Chapter 8

Comprehensive Identification of RNA-Binding Proteins by RNA Interactome Capture

Alfredo Castello, Rastislav Horos, Claudia Strein, Bernd Fischer, Katrin Eichelbaum, Lars M. Steinmetz, Jeroen Krijgsveld, and Matthias W. Hentze

Abstract

RNA associates with RNA-binding proteins (RBPs) from synthesis to decay, forming dynamic ribonucleoproteins (RNPs). In spite of the preeminent role of RBPs regulating RNA fate, the scope of cellular RBPs has remained largely unknown. We have recently developed a novel and comprehensive method to identify the repertoire of active RBPs of cultured cells, called RNA interactome capture. Using in vivo UV cross-linking on cultured cells, proteins are covalently bound to RNA if the contact between the two is direct ("zero distance"). Protein-RNA complexes are purified by poly(A) tail-dependent oligo(dT) capture and analyzed by quantitative mass spectrometry. Because UV irradiation is applied to living cells and purification is performed using highly stringent washes, RNA interactome capture identifies physiologic and direct protein-RNA interactions. Applied to HeLa cells, this protocol revealed the near-complete repertoire of RBPs, including hundreds of novel RNA binders. Apart from its RBP discovery capacity, quantitative and comparative RNA interactome capture can also be used to study the responses of the RBP repertoire to different physiological cues and processes, including metabolic stress, differentiation, development, or the response to drugs.

Key words RNA, RNA-binding protein, Proteomics, Proteome, RNA interactome capture, Posttranscriptional regulation, Gene expression

1 Introduction

In the last decade, in vitro and in silico approaches have been developed in order to determine the complete repertoire of RNA-binding proteins (RBPs), referred to here as the RNA interactome. While in vitro approaches served to identify dozens of novel RBPs [1, 2], the abundant negative charges of the sugar-phosphate backbone of the RNA have the capacity to mediate unspecific binding of basic proteins in vitro and, therefore, extensive validation is required. Biocomputational methods can recognize proteins bearing classical RNA-binding domains (RBDs) such as the RNA

Erik Dassi (ed.), *Post-Transcriptional Gene Regulation*, Methods in Molecular Biology, vol. 1358,
DOI 10.1007/978-1-4939-3067-8_8,

recognition motif (RRM) and K-homology domain (KH) [3]. However, their capacity to identify unorthodox RBDs is very limited. To address these limitations we and others developed RNA interactome capture, which combines UV cross-linking and oligo(dT) capture to pull down proteins bound to polyadenylated RNA in living cells. Applied to HeLa [4] and HEK293 cells [5], RNA interactome capture determined the first near-complete RNA interactomes of a human cell line. More recently, this approach has also been applied successfully to mouse embryonic stem cells [6] and *Saccharomyces cerevisiae* [7].

In a first step, protein-RNA interactions are "fixed" applying two different approaches: (a) Irradiation with ultraviolet (UV) light at 254 nm of cell monolayers induces short-lived radicals at the nucleotide base that can attack amino acids in close proximity forming covalent bonds. (b) The photoactivatable-ribonucleoside-enhanced cross-linking (PAR-CL) protocol employs 4-thiouridine (4SU) [8], which is taken up by cells and incorporated into nascent RNAs. Protein-RNA cross-linking is achieved by irradiation at 365 nm. Following UV cross-linking by either approach, lysis under denaturing conditions, and homogenization, polyadenylated RNAs and their covalently bound proteins are isolated with oligo(dT) magnetic beads through highly stringent washes. RBPs bound to polyadenylated RNA are then released by RNase treatment and identified by proteomics.

RNA interactome capture has notable advantages over previous RBP identification methods: (1) As UV irradiation is applied to cell monolayers, protein-RNA interactions discovered have occurred within the native context and without overexpression. (2) Because the free radicals are induced at the nucleotide base, UV irradiation promotes exclusively protein-RNA and not protein-protein cross-links [9, 10]. (3) The short-lived nature of the free radicals (nanosecond range) limits covalent bond formation to amino acids at "zero distance" (~2 Å) [9]. (4) Due to the stability of nucleic acid hybrids (i.e., poly(A) tails binding to oligo(dT) magnetic beads) in the presence of high-salt and chaotropic detergents, very stringent washing steps can be applied to remove all non-cross-linked polypeptides. On the other hand, RNA interactome capture will fail to detect RBPs when they are (1) not bound to polyadenylated RNAs, (2) not expressed in the cell type under study, (3) not active in RNA binding under the experimental conditions, or (4) not cross-linked efficiently by UV irradiation.

