Published in final edited form as:

Methods. 2012 October ; 58(2): 88–93. doi:10.1016/j.ymeth.2012.06.020.

Biochemical Analysis of Long Non-coding RNA-containing Ribonucleoprotein Complexes

Chenguang Gong, Maximilian Wei-Lin Popp, and Lynne E. Maquat*

Department of Biochemistry and Biophysics, School of Medicine and Dentistry, Center for RNA Biology, University of Rochester, Rochester, New York 14642, USA

Abstract

Long non-coding RNAs (lncRNAs), once relegated to junk products of the genome, are becoming better appreciated for the myriad functions they play in cellular processes. It is clear that for most of the cases studied, lncRNAs carry out their functions at least in part through interactions with proteins. Here we present two complementary biochemical methods for the analysis of lncRNA-containing ribonucleoprotein complexes, hereafter referred to as RNPs. The first strategy offers users the ability to purify RNPs based on a protein component and to analyze the spectrum of lncRNAs, other proteins, and, if present, other types of RNAs that are bound to it. The second makes use of a bacteriophage MS2 binding-site affinity-handle grafted onto a lncRNA of interest to investigate the proteins and RNAs that co-purify with the tagged RNA.

Keywords

long non-coding RNA (lncRNA); immunoprecipitation; MS2 coat-binding protein; ribonucleoprotein complex (RNP); RT-PCR; biochemistry

1. Introduction

Recently it has become clear that a significant portion of the transcriptome encodes gene products that are likely unable to be translated into proteins. Such transcripts greater than ~200-nt in length are termed long non-coding RNAs (lncRNAs), and they may either lie in intergenic regions (lincRNAs) or overlap with protein-coding genes. That the number of lncRNAs encoded by the human genome likely far outstrips the number of protein-coding mRNAs [1] indicates that lncRNAs generally are the products of an impressive level of faulty transcription or, alternatively, serve yet-to-be-discovered cellular functions. Differentiating between these two possibilities requires the careful application of biochemical techniques. For those few lncRNAs that have an ascribed function [2], a combination of genetic and biochemical experiments has moved the field forward and illustrated the pervasive nature of functional lncRNAs. We now know, for example, that lncRNAs such as HOTAIR can modulate the state of chromatin by recruiting the Polycomb chromatin remodeling complex PRC2 [3]. Others such as Evf2 can activate transcription by recruiting transcription factors [4], and still others such as MALAT1 impinge upon pre-

© 2012 Elsevier Inc. All rights reserved.

*Correspondence: lynne_maquat@urmc.rochester.edu, Phone: 585-273-5640, Fax: 585-271-2683.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

mRNA splicing by sequestering splicing factors [5]. lncRNAs can affect not only RNA synthesis but also RNA degradation. For example, lncRNAs may duplex with mRNAs targeted for Staufen1-mediated mRNA decay and in so doing form a binding platform that recruits the double-stranded RNA-binding protein Staufen1 [6].

The importance of lncRNAs is underscored by their link to human diseases [7]. In certain metastatic types of non-small-cell lung cancer (NSCLC), MALAT-1 is up-regulated [8]. By recruiting PRC2 to particular genes, HOTAIR can modulate the metastatic breast-cancer program [3]. In part fueled by demonstrating for a handful of lncRNAs their function and involvement in disease, new candidate lncRNAs are rapidly being identified using high-throughput computational methods [9,10]. Although such methods are powerful at predicating the existence of lncRNAs, elucidating the function of each lncRNA in detail falls under the purview of conventional biochemistry. Here we describe two methods for the biochemical analysis of lncRNAs that can be used for just this purpose (Figure 1).

2. Description of methods

2.1 Enumeration of lncRNAs associated with a particular protein

Protein complexes can be immunoprecipitated from a heterogeneous cellular homogenate provided that a suitable antibody exists and non-denaturing conditions are employed. Any RNAs, including lncRNAs that are stably associated with the complex, will also be isolated, and the abundance of particular lncRNAs can be determined using polymerase chain reaction (PCR)-based methods. Although low-throughput, this approach is inexpensive, requires no specialized equipment, and allows for careful quantitation of the relative levels of co-precipitating lncRNAs.

