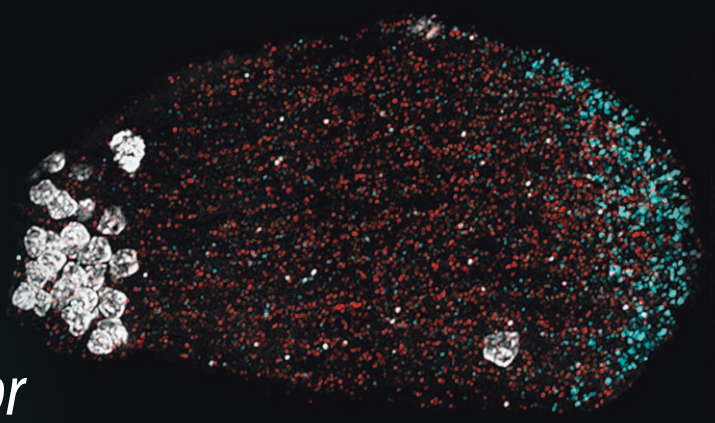
Methods in Molecular Biology 2666

Springer Protocols



Ren-Jang Lin *Editor*

RNA-Protein Complexes and Interactions

Methods and Protocols

Second Edition

Humana Press

Chapter 16

Streamlined Purification of RNA–Protein Complexes Using UV Cross-Linking and RNA Antisense Purification

Nhu Trang, Tong Su, Simone Hall, Nada Boutros, Bobby Kong, Calvin Huang, and Colleen A. McHugh

Abstract

RNA–protein interactions are important in development and disease, but identification of novel RNA–protein interactions remains challenging. Here, we describe an updated capture method to identify direct and specific RNA–protein interactions. First, RNA and protein are covalently cross-linked in living cells by treatment with UV light at 254 nanometers wavelength. The antisense purification approach is dependent upon nucleic acid hybridization between biotinylated DNA probes and a target RNA. Target protein:RNA:DNA complexes are enriched by capture on streptavidin magnetic beads and purified through several denaturing washes that remove nonspecific protein and nucleic acid interactors. Mass spectrometry is used to identify proteins that are specifically enriched in the target RNA capture. This method has been applied to discover the protein interactions of noncoding RNAs but can be used to capture any RNA where the target sequence is known.

Key words RNA–protein interactions, UV cross-linking, Endogenous complexes, Macromolecular complex purification, Nucleic acid biology, Mass spectrometry

1 Introduction

Noncoding RNAs are important molecules controlling cell growth and organismal development. Many human cancers display dysregulated expression of noncoding RNAs contributing to disease phenotypes, and the interactions between RNA and proteins are of great interest as novel targets for molecular medicine. A greater understanding of the RNA–protein interactome is needed to enable the development of specific treatments for RNA-mediated diseases. However, it has been challenging to confidently identify direct interacting factors of noncoding RNAs because it is difficult to purify only the interacting partners that are specific to a given cellular RNA [1]. One challenge is limited abundance of the target RNA. We describe a method for non-adherent culture of cancer

Ren-Jang Lin (ed.), *RNA-Protein Complexes and Interactions: Methods and Protocols*,
Methods in Molecular Biology, vol. 2666, https://doi.org/10.1007/978-1-0716-3191-1_16,

cells grown in large volumes of medium for easy scale-up of the experiment. Stirred cultures of cells grown in suspension allow production of up to 1×10^6 cells per milliliter in a few days with minimal medium cost and handling time. A second challenge is difficulty in separating direct interaction partners from indirect or nonspecific interactors. Advancements in high-sensitivity and high-resolution mass spectrometry have made it possible to identify very small quantities of proteins in a sample. A drawback of this high sensitivity is that identification of a short list of direct and specific interactions for a target RNA molecule is made more challenging when thousands of proteins are detected in a single sample [2]. Nonspecific in vitro interactions of highly expressed RNA-binding proteins occurring in solution can also contribute to false-positive protein interactor identifications [3]. To address these issues, we adapted a streptavidin and biotin-based oligonucleotide hybridization strategy to purify RNA–protein complexes.

We describe here an updated RNA–protein complex capture method based on the RNA antisense purification protocol coupled with mass spectrometry [4]. To increase abundance of the target RNA, input cell numbers can be scaled up to 200–500 M cells per sample or the target RNA can be overexpressed. First, live cells are treated with UV_{254nm} light to covalently cross-link specific RNA–protein interactions in their relevant complexes in each cell [5]. Next, cells are lysed and long biotinylated single DNA strands are added, consisting of nucleotide sequences complementary to the target RNA (Fig. 1a). Hybrids of the biotinylated DNA and target RNA, along with any cross-linked proteins, are enriched by separation from lysate using streptavidin-coated magnetic beads (Fig. 1b). Any non-cross-linked nonspecific and in vitro interactors are removed from the target RNA by washing with a strongly denaturing buffer containing 4 M urea and several different detergents. Finally, enriched proteins and RNA are eluted from the magnetic bead supports and evaluated by mass spectrometry or Western blot, and quantitative real-time PCR (qPCR), respectively. This streamlined method has been used successfully in mouse and human cells to identify noncoding RNA–protein complexes. This updated protocol includes additional important information, such as instructions on how to harvest and cross-link large quantities of cancer cells, details on RNA sample preparation, and guidelines for mass spectrometry data analysis. This simplified and streamlined protocol will make RNA capture experiments accessible to researchers who wish to identify direct RNA–protein interactions with lower cost and effort than previous versions of this technique. Although our lab has applied this method to noncoding RNAs, it can be used for any RNA target where the sequence is known.