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents. Store buffers at 4 °C and samples at −70–80 °C (unless indicated otherwise). All buffers should be filtered and autoclaved

before usage. For buffers containing LiDS or DTT, autoclaving should be performed prior to the addition of these heat-sensitive components.

2.1 Buffers

1 Lysis buffer: 20 mM Tris–HCl pH 7.5, 500 mM LiCl, 0.5 % LiDS (wt/v, stock 10 %), 1 mM EDTA, 5 mM DTT.

2 Buffer 1: 20 mM pH 7.5 Tris–HCl, 500 mM LiCl, 0.1 % LiDS (wt/v), 1 mM EDTA, 5 mM DTT.

3 Buffer 2: 20 mM pH 7.5 Tris–HCl, 500 mM LiCl, 1 mM EDTA, 5 mM DTT.

4 Buffer 3: 20 mM pH 7.5 Tris–HCl, 200 mM LiCl, 1 mM EDTA, 5 mM DTT.

5 Elution buffer: 20 mM pH 7.5 Tris–HCl, 1 mM EDTA.

6 Buffer 4: 50 mM NaCl.

7 10× RNase buffer: 50 mM pH 7.5 Tris–HCl, 1.5 M NaCl, 5 mM DTT.

8 5× Proteinase K buffer: 50 mM pH 7.5 Tris–HCl, 750 mM NaCl, 1 % SDS, 50 mM EDTA, 2.5 mM DTT, 25 mM $CaCl_2$.

9 Phosphate-buffered saline (PBS).

2.2 Reagents

1. Oligo (dT_{25}) magnetic beads (New England Biolabs, S1419S).
2. 4-Thiouridine (4SU, Sigma).
3. Lithium dodecyl sulfate (LiDS).
4. Lithium chloride (LiCl).
5. DTT.
6. Amicon Ultra® Centrifugal Filters (50 ml, 10 KDa cutoff, Millipore UFC901024).
7. Ribonuclease T1 from *Aspergillus oryzae* (RNase T1).
8. Ribonuclease A from bovine pancreas (RNase A).
9. RNeasy kit (Qiagen).

2.3 Equipment

1. Humidified 37 °C, 5 % CO_2 incubator.
2. Microbiological cabinet class 2.
3. 15 cm dishes or 500 mm^2 square dishes.
4. Cross-linking devices: cCL = 254 nm bulbs; PAR-CL= 365 nm bulbs. Spectrolinker UV Cross-linkers (Spectroline).
5. Needle (27G, 3/4-inch; no. 20, 0.4 mm × 19 mm).
6. Sterile syringe (5 ml).
7. 50 ml Magnetic separation rack and 12-tube (2 ml) magnetic separation rack.
8. Refrigerated bench-top centrifuge.

3 Methods

Experimental design: Include a non-irradiated (noCL) sample to control the signal obtained for UV-irradiated samples for background noise derived from the purification with oligo(dT) magnetic beads [11]. In the same vain, include 4SU-treated, non-irradiated cells (4SU noCL) as a control for PAR-CL. The overview of the cross-linking and purification protocol is schematized in Fig. 1a–c.

1. For conventional cross-linking (cCL), seed cells in 5×500 cm^2 dishes (1500 cm^2 of total growth area) with normal medium (e.g., DMEM and 5 % fetal calf serum, FCS) to reach 80–90 % confluence after overnight incubation (~1.9×10^7 cells per dish; *see* **Note 1**) (Fig. 1a). For PAR-CL, follow the same seeding protocol, but supplement the medium with 100 μM 4SU (*see* **Note 2**).
2. After overnight incubation, wash cells twice with 30 ml of PBS (room temperature) until the PBS remains colorless.
3. Remove the PBS and place the culture dishes without their lids on ice at ~15 cm from the UV source. Irradiate with 150 mJ/cm^2 at 254 nm UV light for cCL or at 365 nm UV light for PAR-CL (*see* **Note 3**). After irradiation, add 15 ml of ice-cold PBS per dish and keep the already irradiated dishes at 4 °C while processing the rest of the dishes.
4. Scrape the cells into the PBS added in **step 3** with a rubber policeman and centrifuge at $400 \times g$ for 3 min at 4 °C (*see* **Note 4**). Remove and discard the supernatant.
5. Add 10 ml of ice-cold lysis buffer, resuspend the pellet pipetting up and down, and supplement the lysate with 30 additional ml of lysis buffer. Mix the lysate by inverting the 50 ml tube (*see* **Note 5**).
6. Pass the lysate through a 5 ml syringe with a narrow needle (gauge 0.4 mm diameter) to homogenize. Repeat the process two additional times until the viscosity of the lysate is significantly reduced (*see* **Notes 6** and **7**). Keep the sample in ice while processing the other samples and the oligo(dT) beads are equilibrated (see below).
7. Equilibrate 2 ml of oligo(dT)$_{25}$ magnetic beads per tube by washing three times with 5× volumes of lysis buffer. Resuspend the bead pellet in 2 ml of lysis buffer (*see* **Note 5**).
8. Add resulting bead suspension (2 ml) to the sample and incubate for 1 h at 4 °C with gentle rotation (*see* **Note 5**).
9. Place the tubes on a magnet at 4 °C and wait until the beads are collected (this can take up to 30 min). Recover the supernatant and store it in a new tube at 4 °C for the following two cycles of oligo(dT) capture (see below).