2.1.1 PCR primer design—Levels of candidate lncRNAs present in immunoprecipitated RNPs may be measured using reverse-transcription coupled to either semi-quantitative PCR (RT-sqPCR, see section 2.1.2) or to real-time quantitative PCR (RT-qPCR). In either case, primer design is a critical aspect of the experimental procedure and must be optimized before subjecting precious immunoprecipitated samples to analysis. Like mRNAs, lncRNAs are subject to post-transcriptional processing events (splicing, capping, polyadenylation, etc.). Therefore, special care must be taken to design primers to detect not only the desired lncRNA but also only the desired form of the lncRNA. For detecting mature, spliced lncRNAs, primer sets are designed with at least one primer annealing to a region straddling an exon-exon junction and the other avoiding introns. For detecting unspliced lncRNAs, primers are designed such that at least one anneals to an intronic region. Primer sets used for RT-PCR analysis typically adhere to the following rules: (*i*) the T_m should be $\sim 60^\circ\text{C}$, (*ii*) the G-C content should be $\sim 50\%$, and (*iii*) the product size should be 200–400 bp.

2.1.2 Immunoprecipitation and RT-PCR—Immunoprecipitation (IP) of RNPs proceeds under mild, non-denaturing conditions in order to preserve protein–protein, protein–RNA and RNA–RNA interactions. Dividing cells are harvested and washed with ice-cold phosphate-buffered saline (PBS). Gentle hypotonic lysis is achieved by incubating cell pellets in lysis buffer (20mM Tris-Cl pH7.4, 10mM NaCl, 2mM EDTA, 0.5% Triton X-100) supplemented with an RNase inhibitor (40 U/ml RNaseOUT, Invitrogen) and protease inhibitors (2mM benzamidine hydrochloride, 1mM PMSF, 1X Roche protease inhibitor tablet) for 10 minutes at 4°C , followed by supplementation with 150–300mM NaCl and further incubation for 5 minutes on ice. Increased salt concentrations in the lysis and subsequent wash steps may decrease the level of contaminating RNAs at the expense of also decreasing the level of specific RNAs. Insoluble material is then separated from the soluble homogenate by centrifugation ($18000 \times g$, 10 minutes, 4°C). Complete separation of

insoluble material is critical to eliminate background in the downstream RT-PCR analysis. After clarifying the homogenate, total-protein content is quantitated using, e.g., the Bradford Assay. Next, lysates are pre-cleared using 50 μ l of Protein-A agarose or Protein-G agarose beads (depending on the origin of the antibody used and, in either case, washing extensively in lysis buffer that contains salt) per 4–8 mg of lysate in an effort to further reduce background. Samples are incubated for 1 hour at 4°C with end-over-end rotation followed by removal of the beads by centrifugation. RNA is extracted from 1/20th of the supernatant, and another 1/20th of the pre-cleared sample is saved and used later in using Western blotting to estimate the efficiency with which the rest of the sample is immunoprecipitated.