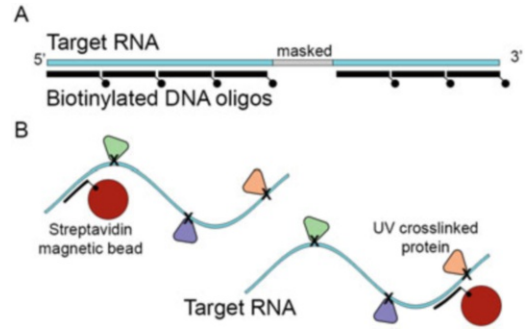


Fig. 1 Overview of probe design and capture of UV-cross-linked RNA–protein complexes. (**a**) Schematic of single tiling path biotinylated DNA oligos complementary to the target RNA. Regions of the target RNA that are low complexity or contain repetitive genetic elements are masked, and probes are not designed to target these sequences. (**b**) Target RNA that has been cross-linked to directly interacting proteins in living cells is captured and enriched by hybridization with a complementary set of biotinylated DNA oligonucleotide probes. These probes are captured and separated from lysate on magnetic streptavidin beads, while nonspecific RNA and proteins are removed with a denaturing wash, so that only directly interacting RNA and protein factors remain in the purified complexes

2 Materials

2.1 Specialized Equipment and Materials

1. Custom 5′ biotin-modified DNA oligonucleotides complementary to target RNA sequence.
2. Sonicator with microtip.
3. Thermomixer or other incubating shaker.
4. Magnetic racks for 96-well plates, Eppendorf tubes, and 50 mL tubes.
5. TURBO DNase or other salt-tolerant DNase.
6. SpeedVac or other vacuum lyophilizer.
7. Streptavidin-conjugated magnetic beads.
8. Protease inhibitor cocktail set III, EDTA-free (can be obtained from EMD Millipore).
9. Murine RNase inhibitor (can be obtained from New England Biolabs).
10. Ethylenediaminetetraacetic acid (EDTA).
11. Egtazic acid (EGTA).
12. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP).
13. RLT buffer (can be obtained from Qiagen).
14. Acetone.

15. Trichloroacetic acid (TCA).
16. Benzonase nuclease.
17. Silane-modified magnetic beads.
18. HiPPR detergent removal columns (can be obtained from Pierce).
19. Sequencing grade modified trypsin.
20. ZipTips with C_{18} resin (can be obtained from EMD Millipore).
21. RapiGest SF anionic surfactant (can be obtained from Waters Corporation).
22. Ammonium bicarbonate.
23. Iodoacetamide.
24. Hydrochloric acid (HCl).
25. Sequencing grade trypsin (can be obtained from Promega).
26. Formic acid.
27. Acetonitrile.
28. Ammonium hydroxide.

2.2 Materials for Tissue Culture and Cell Harvesting

1. Complete IMDM culture medium with FBS and L-glutamine.
2. UV cross-linker with 254 nm wavelength bulbs.
3. Corning disposable spinner flasks, 125 mL and 1 L volumes.

2.3 Buffers for Cell Lysis and RNA-Protein Complex Purification

1. Total cell lysis buffer: 10 mM Tris–HCl, pH 7.5, 500 mM LiCl, 0.5% n-dodecyl β-D-maltoside, 0.2% sodium dodecyl sulfate, and 0.1% sodium deoxycholate.
2. DNase salt stock: 500 mM $MgCl_2$ and 100 mM $CaCl_2$.
3. 4 M urea hybridization buffer: 10 mM Tris–HCl pH 7.5, 500 mM LiCl, 5 mM ethylenediaminetetraacetic acid, 0.5% n-dodecyl β-D-maltoside, 0.2% sodium dodecyl sulfate, 0.1% sodium deoxycholate, 4 M urea, and 2.5 mM Tris (2-carboxyethyl) phosphine.
4. 6 M urea buffer: 10 mM Tris–HCl pH 7.5, 500 mM LiCl, 5 mM ethylenediaminetetraacetic acid, 0.5% n-dodecyl β-D-maltoside, 0.2% sodium dodecyl sulfate, 0.1% sodium deoxycholate, 6 M urea, and 2.5 mM Tris (2-carboxyethyl) phosphine.
5. NLS elution buffer: 20 mM Tris–HCl pH 8.0, 10 mM ethylenediaminetetraacetic acid, 2% N-lauroylsarcosine, 2.5 mM Tris (2-carboxyethyl) phosphine.
6. Benzonase elution buffer: 20 mM Tris–HCl pH 8.0, 0.05% N-lauroylsarcosine, 2 mM $MgCl_2$, and 2.5 mM Tris (2-carboxyethyl) phosphine.