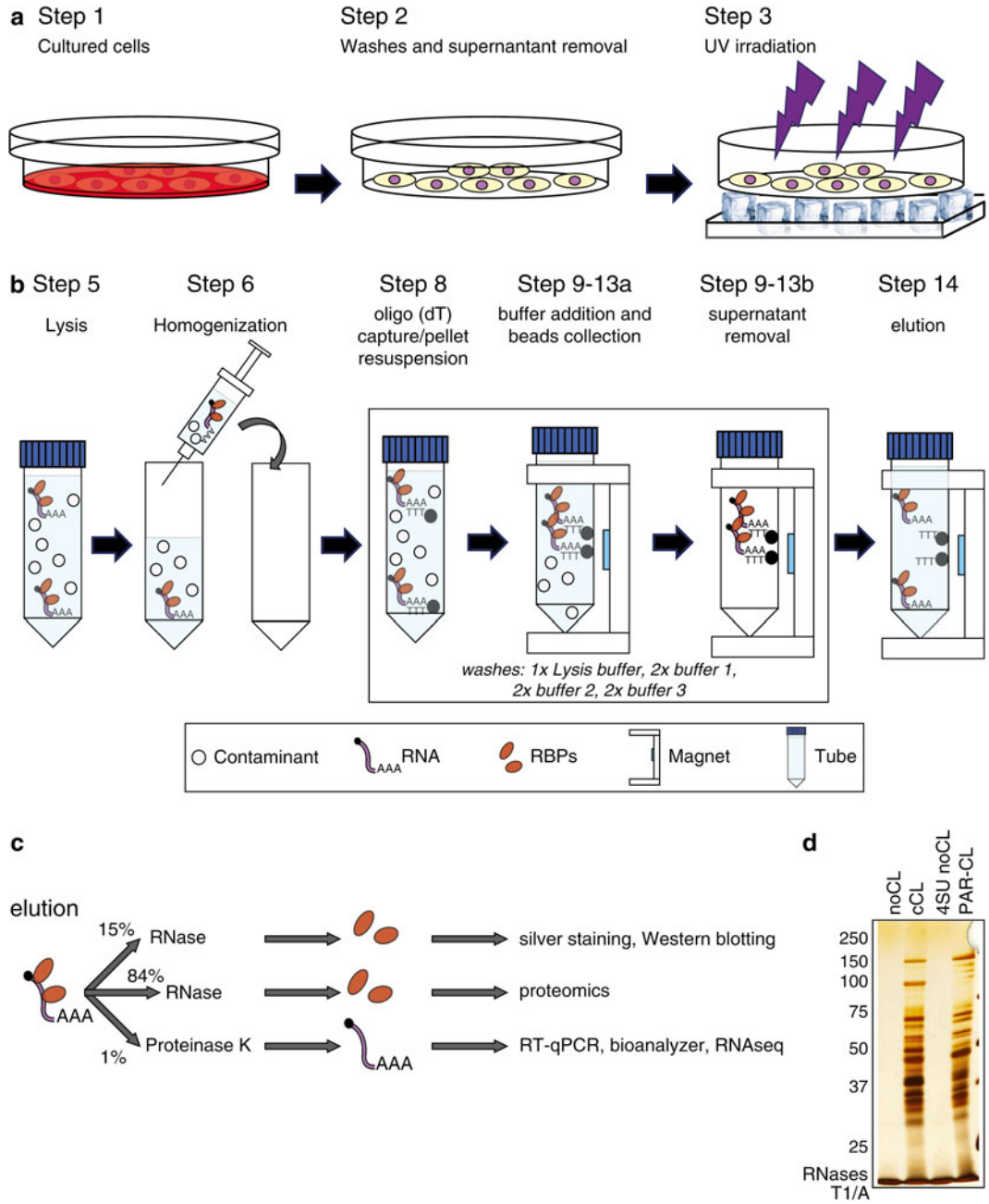

Fig. 1 RNA interactome capture workflow: Schematic representation of in vivo UV cross-linking (**a**), oligo (dT) capture (**b**), and downstream processing of the eluates (**c**). After elution, samples are treated with either proteinase K for RNA quality controls or RNases for protein quality controls and mass spectrometry. (**d**) Representative silver staining of RNase-treated eluates