The antibody used in the IP is added to the pre-cleared sample. In general, the amount of antibody needed to achieve maximal IP efficiency must be determined empirically, however, 10 μ g of antibody per 8 mg of cell lysate is an appropriate starting point for optimization. After at least a 2-hour incubation at 4°C with end-over-end rotation, the antibody and lysate mixture is microcentrifuged at maximal speed for 10 minutes to remove any precipitate that formed during the course of incubation. Again, this step is critical to reduce background in the RT-PCR analysis. The supernatant is added to 50 μ l of washed Protein-A or Protein-G agarose beads and incubated for another hour at 4°C. The Protein-A or Protein-G agarose–antibody complexes are then isolated by centrifugation at 3000 $\times$ g for 1 minute, and the depleted supernatant is removed and discarded. The beads are incubated in 1 ml of wash buffer (50mM Tris-Cl pH7.4; 150–300mM NaCl; 0.05% Triton X-100; 1mM PMSF; 2mM Benzamidine, 1X Roche protease inhibitor tablet; 40 U/ml RNaseOUT) for 2 minutes with rotation at 4°C, followed by isolation using centrifugation. This washing procedure is repeated a minimum of 10 times. After the final wash, the buffer is completely removed using a 27-gauge needle, and the beads are incubated in 50 μ l of 2X Laemmli sample buffer for 5 minutes at 95°C. Half of the sample is used for Western blotting to estimate IP efficiency. RNA is extracted from the other half using TRIzol. Glycogen (1 μ g) is added to the RNA extraction as carrier, followed by isopropanol precipitation.

Contaminating DNA must be removed from the extracted RNA (both before and after IP) by incubation with RQ1 (RNA Qualified) RNase-free DNase according to the manufacturer's directions (Promega), followed by phenol-chloroform extraction and sodium acetate/ethanol precipitation. The DNA-depleted RNA samples are used as templates for cDNA synthesis in an RT reaction that is primed with random hexamers and uses Superscript III (Invitrogen) according to the manufacturer's directions. This reaction is easily saturated. Thus, to ensure accurate downstream quantitation and to provide a means of gauging the relative amounts of the candidate lncRNAs in the immunoprecipitated samples, the sample of total-cell RNA is subjected to two-fold serial dilutions, and each dilution is used in a separate RT reaction. Finally, the levels of candidate lncRNAs in each immunoprecipitated sample are determined using radioactive sqPCR and primers designed according to section 2.1.1. Generally, 2 μ l of the RT reaction is used as template in a standard 50 μ l PCR reaction that contains 2 μ Ci of [α -³²P]dATP. Products are resolved in a 5% acrylamide gel and visualized using phosphorimaging. The number of cycles used in the sqPCR step must be determined empirically; for abundant lncRNAs, 19 cycles may suffice for detection, while for less abundant lncRNAs as many as 30 cycles may be necessary. It is critical that the titrated pre-IP RNA samples provide a linear frame of reference for quantitating the lncRNAs present in the immunoprecipitated samples.

2.2 Identification of proteins and/or RNAs associated with a particular lncRNA

This procedure is aimed at retrieving cytoplasmic and polyadenylated lncRNA-containing RNP complexes from mammalian cells through affinity purification of the lncRNA of interest. The technique relies on the interaction of the bacteriophage MS2 coat-protein with a specific viral RNA stem-loop structure that is inserted into or grafted onto the 3' end of the

lncRNA. Full-length lncRNAs to be cloned into a plasmid-based expression vector can be obtained from either cDNA libraries or cDNA syntheses using cellular RNA. Twelve copies of the MS2 coat-protein binding site are then introduced into or downstream of the region encoding the lncRNA to generate a lncRNA_MS2bs expression plasmid. This expression plasmid and a second plasmid encoding a FLAG-tagged MS2 coat-protein are transiently introduced into cultured cells. *In vivo*, FLAG-tagged MS2 coat-proteins specifically bind an MS2 coat-protein-binding site (MS2bs) covalently linked to the lncRNA sequence. Anti-FLAG IP is used for retrieval of the lncRNA_MS2bs RNP. The identity of the protein components can be determined by mass spectrometry, or candidate proteins can be confirmed by Western blotting; the lncRNA and any other RNA components can be determined by RNA-Seq or, using RT-PCR and RNA-specific PCR primers (see section 2.1.2).