7. 10× TNE buffer: 50 mM Tris–HCl, 1000 mM NaCl, 1 mM EDTA).
8. Peptide resuspension buffer: 5% acetonitrile and 0.5% trifluoroacetic acid.

3 Methods

3.1 Growth of Large-Scale Cultures of Cancer Cells

1. Obtain suspension cells from American Type Culture Collection or other reliable supplier that performs regular genotyping of cell lines. We use K562 human leukemia cells in these experiments.
2. Thaw K562 cells from frozen stock into 9 mL filter sterilized IMDM medium.
3. Culture at 4×10^5 cells/mL in T75 flasks with 10 mL of medium.
4. Count cells, and if live cells make up >95% of the culture, then seed a 125 mL stirred flask at 4×10^5 cells/mL. To seed the stirred flask, combine 26 mL of fresh medium and 26 mL of cells for 52 mL total volume.
5. Bring culture volume up to 100 mL in 125 mL spinner flask.
6. Grow cells to around 1×10^6 cells/mL. This usually takes 2–3 days.
7. Cells can be harvested from 125 mL spinner flasks according to the protocol below. For larger culture volumes, combine two 125 mL spinner cultures into a 1 L spinner flask and dilute with fresh medium to 4×10^5 cells/mL (Fig. 2a).

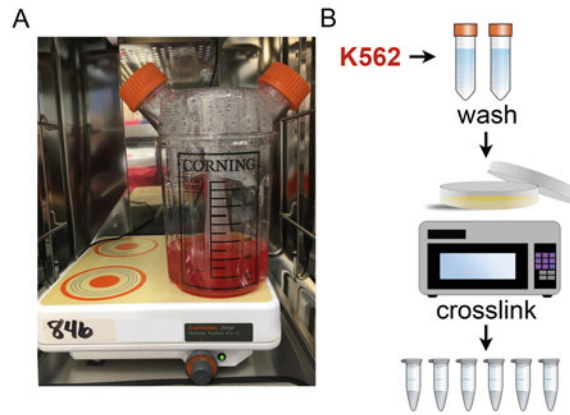


Fig. 2 Cross-linking and harvesting of cell cultures. (**a**) K562 cell cultures grown in 1 L spinner flask on a magnetic stir plate inside a CO_2 incubator. (**b**) Schematic overview of protocol to harvest, wash, cross-link, and pellet the cultured human cells to prepare for cell lysis and capture experiments

8. Keeping sterile, the cells can now be harvested at 1×10^6 cell/mL by pouring off desired volume of culture and replacing with fresh medium to a density of 4×10^5 cells/mL every 2–3 days. Up to 1×10^7 cells can be harvested at a time with this method by one person in approximately 2 h (1 L of culture at 1×10^6 cells/mL).

3.2 Cross-Linking and Harvesting Cells

1. For each 5×10^7 cells to be harvested, use the following volumes: collect and wash cells with 12.5 mL of ice-cold 1× phosphate buffered saline (PBS).
2. Remove the first wash and resuspend cell pellet in 12.5 mL of ice-cold 1× PBS.
3. Prepare trays of crushed ice (cafeteria tray, or other similar plastic tray) that will fit inside the cross-linking instrument. We use a Spectrolinker with 254 nm wavelength bulbs (Fig. 2b).
4. Pipet 10 mL of cell suspension onto a clear 15 cm culture plate with lid removed. Place plates onto tray of crushed ice inside the Spectrolinker, leaving lids off.
5. Perform UV cross-linking of cells at 254 nm for 0.4 J/cm^2.
6. Rock the plate gently to mix cell suspension, then repeat the cross-linking under the same conditions for a total dose of 0.8 J/cm^2.
7. Scrape the plate with a cell lifter and collect cells into a clean 15 mL tube on ice.
8. Centrifuge at 600 × *g* for 4 min at 4 °C to pellet cells. Aspirate the supernatant.
9. Resuspend cell pellets in 1 mL of ice-cold 1× PBS for each 5×10^7 cells.
10. Centrifuge at 1000 × *g* for 4 min at 4 °C to pellet cells. Aspirate the supernatant.
11. Flash freeze cell pellets in liquid nitrogen or a dry ice/ethanol bath and store at −80 °C. Adjust the final resuspension and aliquot volumes as needed to achieve the desired cell number in each pellet.

