 linker ligation reaction:

40 µl of L3 RNA linker (4 pmol/µl)

14 µl water

8 µl 10×T4 RNA ligase buffer

8 µl BSA (0.2 µg/µl)

8 µl ATP (10 mM)

2 µl T4 RNA ligase (10 u/µl)

Incubate at 16°C overnight mixing each 5 min for 10 sec at 1000 rpm (program thermomixer).

Wash 2 times with: Buffer C, T4 PNK Reaction Buffer.

3.10 Phosphorylation of the 5' end of RNA

Add 80 µl PNK reaction:

40 µl water

8 µl 10xPNK buffer

2 µl ATP (10mM)

4 µl T4 PNK enzyme (10 u/µl)

Incubate at 37°C for 15 min mixing each 3 min for 10 sec at 1000 rpm.

Wash 2 times with each: Buffer C, T4 RNA Ligation Reaction Buffer.

3.11 L5 linker ligation

Add 80 - 300µl L5 linker ligation reaction:

40 µl of L5 RNA linker (4 pmol/µl)

14 µl water

8 µl 10×T4 RNA ligase buffer

8 µl BSA (0.2 µg/µl)

8 µl ATP (10 mM)

2 µl T4 RNA ligase (10 u/µl)

Incubate at 16°C overnight, mixing each 5 min for 10 sec at 1000 rpm.

Wash 2 times with: Buffer C.

3.12 Elution of the Protein::RNA complex

Elute protein::RNA complex from beads by adding 100 – 500 µl of HA-peptide diluted in PBS (100 mg/ml).

Spin and transfer supernatant into a fresh tube.

Reverse crosslinks by heating the sample for 45 min at 70°C.

3.13 cDNA synthesis

8 µl sample after the reversion of crosslinking

2 µl of P3 DNA primer (5 pmol/µl)

heat at 65°C for 5 min, than chill to 4°C.

Add:

3 µl of 3 mM dNTPs

1 µl of 0.1 M DTT

4 µl of 5×SuperScript RT buffer

1 µl RNasin

1 µl SuperScript III

Incubate at 50°C for 30 min and 90°C for 3 min.

3.14 PCR amplification:

2 µl of the first strand cDNA reaction

27 µl Accuprime Pfx Supermix

0.5 µl P5 DNA primer (30 pmol/µl)

0.5 µl P3 DNA primer (30 pmol/µl)

5 min 95°C; 30 cycles of (20 s for 95°C, 30 s for 61°C, and 20 s for 68°C); 5 min 68°C.

3.15 Size selection of DNA fragments

Load the PCR sample on a standard 1.5 - 2% agarose gel, visualize after electrophoresis, cut out the DNA of 40 – 200 bp.

Extract DNA with the QIAEX II kit [9]:

Add 6 volumes of Buffer QX1 to 1 volume of gel fragments.
Add 15 µl QIAEX II to the sample (resuspend QIAEX II by vortexing for 30 s).
Incubate at 50°C for 10 min, mixing by vortexing every 2 min.
Centrifuge the sample for 30 s and remove supernatant.
Wash the pellet with 500 µl of Buffer QX1
Wash the pellet twice with 500 µl of Buffer PE
Air-dry the pellet for 10 – 15 min.
Add 20 µl of 10 mM Tris-HCl, pH 8.5 and resuspend the pellet by vortexing.
Incubate at room temperature for 5 min.
Centrifuge for 30 s and transfer the supernatant into a clean tube.

Now this sample is ready for massive parallel sequencing, e.g. 454 or solexa, however, it is advisable to check out the quality of your sample by cloning and sequencing of approximately 100 independent clones.

3.16 Cloning Step: Generation of overhangs

7 µl QIAEXII extracted DNA
1 µl 10xTaq Buffer
1 µl 10mM dATP
1 µl Taq Polymerase
Incubate at 72°C for 20 min.

3.17 TOPO TA CLONING REACTION:

4.5 µl DNA with A overhangs
1 µl salt solution (1.2 M NaCl, 0.06 M $MgCl_2$)
0.5 µl pCR2.1-TOPO vector
Incubate for 10 min at room temperature.

3.18 Transformation of E. coli

Add 3 µl of TOPO reaction into one vial (50 µl) of competent cells
10 min on ice
30 -4 5 sec 42°C
2 min on ice
Add 250 µl of SOC or LB medium preheated at 37°C.

60 min shaking at 37°C

Spread 50 - 200 µl per plate (50 µg/ml Amp, 20-40 µl of Xgal (40 mg/ml stock)

Blue/white colony selection (Pick white colonies).

Isolate DNA by standard Mini-prep procedures and sequence the insert by using M13 primer.

4. Example of an experiment

In this experiment we used an Rct1 knock out strain of S. pombe which is complemented by an HA-tagged Rct1 gene [7] and performed a CLIP experiment as described above. Figure 1 presents a Western blot to demonstrate crosslinking and Figure 2 shows an agarose gel after electrophoresis of the CLIP samples processed with different concentrations of RNase T1.