2.2.1 Generating the lncRNA_MS2bs expression plasmid—Although lncRNA sequences can generally be obtained from NCBI, in many cases the transcriptional start site of a lncRNA is unknown. To determine the exact 5' end of a lncRNA, RNase protection, nesting RT-PCR, or primer-extension assays can be adopted. Although the lncRNA studied here is polyadenylated, the endogenous poly(A) signal (e.g. AAUAAA) is excluded once the lncRNA cDNA is cloned into a mammalian-cell expression plasmid (e.g. pcDNA3, Invitrogen). Twelve copies of the MS2bs are inserted between the lncRNA and the vector-derived poly(A) signal to generate what will be called a pcDNA3_lncRNA_MS2bs plasmid. As an alternative to using the vector derived poly(A) signal, the endogenous poly(A) signal could be used, with the 12 MS2bs placed between the body of the lncRNA and the endogenous poly(A) site. Two other plasmids, pcDNA3_lncRNA, which lacks any MS2bs, and pcDNA3_FLUC_MS2bs, which contains an irrelevant luciferase (LUC) RNA molecule linked to twelve copies of the MS2bs, are also generated and serve as negative controls in the following steps. The twelve MS2bs may be inserted into either the 5' or 3' end of the lncRNA so that binding of the FLAG-MS2 coat protein does not interfere with the binding of other RNP components.

2.2.2 Optimizing cell transfections—Before starting the purification of lncRNA_MS2bs RNP complexes, two test transfections should be done. In the first, cells should be transfected with different amounts of the pcDNA3_lncRNA_MS2bs plasmid. The level of plasmid-derived lncRNA should not exceed that of the endogenous lncRNA since overexpression can result in the formation of nonphysiological RNPs and complicate interpretation of the data. In the second, cells should be cotransfected with the optimized amount of pcDNA3_lncRNA_MS2bs and different amounts of pFLAG-MS2 (FLAG-tagged MS2 coat-protein expression plasmid), and an anti-FLAG IP of the resulting lysates should be performed (see section 2.1.2). By so doing, the co-IP efficiency of the lncRNA_MS2bs with the FLAG-MS2 coat-protein can be determined. To achieve maximal recovery of lncRNA_MS2bs-containing complexes, the cellular ratio of free lncRNA_MS2bs to FLAG-MS2-bound lncRNA_MS2bs should be minimized by increasing the amount of pFLAG-MS2 plasmid used in the transfection. The cellular ratio of free FLAG-MS2 protein to lncRNA_MS2bs-bound FLAG-MS2 should also be minimized, however, by decreasing the amount of pFLAG-MS2 used in transfection. The goal of these experiments is to provide an optimal range where unbound forms of both exogenously expressed components — the FLAGMS2 and lncRNA_MS2bs — are minimized.

2.2.3 Formaldehyde crosslinking prior to cell lysis—To increase the specificity of lncRNA_MS2bs-containing RNP recovery, transfected cells are crosslinked using formaldehyde. Formaldehyde is a very short-length reversible amine-reactive crosslinker that is used to covalently capture cellular protein-RNA interactions. Cells are washed with

room-temperature 1X PBS and then incubated in 1% formaldehyde (in 1X PBS) for 10 minutes at room temperature. Crosslinking is subsequently quenched by adding glycine to a final concentration of 0.25M and incubating for 5 minutes at room temperature. Cells are harvested and resuspended in salt-supplemented lysis buffer (section 2.1.2). Sonication is required for sufficient lysis of the crosslinked cells. As an example, six-to-eight rounds of 30 bursts each with a Branson micro-tip sonicator using an output of 4 and a duty cycle of 30% efficiently lyses > 90% of the cells.

The anti-FLAG IP is performed as described (section 2.1.2) except the wash buffer includes 500 mM NaCl to reduce the background that accompanies formaldehyde crosslinking.

Incubation of the immunoprecipitated sample for 5 minutes at 95°C in 2X Laemmli sample buffer, followed by incubation for 1 hour at 65°C and further incubation for 5 minutes at 95°C can reverse the formaldehyde-induced crosslinks, liberating the purified RNAs from associated proteins. Again, half of the immunoprecipitated sample is used for protein analysis, and the other half is used for RNA analysis as described (section 2.1.2).