10. Resuspend the bead pellet in 35 ml ice-cold lysis buffer. Incubate for 5 min at 4 °C with gentle rotation and pellet the beads with the magnet. Discard the supernatant.
11. Add 35 ml of ice-cold buffer 1, resuspend the beads, and incubate for 5 min at 4 °C with gentle rotation. Pellet the beads with the magnet and discard the supernatant (*see* **Notes 8** and **9**). Repeat this step once.
12. Add 35 ml of ice-cold buffer 2 and resuspend the beads. Mix by inverting the sample ten times. Pellet the beads with the magnet and discard the supernatant (*see* **Notes 8** and **9**). Repeat this step once.
13. Add 35 ml of ice-cold buffer 3, and resuspend the beads. Mix by inverting the sample ten times. Pellet the beads with the magnet and discard the supernatant. Repeat this step once.
14. Resuspend the bead pellet in 500 μl of elution buffer. Transfer the sample to a sterile 1.5 ml tube and elute the RNA-protein complexes by incubating at 55 °C for 3 min. Collect the beads in a magnet and transfer the supernatant to a new tube. Pellet any residual beads again in the magnet and collect the supernatant and transfer it to a new sterile tube (*see* **Note 10**). Determine the resulting RNA content using a Nanodrop device.
15. Recycle the beads following the manufacturer's recommendation. Add recycled beads to the lysate stored at 4 °C (*see* **step 9**) and repeat the isolation (from **steps 8** to **14**) twice (three isolation cycles in total).
16. Pool the eluates from the three successive oligo (dT) capture cycles (final volume 1.5 ml).
17. Take 20 μl of the pooled eluate from **step 16**. Add 5 μl of 5× proteinase K buffer and 1 μg of proteinase K, incubate for 30 min at 37 °C and 30 min at 50 °C. Next, isolate the RNA with RNeasy kit (Qiagen) or trizol (Invitrogen). Use purified RNA for RNA quality control analyses (e.g., RT-qPCR using primers against mRNAs and rRNAs, bioanalyzer, or RNAseq) [4, 11].
18. Take the rest of the elution and add 150 μl of 10× RNase buffer, and ~50–100 U of RNase T1 and RNase A. Incubate at 37 °C for 1 h followed by 15 min at 55 °C.
19. Transfer the eluate into an Amicon Ultra 10^{-3} KDa cutoff (*see* **Note 11**). Top up the filter device with buffer 4 and centrifuge at 4000 × *g* for 45 min at 4 °C.
20. Discard the flow through and top up the filter device again with buffer 4. Centrifuge at 4000 × *g* for 45 min at 4 °C.
21. Recover the sample from the filter unit in about 200 μl.
22. Use 30 μl of the sample for protein quality analyses (e.g., silver staining in Fig. 1d) [4, 11].

23. Once quality controls are performed satisfactorily, the rest of the sample can be analyzed by quantitative mass spectrometry. RNA interactome capture is compatible with all the state-of-the-art quantitative proteomic approaches, including label-free quantification [4], SILAC [5], dimethyl labeling [6], and isobaric labeling (tandem mass tags, TMT).

4 Applications

RNA interactome capture has been used to determine the RNA-bound proteome of HeLa [4], HEK293 [5], mouse embryonic stem cells [6], and *S. cerevisiae* [7]; it can readily be applied to other cell lines and primary cells, and likely be adapted to organisms. Finally, RNA interactome capture can be used in a quantitative and comparative way to explore the plasticity of mRNA interactomes in response to different physiological conditions and biological cues. Moreover, RNA interactome capture has been used to study the RNA-binding capacity of a given protein in vivo. In brief, the protein of interest is fused to eGFP and expressed in cultured cells. Upon UV cross-linking and oligo (dT) capture, eGFP signal is measured in a plate reader and used as a proxy for RNA binding. This protocol requires 1/5 of the cells, buffer volume, and beads indicated above [12].