3.3 Designing Antisense Capture Probes

1. Determine desired length for 5′ biotinylated DNA oligonucleotide capture probes. We have successfully used 120mer, 90mer, 70mer, or 60mer sequences to capture target RNAs. Control probe sequences for U1 snRNA (positive) and luciferase (negative) are shown in Table 1. Shorter probes have decreased cost per probe but may result in increased variability in target RNA recovery in individual experiments (Fig. 3). Probes are designed in a similar manner as described in previous antisense capture protocols [6], but we currently use

Table 1
Sequences of positive and negative control oligonucleotide DNA capture probes

Name	DNA sequence 5′-3′
hU1-120-1	[Biotin~5]AGTCGAGTTTCCCACATTTGGGGAAATCGCAGGGGTCAGCACATCCGGAGTGCAATGGATAAGCCTCGCCCTG GGAAAACCACCTTCGTGATCATGGTATCTCCCCTGCCAGGTAAGTAT
hU1-120-2	[Biotin~5]TACCACAAATTATGCAGTCGAGTTTCCCACATTTGGGGAAATCGCAGGGGTCAGCACATCCGGAGTGCAATGGATA AGCCTCGCCCTGGGAAAACCACCTTCGTGATCATGGTATCTCCC
hU1-120-3	[Biotin~5]AACGCAGTCCCCCACTACCACAAATTATGCAGTCGAGTTTCCCACATTTGGGGAAATCGCAGGGGTCAGCACATCC GGAGTGCAATGGATAAGCCTCGCCCTGGGAAAACCACCTTCGTG
hU1-70-1	[Biotin~5]TGCAATGGATAAGCCTCGCCCTGGGAAAACCACCTTCGTGATCATGGTATCTCCCCTGCCAGGTAAGTAT
hU1-70-2	[Biotin~5]CCCACTACCACAAATTATGCAGTCGAGTTTCCCACATTTGGGGAAATCGCAGGGGTCAGCACATCCGGAG
hU1-70-3	[Biotin~5]CAGGGGAAAGCGCGAACGCAGTCCCCCACTACCACAAATTATGCAGTCGAGTTTCCCACATTTGGGGAAA
hU1-60-1	[Biotin~5]AAGCCTCGCCCTGGGAAAACCACCTTCGTGATCATGGTATCTCCCCTGCCAGGTAAGTAT
hU1-60-2	[Biotin~5]AGTCGAGTTTCCCACATTTGGGGAAATCGCAGGGGTCAGCACATCCGGAGTGCAATGGAT
hU1-60-3	[Biotin~5]CAGGGGAAAGCGCGAACGCAGTCCCCCACTACCACAAATTATGCAGTCGAGTTTCCCACA
Luc-70-1	[Biotin~5]CTCCAGCGGTTCCATCCTCTAGAGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCAT
Luc-70-2	[Biotin~5]AAGGGCCACACCCTTAGGTAACCCAGTAGACCCAGAGGAATTCATTATCAGTGCAATTGTTTTGTCACGA
Luc-70-3	[Biotin~5]GCGTCGAAGATGTTGGGGTGTTGTAACAATATCGATTCCAATTCAGCGGGGGCCACCTGATATCCTTTGT
Luc-60-1	[Biotin~5]TCCATCCTCTAGAGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCAT
Luc-60-2	[Biotin~5]CCCTTAGGTAACCCAGTAGACCCAGAGGAATTCATTATCAGTGCAATTGTTTTGTCACGA
Luc-60-3	[Biotin~5]TGTTGGGGTGTTGTAACAATATCGATTCCAATTCAGCGGGGGCCACCTGATATCCTTTGT

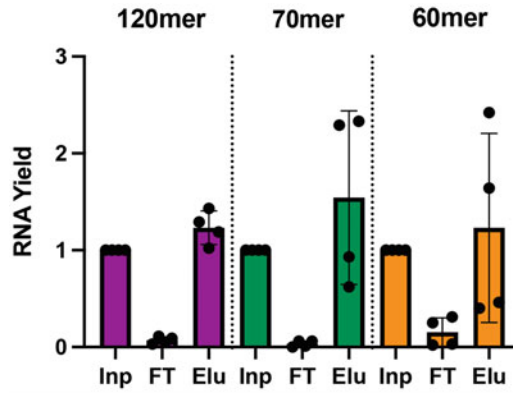


Fig. 3 Target RNA recovery with different probe lengths. Four independent captures of U1 snRNA were performed with 120mer, 70mer, and 60mer biotinylated complementary DNA oligonucleotide probes, and recovery of the target RNA in the input, flow-through, and elution samples was plotted. Increased variability of target RNA yield and decreased depletion from flow-through sample were noted with the shortest DNA probes, although these probes were still able to capture target RNA under the denaturing buffer conditions

shorter probes with only a single tiling path (continuous probe coverage) along the target RNA sequence to reduce cost and number of probes purchased (Fig. 1a).