3. Results

3.1 IP of RNP via a protein component

For Figure 2, HeLa cells were transiently transfected with the plasmid pcDNA3_STAU1-HA₃. STAU1-HA₃-containing RNP was retrieved via the HA epitope using anti-HA agarose. IP using rat (r)IgG served as a negative control, i.e. a control for anti-HA IP specificity. Anti-HA is shown to immunoprecipitate STAU1-HA₃ but not the irrelevant protein Calnexin. lncRNAs that are involved in STAU1-binding site (SBS) formation, such as ½-sbsRNA2, ½-sbsRNA3 and ½-sbsRNA4 [6], were co-immunoprecipitated with STAU1-HA₃ and detected using RT-sqPCR.

3.2 Purification of RNP via the lncRNA component

For Figure 3, HeLa cells were transiently transfected with pre-optimized amounts of the pFLAG-MS2 and pcDNA3_½-sbsRNA2_MS2bs plasmids, or as negative controls, either pcDNA3_½-sbsRNA2, which lacks any MS2bs, or pcDNA3_FLUC_MS2bs, which produces FLUC mRNA containing twelve copies of the MS2bs. Two days later, cells were crosslinked using formaldehyde and harvested. Cell lysates were immunoprecipitated using anti-FLAG, or as a negative control, mouse (m)IgG. As expected, prior to IP, expression of both ½-sbsRNA2 and ½-sbsRNA2_MS2bs decreased the abundance of CDCP1 mRNA but not BAG5 mRNA, showing that fusion of twelve copies of MS2bs to the 3' end of ½-sbsRNA2 does not perturb its physiological role in the STAU1-mediated mRNA decay (SMD) of CDCP1 mRNA [6]. Retrieval of the FLAG-MS2 protein via anti-FLAG-agarose was nearly quantitative. ½-sbsRNA2_MS2bs and FLUC_MS2bs were co-immunoprecipitated using FLAG-MS2 but not with ½-sbsRNA2, which lacked any MS2bs. The target of ½-sbsRNA2, CDCP1 mRNA, as well as the key SMD proteins STAU1 and UPF1 were detected only in the ½-sbsRNA2_MS2bs IP. From these and other studies [6], we conclude that the Alu element within ½-sbsRNA2 base-pairs with a partially complementary Alu element within CDCP1 mRNA to create an SBS. In contrast, irrelevant proteins, such as Calnexin, the dsRNA-binding protein ILF3, the single-stranded RNA-binding protein FMR1, and mRNAs that are not predicted to base-pair with ½-sbsRNA2, such as SMG7 mRNA, and an mRNA that is not targeted by ½-sbsRNA2, such as BAG5 mRNA, did not co-immunoprecipitate with ½-sbsRNA2_MS2bs.

4. Discussion

Here we provide two complementary examples of how protein-lncRNA complexes can be biochemically characterized. Both approaches have inherent strengths and weaknesses. The first identifies lncRNAs in protein immunoprecipitates. Provided that an antibody to the cellular protein of interest exists, interactions that take place in a physiologically relevant and unperturbed cellular setting can be examined. This method may be challenging when studying low-abundance lncRNAs or when antibodies of dubious quality and/or specificity are used. Notably, low-affinity or transient protein-RNA interactions may be captured by formaldehyde crosslinking (see section 2.2.3), partially overcoming issues associated with transient or nonspecific interactions. However, the use of formaldehyde reduces IP efficiency, necessitating the analysis of more cell lysate.

The second approach involves the co-expression of a lncRNA_MS2bs fusion and FLAG-MS2 protein. RNP containing the lncRNA_MS2bs fusions are then immunoprecipitated using an anti-FLAG antibody. Although this approach requires careful consideration of the levels at which the lncRNA_MS2bs fusion and FLAG-MS2 protein are expressed, the IP efficiency of anti-FLAG is consistently excellent (i.e. nearly quantitative).