5 Notes

1. The cell number indicated in the protocol refers to HeLa cells and this may vary between cell lines due to differences in cell volume and RNA content per cell. A successful large scale RNA interactome capture experiment will yield ~100–300 μg of RNA upon oligo (dT) pull down. Knowing the amount of RNA isolated from a defined number of cells, it is possible to calculate the quantity of cells required to capture ~100–300 μg of RNA.
2. We observed that 100 μM is the optimal concentration of 4SU for most of the cell lines tested. Nevertheless, the 4SU dose may require optimization for certain cell lines.
3. Irradiation with 150 mJ/cm^2 of 254 nm UV light yields a relatively high UV cross-linking efficiency while keeping the RNA intact. In our hands, this dose is optimal for most adherent cell types. However, we have found few remarkable exceptions, suggesting that UV dosage may require optimization when working with different cell lines in order to maximize protein yield after oligo (dT) capture.
4. Some cells are sensitive to scraping. In these cases, we recommend to perform direct on-plate lysis to avoid loss of material.

Add the lysis buffer directly onto the cell monolayers and scrape with a rubber policeman. Skip the downstream centrifugation step and proceed with the homogenization.

5. If RNA quality controls reveal a poor enrichment of mRNA over rRNA, we recommend (a) increasing the volume of lysis buffer in **step 5**, (b) reducing the amount of beads, or (c) performing the hybridization at room temperature. These alterations of the original protocol may help to reduce nonspecific adherence to the oligo (dT) beads.
6. During homogenization maintain a good flow rate of the lysate through the narrow needle applying constant pressure on the syringe plunger. If the sample is still very viscous in spite of the three rounds of homogenization, increase either the number of homogenization cycles or the lysis volume.
7. At this step, it is possible to freeze the sample at –70–80 °C, although we recommend to avoid unnecessary freezing and, if possible, to proceed with the oligo (dT) capture after homogenization.
8. If RNA interactome capture is performed successfully, a halo will be visible around the bead pellet while washing with buffers 1, 2, and 3 in a UV cross-linking-dependent manner. Early appearance of the halo correlates with high protein content in the oligo (dT) pull down.
9. If the purification leads to a significant loss of magnetic beads, we recommend to add 0.025 % NP-40 (Igepal) to the buffers 1 and 2. Note that addition of detergent in these buffers will prevent the generation of halo (*see* **Note 8**). Avoid the use of NP-40 in buffer 3 since this detergent will impair downstream mass spectrometric analyses.
10. Removal of residual beads is key to prevent the blockage of the filter unit in downstream steps.
11. Protein can be concentrated by alternative methods such as ethanol or TCA precipitation. Nevertheless, we recommend to test these protocols for potential effects on downstream mass spectrometric analyses.

Acknowledgements

A.C. acknowledges support by the MRC Career Development Award Ref: MR/L019434/1. M.W.H. acknowledges support by the ERC Advanced Grant ERC-2011-ADG_20110310 and by a grant from the Virtual Liver Network of the German Ministry for Science and Education. A.C. is funded by the MRC Career Development Award number MR/L019434/1

References

1. Scherrer T, Mittal N, Janga SC, Gerber AP (2010) A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. PLoS One 5, e15499
2. Tsvetanova NG, Klass DM, Salzman J, Brown PO (2010) Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae. PLoS One 5:e12671, 12610.11371/journal.pone.0012671
3. Anantharaman V, Koonin EV, Aravind L (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res 30:1427–1464
4. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149:1393–1406
5. Baltz AG, Munschauer M, Schwanhausser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell 46:674–690
6. Kwon SC, Yi H, Eichelbaum K, Fohr S, Fischer B, You KT, Castello A, Krijgsveld J, Hentze MW, Kim VN (2013) The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol 20:1122–1130
7. Mitchell SF, Jain S, She M, Parker R (2013) Global analysis of yeast mRNPs. Nat Struct Mol Biol 20:127–133
8. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141:129–141
9. Pashev IG, Dimitrov SI, Angelov D (1991) Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. Trends Biochem Sci 16:323–326
10. Suchanek M, Radzikowska A, Thiele C (2005) Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. Nat Methods 2:261–267
11. Castello A, Horos R, Strein C, Fischer B, Eichelbaum K, Steinmetz LM, Krijgsveld J, Hentze MW (2013) System-wide identification of RNA-binding proteins by interactome capture. Nat Protoc 8:491–500
12. Strein C, Alleaume AM, Rothbauer U, Hentze MW, Castello A (2014) A versatile assay for RNA-binding proteins in living cells. RNA 20:721–731