2. Design potential probe sequences across the target RNA sequence. Use Repeat Masker (https://www.repeatmasker.org) and Tandem Repeat Finder (https://tandem.bu.edu) to mask any potentially problematic sequence regions.
3. Remove any probes containing greater than 8 bases of repetitive or low complexity sequence, homopolymers greater than 8 bases in length, or greater than 12 of the same nucleotide within a 15 nucleotide window.
4. Query the reference genome with potential probe sequences in a BLAT search (a web version can be accessed through the UCSC genome browser), and remove any probes with matches greater than 25 nucleotides in length to other locations in the genome.
5. Ensure that final oligo sequences are the reverse complement of the target RNA sequence.
6. Purchase DNA oligonucleotide capture probes from IDT or similar supplier with a 5′ Biotin-C6 spacer modification. For ease of handling, the oligonucleotides can be requested to be shipped at a standardized concentration of 1 mM in RNase-free water.
7. Mix all individual probes into a single pool and dilute with ultrapure water to a final concentration of around 350 ng/μL for working stocks. Probe mixes may be stored at −20 °C (*see* **Note 1**).

3.4 Preparation of Whole Cell Lysates

1. The following volumes are for each 5×10^7 cell pellet; scale accordingly to number of cells and samples needed for the experiment (*see* **Note 2**). To each cell pellet, add 11.5 μL of Protease Inhibitor Cocktail and 28.75 μL murine RNase inhibitor.
2. Resuspend cell pellet in 2.25 mL of ice-cold total cell lysis buffer. Transfer suspension from microcentrifuge tube to a 50 mL conical tube.
3. Incubate cell suspension on ice for 10 min. During this incubation, pass cell lysate through an 18-gauge needle at least six to eight times to shear membranes and DNA strands (*see* **Note 3**).
4. Sonicate cell lysate at 3.5 W power output, 35% duty. Perform two sets of sonication at 1 min each, with a 10-min incubation on ice in between sets.
5. Add 12 μL DNase salt stock and 37.5 μL (75 U) TURBO DNase. Incubate at 37 °C for 15 min.
6. Return the tube to ice and immediately quench by adding 50 μL of 0.5 M EDTA, 25 μL of 0.5 M EGTA, and 7.5 μL of 1 M TCEP.
7. Add 5 mL of 6 M urea buffer or until the lysate sample has a final concentration of 4 M urea.
8. Incubate on ice for 15 min with periodic mixing.
9. Centrifuge at 8500 × *g* for 20 min at 4 °C to pellet insoluble cell matter.
10. Transfer the supernatant to a fresh tube.
11. Perform preclear step to reduce background. For each 5×10^7 cells, transfer 250 μL magnetic streptavidin beads to microcentrifuge tube. Separate and discard the supernatant from beads.
12. Wash beads three times with 250 μL of 10 mM Tris–HCl pH 7.5 and two times with 250 μL of 4 M urea hybridization buffer.
13. After removing last wash, resuspend beads in the lysate, and incubate at 45 °C for 30 min with intermittent shaking.
14. Separate beads and discard. Save the supernatant as the precleared whole cell lysate sample.
15. Flash freeze precleared whole cell lysate in liquid nitrogen, and store at −80 °C until ready for RNA capture.

3.5 Capture and Purification of UV Cross-Linked RNA–Protein Complexes

1. Thaw precleared whole cell lysate at 45 °C on thermomixer.
2. Denature biotinylated DNA probe mix at 85 °C for 3 min, then immediately place on ice. For U1 snRNA experiments shown here, we used 5 μg of DNA oligonucleotide probes per capture from 5×10^7 cells worth of lysate (*see* **Note 4**).

3. Transfer probes to lysate and incubate at 45 °C for 2 h with 1000 rpm intermittent mixing.
4. During the last 20 min of incubation, prepare 500 µL magnetic streptavidin beads by washing twice with 500 µL of 10 mM Tris–HCl pH 7.5, then twice with 500 µL of 4 M Urea hybridization buffer (*see* **Note 5**).
5. After the incubation, transfer 1×10^5 cells worth of lysate to PCR strip tube and save at −20 °C as "*RNA Input + Probe*" sample.
6. Transfer washed beads to lysate and incubate at 45 °C for 30 min with 1000 rpm intermittent mixing.
7. Magnetically separate beads from the supernatant. Transfer 1×10^5 cells worth of supernatant to PCR strip tube and save at −20 °C as "*RNA Flow-Through*" sample. Save the remaining supernatant at −20 °C.
8. Wash beads three times with 500 µL of 4 M urea hybridization buffer. Incubate each wash at 45 °C for 5 min with 1000 rpm intermittent mixing.
9. Before removing last wash, split sample into "*RNA Elution*" and "*Protein Elution*" samples. Transfer 1×10^6 cells worth of beads to a fresh PCR strip tube and save at −20 °C as "*RNA Elution*" sample. The remaining beads will be used for the elution of captured proteins.