Additional methods for lncRNA purification not presented could involve generating and expressing in cells RNA aptamers that bind specifically to streptavidin matrices and have been fused to a lncRNA of interest [11]. While the length of an RNA aptamer is generally much shorter than the 12 copies of the MS2bs employed in the second method, folding of the MS2bs is more likely to be independent of the lncRNA sequence and thus probably more likely to be accessible to binding by the molecule essential for RNP purification. An alternative method using a set of biotin-labeled DNA oligonucleotides that hybridize to and tile across the lncRNA to be purified avoids the artifacts that may be introduced by exogenous expression of fusion constructs [12]. Although costly, the DNA-oligo tiling method, which purifies the endogenous lncRNA, does not require knowledge of the complete lncRNA sequence.

With the application of the biochemical tools presented here, our understanding of the function of cellular lncRNAs will undoubtedly continue to grow in the future.

Acknowledgments

Research on STAU1 and SMD in the Maquat lab is supported by NIH R01 GM074593. C.G. is partially supported by a Messersmith Graduate Student Fellowship. M.W.P. is a HHMI Fellow of the Damon Runyon Cancer Research Foundation, DRG-2119-12.

Abbreviations

lncRNA	long non-coding RNA
lincRNA	long intergenic non-coding RNA
mRNA	messenger RNA
NSCLC	non-small-cell lung cancer
RT-PCR	reverse-transcription coupled to the polymerase chain reaction
EDTA	Ethylenediaminetetraacetic acid
PMSF	Phenylmethanesulfonylfluoride
IP	Immunoprecipitation

RNA-seq RNA sequencing (high throughput cDNA sequencing)

References

1. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, et al. Science. 2005; 309:1559–1563. [PubMed: 16141072]
2. Wang KC, Chang HY. Mol Cell. 2011; 43:904–914. [PubMed: 21925379]
3. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, et al. Nature. 2010; 464:1071–1076. [PubMed: 20393566]
4. Bond AM, Vangompel MJ, Sametsky EA, Clark MF, Savage JC, Disterhoft JF, Kohtz JD. Nat Neurosci. 2009; 12:1020–1027. [PubMed: 19620975]
5. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, et al. Mol Cell. 2010; 39:925–938. [PubMed: 20797886]
6. Gong C, Maquat LE. Nature. 2011; 470:284–288. [PubMed: 21307942]
7. Wapinski O, Chang HY. Trends Cell Biol. 2011; 21:354–361. [PubMed: 21550244]
8. Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, Tidow N, et al. Oncogene. 2003; 22:8031–8041. [PubMed: 12970751]
9. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, et al. Nature. 2009; 458:223–227. [PubMed: 19182780]
10. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, et al. Nat Biotechnol. 2010; 28:503–510. [PubMed: 20436462]
11. Srisawat C, Engelke DR. RNA. 2001; 7:632–641. [PubMed: 11345441]
12. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Mol Cell. 2011; 44:667–678. [PubMed: 21963238]

Appendix: Equipment and Supply List

*supplier is in parenthesis

Tris base (American Bioanalytical)

NaCl (American Bioanalytical)

Ethylenediaminetetraacetic acid (EDTA, Sigma-Aldrich)

Triton X-100 (Sigma-Aldrich)

RNase OUT (Invitrogen)

Phenylmethanesulfonylfluoride (PMSF) (Sigma-Aldrich)

Benzamidinium HCl (Sigma-Aldrich)

Mini protease inhibitor tablet cocktail –EDTA (Roche)

Superscript III Reverse Transcriptase (Invitrogen)

Protein-A agarose (Roche)

Protein-G agarose (Roche)

Glycogen (Fermentas)

RQ1 DNase (Promega)

Sodium acetate (Sigma-Aldrich)

Ethanol, absolute (Fisher)