5. Troubleshooting

Problem	**Solution**
No cross-linking detected	Use higher concentrations of formaldehyde Try prolonged times of cross-linking treatment.
No PCR product detected	Test ligase and PNK for activity. Use lower concentration of RNase. Use SYBR Gold for gel staining. Add more PCR cycles or use PCR product as template for second PCR reaction.
No size difference between samples treated with high and low concentration of RNase	Use lower and higher concentration of RNase to find optimal concentrations which will produce different sizes of. Use different RNase e.g. RNase A.
Primer dimmers cloned	Use lower concentration of DNA primers. Use higher annealing temperature. Perform size selection of PCR product, cut out the DNA of 50 – 200 bp.

6. REFERENCES

[1] Lin, R.J.(2008). RNA- Protein Interaction Protocols(Duarte: Humana Press).
[2] Djordjevic, M. (2007). SELEX experiments: new prospects, application and data analysis in inferring regulatory pathways. Biomol. Engineering 24, 179-189
Buch RNA-Protein interactions
[3]Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science. 14;302(5648):1212-5.
[4] Wang Z, Tollervey J, Briese M, Turner D, Ule J.(2009)CLIP: construction of cDNA libraries for high-throughput sequencing from RNAs cross-linked to proteins in vivo. Methods. 48(3):287-93.
[5] Ule J, Jensen K, Mele A, Darnell RB.(2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods. 37(4):376-86.
[6] Niranjanakumari S, Lasda E, Brazas R, Garcia-Blanco MA. (2002). Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. Methods 26(2):182-90.
[7] Gullerova M, Barta A, Lorkovic JZ. (2007). Rct1, a Nuclear RNA Recognition Motif-Containing Cyclophilin, Regulates Phosphorylation of the RNA Polymerase II C-Terminal Domain Mol Cell Biol. 27(10):3601-11. Epub 2007 Mar 5.
[8] Lorković ZJ, Barta A.(2002) Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana.Nucleic Acids Res. 1;30(3):623-35.
[9] QIAEX II® Handbook avalibe online: *www1.qiagen.com/HB*

Acknowledgements:

The authors are grateful to Z. Lorkovic for fruitful discussions. This work was funded by the EU FP6 Programme Network of Excellence on Alternative Splicing (EURASNET) [LSHG-CT-2005-518238]; the Austrian Science Foundation (FWF: SFB-F017/10/11) and the Austrian GEN-AU program (ncRNAs).

Figure legends

Figure 1. Western blot of crosslinked extracts probed with Anti-Rct1 antibody. Lane 1: control sample (before crosslinking); lane 2: sample crosslinked with formaldehyde.

Figure 2. Electrophoresis of PCR reactions from CLIP samples (3.14) processed with different concentrations of RNase T1. Lane 1: marker GeneRuler (Fermentas), lane 2: untreated by RNAse, lane 3: positive control for PCR reaction, lane 4: treated with low concentration (dilution 1:5000), lane 5: treated with high concentrations (dilution 1:100) of RNase T1.

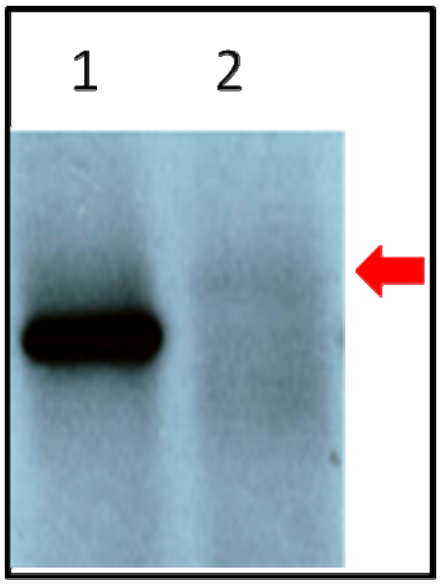

Figure 1:

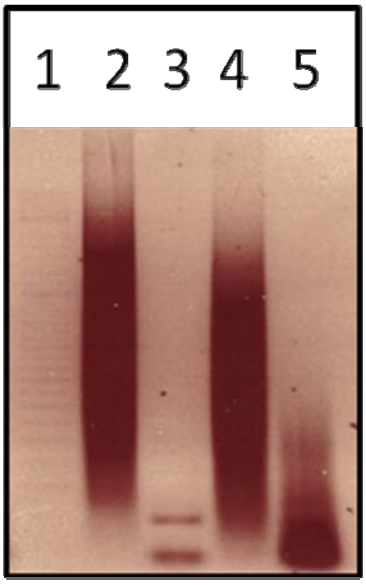

Figure 2:

Overview:
RNA::Protein Crosslinking and Immunoprecipitation *in vivo* (CLIP)