3.6 Elution of Captured RNA

1. For "*RNA Elution*" sample, magnetically separate and discard the wash.
2. Resuspend beads in 18 µL of NLS elution buffer.
3. Heat the beads at 95 °C for 2 min.
4. Dilute "*RNA Input + Probe*" and "*RNA Flow-Through*" samples with NLS elution buffer to 18 µL final volume.
5. For all samples, digest proteins by adding 2 µL Proteinase K and incubate at 50 °C for 1 h.
6. Magnetically separate and transfer the supernatant from "*RNA Elution*" sample to a new PCR strip tube.
7. Purify RNA samples using silane-modified magnetic beads:
   (a) Per 20 µL RNA sample, transfer 20 µL silane beads to a PCR strip tube. Magnetically separate and discard the supernatant.
   (b) Resuspend beads in 60 µL of RLT buffer.
   (c) Transfer beads to RNA sample and mix by pipetting.
   (d) Add 120 µL of 100% ethanol to each sample and mix by pipetting.

(e) Incubate 2–5 min to allow for nucleic acids to bind to the silane beads.

(f) Wash beads twice with 120 μL of 70% ethanol.

(g) Magnetically separate and discard the supernatant. Allow beads to air-dry until the bead surface is matte (*see* **Note 6**).

8. Perform DNase treatment by resuspending beads in 19 μL of 1× TURBO DNase buffer and adding 1 μL (2 U) TURBO DNase. Incubate at 37 °C for 20 min.
9. Perform a second cleanup by repeating **steps 7b** to **7g** without adding new beads. Add the RLT buffer directly to the 20 μL DNase-treated sample with existing beads.
10. After the second cleanup, elute the RNA in 10 μL of ultrapure water.
11. Reverse transcribe the RNA to complementary DNA (cDNA) using preferred protocol, and use quantitative real-time PCR (qPCR) to analyze captured RNA (*see* **Note 7**).

3.7 Elution of Captured Protein

1. For "*Protein Elution*" sample, magnetically separate and discard the last wash from remaining beads.
2. Resuspend beads in 500 μL of benzonase elution buffer.
3. Add 25–100 U of benzonase nuclease.
4. Incubate at 37 °C for 2 h with 1000 rpm intermittent mixing.
5. Magnetically separate and transfer the supernatant to a fresh microcentrifuge tube. Repeat this **step 5** more times to completely remove all magnetic beads from the sample. Save the last supernatant as the "*Protein Elution*" sample.
6. Add 80% TCA to a final concentration of 20% TCA to protein sample.
7. Briefly vortex the tube and incubate at 4 °C overnight.
8. Centrifuge at 10,000 × *g* for 15 min at 4 °C to pellet the protein.
9. Remove and discard the supernatant. Wash the protein pellet with 1 mL cold acetone.
10. Centrifuge at 10,000 × *g* for 15 min at 4 °C, then remove and discard the wash.
11. Air-dry the protein pellet and store at −20 °C until ready for mass spectrometry preparation or Western Blot analysis (*see* **Note 8**).

3.8 Quantitation of Captured Proteins by Mass Spectrometry

1. Resuspend protein pellet with 50 μL of ultrapure water.
2. Add 5 μL of 5% RapiGest in 10× TNE buffer.
3. Boil the sample at 100 °C for 10 min.

4. Return the tube to ice and add 100 µL of 50 mM ammonium bicarbonate, 2 µL of 0.5 M TCEP, and 5 µL of freshly prepared 180 mM iodoacetamide.
5. Incubate at 37 °C for 5 min.
6. Add 0.2–0.5 micrograms of trypsin to the sample and incubate on 37 °C overnight with shaking at 600 rpm.
7. Add 20 µL of 1 M hydrochloric acid and incubate at 37 °C for 30 min.
8. Centrifuge at 4 °C for 30 min at 16,000 × *g*.
9. Transfer the supernatant to a new tube and add 10 µL of 1.5 M ammonium hydroxide to neutralize.
10. Use HiPPR resin spin column to remove any remaining detergent in the sample, according to manufacturer's instructions.
11. Lyophilize sample overnight in SpeedVac.
12. Add 10 µL of peptide resuspension buffer to solubilize the dry peptide pellet.
13. Use a Zip Tip with C_{18} resin to desalt the sample following the manufacturer's instructions.
14. Transfer purified peptide sample to new tube and lyophilize in SpeedVac to dry completely.
15. Store the sample at −20 °C until ready for mass spectrometry.
16. Resuspend peptides in 0.2% formic acid and 5% acetonitrile. Analyze sample in 60-min to 90-min run with LC-MS (*see* **Note 9**).

3.9 Mass Spectrometry Data Analysis

1. Download and install MaxQuant (https://maxquant.org).
2. In "Raw Data" tab: Load MS raw data files, then select "Set Experiment" to rename the files in Experiment field.
3. In the "Group-Specific Parameters" tab, leave default settings. Type: multiplicity = standard, Labels = 1, Modification: variable modifications = Oxidation (M) + Acetyl (Protein N-term), fixed modification = Carbamidomethyl (C), max. number of modification per peptides = 5, Instrument: Default; First search: leave it blank; Digestion: Digestion mode = Specific; Enzyme = Trypsin/P; Max.missed = 2; Label-free quantification = iBAQ (if desired); Misc.: Isobaric weight exponent = 0.75.
4. Download the reference tryptic proteome file from the NCBI database. In "Group-Specific Parameters" tab, under "Sequences" subsection, load the file UP000005640_9606.fasta (for human samples) in FASTA file path; leave remaining parameters on default settings. In "Identification" tab, select 2 for "Min. unique peptides" field. In "Label-Free Quantification" subsection, check the box labeled "iBAQ."