TRIzol for nucleic acid extraction (Invitrogen)

[α -32P] dATP (Perkin Elmer)

GoTaq Green 5X PCR buffer (Promega)

Taq polymerase (Invitrogen)

Anti-FLAG sepharose (Sigma)

Formaldehyde (Sigma-Aldrich)

Branson probe sonicator (Branson)

Highlights

- Immunoprecipitation of lncRNA-containing protein complexes via either a protein component or a lncRNA component
- Insertion of MS2 coat protein-binding sites into lncRNAs to provide an RNA-based affinity-handle
- Analysis of lncRNAs by RT coupled to either semi-quantitative or quantitative PCR

Figure 1.

Schematic representation of the two RNP-purification techniques described in this communication. (Left) Protein “X” (red) with any associated lncRNAs, proteins, or other RNAs, is immunoprecipitated from a heterogeneous cellular lysate that contains irrelevant proteins and nucleic acids (gray) using an anti-Protein X antibody, provided such an antibody exists. Should there be no anti-Protein X antibody, then Protein X can be expressed as an epitope-tagged Protein X; however, it will not be possible to define conditions whereby epitope-tagged Protein X is overexpressed relative to cellular Protein X until an anti-Protein X antibody becomes available. Note that it is unlikely that all of cellular (or epitope-tagged) Protein X is bound to a particular lncRNA, leading to the IP of free Protein

X (or epitope-tagged Protein X) and, possibly, multiple RNA species. (Right) IP of lncRNA-containing RNP (red) via the lncRNA proceeds by fusion of 12 copies of a MS2 coat-protein binding site (MS2bs; all 12 are diagramed as a single hairpin) to the lncRNA. A co-expressed FLAG-MS2 protein (blue) binds each MS2bs tightly and is used to immunoprecipitate the tagged lncRNA as well as any associated RNAs and proteins. Both techniques can use downstream Western blotting and RT-PCR analyses. Alternatively or additionally, associated proteins and RNAs can be defined using, respectively, mass spectrometry and deep sequencing. See text for details.

\$watermark-text

\$watermark-text

\$watermark-text

A **B** **Figure 2.**

Immunoprecipitation of STAU1-HA₃ RNP. (A) Lysates of HeLa cells transiently expressing STAU1⁵⁵-HA₃ protein at a level comparable to the level of cellular STAU1⁵⁵ were subjected to immunoprecipitation (IP) as described. Western blotting confirmed that STAU1-HA₃ was expressed at a similar level to endogenous STAU1, and that the anti-HA affinity resin enriched STAU1-HA₃ but not irrelevant proteins such as Calnexin. Three-fold protein dilutions confirmed that data fell within the linear range required for semi-quantitative analyses. (B) RT-sqPCR analysis, performed as described in the text, showed the presence of 1/2-sbsRNA2, 1/2-sbsRNA3, and 1/2-sbsRNA4 in the STAU1-HA₃ immunoprecipitate and the absence of irrelevant RNAs, such as SMG7 mRNA.

Figure 3.

Purification of $\frac{1}{2}$ -sbsRNA2 using the FLAG-tagged bacteriophage MS2 protein. (A) Schematic representation of the plasmid DNAs used to transiently transfect HeLa cells. See text for details. (B) Western blotting (upper 5 panels) confirmed the presence of endogenous UPF1 and STAU1 in only the sample purified using anti-FLAG from cells expressing $\frac{1}{2}$ -sbsRNA2-MS2bs. Irrelevant proteins ILF3 and Calnexin were absent from immunoprecipitated samples, illustrating IP specificity. RT-sqPCR (bottom 5 panels) demonstrating the presence of $\frac{1}{2}$ -sbsRNA2; CDCP1 mRNA, which is a target of $\frac{1}{2}$ -sbsRNA2 [6]; and the absence of irrelevant SMG7 and FLUC mRNAs. Two-fold cDNA dilutions confirmed that data fell within the linear range of analysis.