Table 2
Table of some common contaminant proteins identified in luciferase control captures from human leukemia cells

Protein name	Gene name
Pyruvate carboxylase, mitochondrial	PC
Propionyl-CoA carboxylase alpha chain, mitochondrial	PCCA
Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	MCCC1
Elongation factor 1-alpha 1	EEF1A1
Heat shock protein HSP 90-beta	HSP90AB1
Alpha-enolase	ENO1
Tubulin alpha-4A chain	TUBA4A
Desmoplakin	DSP
Pyruvate kinase PKM	PKM
ADP/ATP translocase 3	SLC25A6
60 kDa heat shock protein, mitochondrial	HSPD1
Elongation factor 2	EEF2
Putative adenosylhomocysteinase 2	AHCYL1
Calmodulin-like protein 5	CALML5
T-complex protein 1 subunit eta	CCT7

5. Choose number of processors depending on configuration of the computer, then click "Start" to run analysis.
6. When MaxQuant analysis is complete, a folder named "combined" will be generated in the same location as the raw files. The list of captured proteins is in a file named "proteinGroups" inside of the "txt" folder in the "combined" folder.
7. Remove common contaminants that are also found in control captures (Table 2) and identify top proteins by intensity (*see* **Note 10**).

4 Notes

1. *Testing oligonucleotide DNA capture probes.* For new probe sets, perform a test RNA capture from 500 ng of purified total RNA (purified with Qiagen RNeasy Kit, or similar preparation method) before moving onto cell lysate samples, to ensure that the biotinylated DNA oligonucleotides can capture and purify the target RNA as expected.

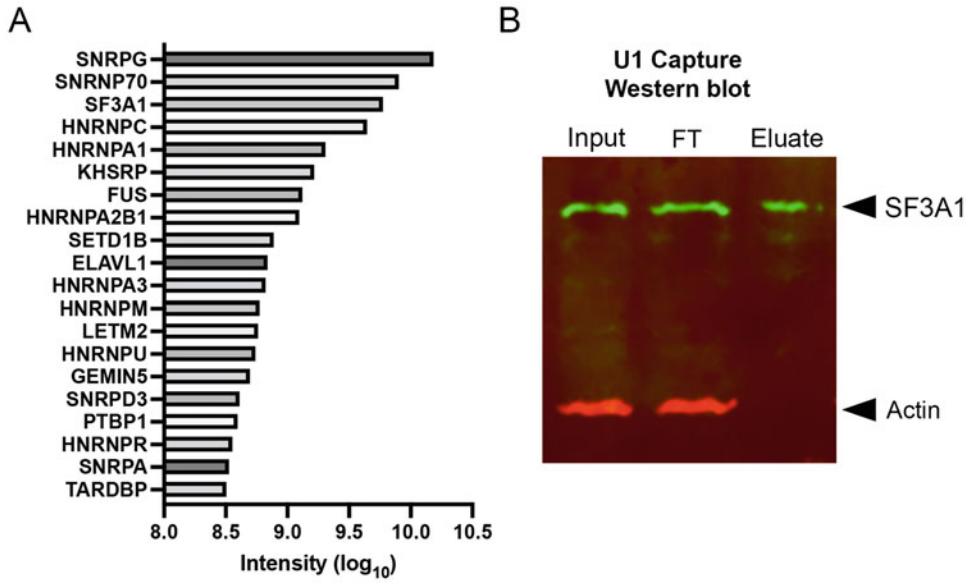


Fig. 4 Detection of U1 snRNA-interacting proteins recovered after capture. (**a**) Top 20 proteins identified in U1 snRNA capture. The known U1-interacting factors are of highest intensity in the sample, and their identification does not require use of SILAC or other specific protein labeling methods. (**b**) Western blot validation of a top U1-interacting protein from the mass spectrometry experiment, SF3A1 (green), which is recovered in the elution. The nonspecific protein actin (red) is found in the flow-through but is not enriched in the protein elution sample

2. *Experimental design*. We generally use between 5×10^7 and 2×10^8 cells for each RNA capture. Always include a positive control (such as U1 snRNA capture) and negative control (such as luciferase capture or non-cross-linked sample) in each experiment to ensure that all parts of the protocol are working as expected. Expected results from U1 snRNA capture are shown in Fig. 4 and common background proteins found in luciferase captures are shown in Table 2.
3. *Cell lysis*. Take care to ensure that the entire cell pellet has been well-solubilized in the lysis buffer by pipetting and passing the cell pellet through the needle. If clumps of cells remain after these steps, the cell lysis may not be efficiently completed and can lead to increased background protein identification.
4. *Quantity of beads and probe*. We have tested the calculated number of beads and probe needed based on the manufacturer's information. However, these quantities are not sufficient for U1 capture, likely due the inability of several large probe and bead complexes to access each bead surface. Empirical testing of bead and probe quantities should be optimized for each target.

5. *Pipetting to wash beads.* Slowly pipette the 4 M urea hybridization buffer on the beads to prevent the creation of detergent bubbles. Excessive bubbles can cause magnetic beads to be trapped and lost throughout the protocol. Samples can be mixed gently by moving the tubes back and forth against the magnetic rack instead of by vigorous pipetting.
6. *Silane cleanups.* A common point of failure is loss of the RNA sample at the silane bead cleanup step. Practice this portion of the protocol with total RNA to ensure that capture of RNA can be achieved. At the last step of the cleanup, avoid overdrying the silane beads. Allow beads to dry until just after they have stopped shining when viewed from an angle; the beads should appear matte but not be allowed to overdry. Otherwise, it can become difficult to elute the RNA from the beads.
7. *Captured RNA yield.* Calculated yields of RNA by qPCR do not always add up to 100% in FT and elution samples as would be expected. As an indication that the experiment is working, we confirm that the nonspecific RNA target (Tubulin, GAPDH, other housekeeping gene) is not captured by the DNA probes from the flow-through and that the specific RNA target (U1, or experimental sample) is captured and enriched significantly in the elution sample. The variability in detection of the target RNA between the flow-through and elution samples might be due to decreased complexity of the sample in elutions or to residual DNA probes that cannot be easily removed from the eluted RNA sample.
8. *Total protein yield.* The total protein recovered from this protocol is generally around 200 nanograms to 1 microgram of protein per capture. Input cell quantities can be adjusted for RNAs of varying abundance, to ensure that enough protein is captured for detection by mass spectrometry. The target RNA can also be overexpressed to increase its abundance in cells; previously, similar proteins were identified from overexpressed Xist RNA captures and endogenous Xist transcript captures in different cell types [4,7].
9. *Expected results.* The total number of identifications obtained per sample is around 200–500 proteins, with the top ~20–50 proteins generally being those specific for the RNA of interest. The complexity and number of proteins recovered after capture varies depending on the life cycle of the target RNA and how many different proteins interact with the specific RNA over its cellular lifetime. Ribosomal RNA capture recovers >150 specific proteins while U1 snRNA capture recovers ~30 specific proteins, including those in the U1 stable snRNP complex as well as known U1-interacting factors such as FUS and processing factors including GEMIN5. While SILAC (stable isotope

labeling with amino acids in cell culture) can be very useful for obtaining comparative quantitative information on protein levels, the purification of RNA–protein complexes by oligonucleotide captures often results in very different lists of protein identifications for each sample. The necessity to use a reference sample for SILAC ratio quantitation means that some proteins will not be quantified, because those proteins do not exist in the comparison sample. Thus, some real interactors may be missed in the SILAC data analysis. Since the purification of cross-linked protein and target RNA is achieved under strongly denaturing buffer conditions, the background protein levels are very low and it is often sufficient to use only the identified protein intensities to determine the top interacting factors for a target RNA. The data shown for U1 snRNA in Fig. 4 are based only on protein intensities from mass spectrometry experiments, not on SILAC ratios.

10. *Removing contaminant proteins.* MaxQuant will indicate common contaminant proteins such as keratins, trypsin, and streptavidin in the "potential contaminant" column of the "proteinGroups" file. Use caution when investigating any proteins that appear on the luciferase probe sample list because it can mean that they are nonspecific interactions. Naturally biotinylated proteins including histones, carboxylases, and mitochondrial matrix proteins can also be removed from the candidate list, since these will be enriched on streptavidin beads.

Acknowledgments

This work was supported by NIH NIGMS Grant GM120494 to Colleen A. McHugh, NCI Cancer Training grant T32 CA009523 to Tong Su, University of California Cancer Research Coordinating Committee grant support C21CR2119 for Calvin Huang, and NIGMS Molecular Biophysics Training Grant T32 GM008326 to Simone Hall.

References

1. McHugh CA, Russell P, Guttman M (2014) Methods for comprehensive experimental identification of RNA-protein interactions. Genome Biol 15(1):203. https://doi.org/10.1186/gb4152
2. Graves PR, Haystead TA (2002) Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev 66(1):39–63; table of contents. https://doi.org/10.1128/MMBR.66.1.39-63.2002
3. Mili S, Steitz JA (2004) Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. RNA 10(11):1692–1694. https://doi.org/10.1261/rna.7151404
4. McHugh CA, Chen CK, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, Sweredoski MJ, Shishkin AA, Su J, Lander ES, Hess S,

Plath K, Guttman M (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521(7551): 232–236. https://doi.org/10.1038/nature14443

5. Greenberg JR (1979) Ultraviolet light-induced crosslinking of mRNA to proteins. Nucleic Acids Res 6(2):715–732

6. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, Plath K, Guttman M (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341(6147): 1237973. https://doi.org/10.1126/science.1237973