Quantitative Profiling of \textit{In Vivo}-assembled RNA-Protein Complexes Using a Novel Integrated Proteomic Approach*\[5

Becky Pinjou Tsai‡, Xiaorong Wang§, Lan Huang§¶, and Marian L. Waterman‡¶

Identification of proteins in RNA-protein complexes is an important step toward understanding regulation of RNA-based processes. Because of the lack of appropriate methodologies, many studies have relied on the creation of \textit{in vitro} assembled RNA-protein complexes using synthetic RNA and cell extracts. Such complexes may not represent authentic RNPs as they exist in living cells as synthetic RNA may not fold properly and nonspecific RNA-protein interactions can form during cell lysis and purification processes. To circumvent limitations in current approaches, we have developed a novel integrated strategy namely \textit{MS2 in vivo biotin} tagged RNA affinity purification (MS2-BioTRAP) to capture bona fide \textit{in vivo}-assembled RNA-protein complexes. In this method, HB-tagged bacteriophage MS2 and stem-loop tagged target or control RNAs are co-expressed in cells. The tight association between MS2 and the RNA stem-loop tags allows efficient HB-tag based affinity purification of authentic RNA-protein complexes. Proteins associated with target RNAs are subsequently identified and quantified using SILAC-based quantitative mass spectrometry. Here the 1.2 kb internal ribosome entry site (IRES) from lymphoid enhancer factor-1 mRNA has been used as a proof-of-principle target RNA. An IRES target was chosen because of its importance in protein translation and our limited knowledge of proteins associated with IRES function. With a conventionally translated target RNA as control, 36 IRES binding proteins have been quantitatively identified including known IRES binding factors, novel interacting proteins, translation initiation factors (eIF4A-1, eIF-2A, and eIF3g), and ribosomal subunits with known noncanonical actions (RPS19, RPS7, and RPL26). Validation studies with the small molecule eIF4A-1 inhibitor Hippuristanol shows that translation of endogenous lymphoid enhancer factor-1 mRNA is especially sensitive to eIF4A-1 activity. Our work demonstrates that \textit{MS2 in vivo biotin} tagged RNA affinity purification is an effective and versatile approach that is generally applicable for other RNA-protein complexes. \textit{Molecular & Cellular Proteomics} 10: 10.1074/mcp.M110.007385, 1–15, 2011.

RNA-protein complexes play central roles in post-transcriptional regulation, but their dynamic nature can make it challenging to identify protein components and define steps in RNA processes. This is especially true for internal ribosome entry site (IRES)\[1 elements. IRESs are long stretches of non-coding RNA in the 5’ untranslated (UTR) regions of a subset of cellular and viral mRNAs (1–4). Unlike most mRNAs which have short 5’UTRs, the highly structured regions of the \textit{LEF1} IRES and other lengthy IRESs impede ribosome scanning (1, 3, 4). Ribosome recruitment and translation of mRNAs with short 5’UTRs rely on the canonical cap-binding complex (eIF4F) which recognizes a 7-methylguanosine cap structure on the 5’ of eukaryotic mRNAs (1, 2, 4). In the case of viral mRNAs, which often do not have a 7-methylguanosine cap, IRESs capture ribosomes for translation via IRES trans-acting factors (ITAFs) or specific RNA secondary structures (5). In the case of IRESs in eukaryotic, cellular mRNAs, the mechanism for recruitment of ribosomes can function either as an alternative to canonical cap-dependent recruitment mechanisms when translation is compromised (e.g. nutrient deprivation, hypoxia, and mitosis), or as a cap-enhancing mechanism to increase translation (1, 3, 6).

IRESs were originally discovered in RNA viruses such as picornaviruses. Viral IRESs are several hundred nucleotides in length, their sequence is highly conserved, and they form tightly folded RNA scaffolds for ITAF assembly and ribosome interactions (5). Many of these RNA structures and interacting ITAFs are well studied. Unlike viral IRESs, our understanding of cellular IRESs is limited. Cellular IRESs are estimated to be present in 3–5% of capped mRNA transcripts, they are highly

---

\[1\] The abbreviations used are: IRES, internal ribosome entry site; HB-tag, Histidine and biotin tag; MS2-HB, bacteriophage MS2 coat protein dimer fused to a HTBH-tag; LEF1, Lymphoid Enhancer Factor-1; MS2-BioTRAP, MS2 \textit{in vivo} Biotin Tagged RNA Affinity Purification; SILAC, Stable isotope labeling with amino acids in cell culture; ITAF, IRES trans-acting factors; Cap, Indicates canonical cap-dependent translation; MAP, mix after purification; PAM-SILAC, purification after mixing.
variable in length (up to several kilobases) and they do not exhibit the same degree of sequence conservation compared with viral IRESs (4). These features make the identification of structures and components of cellular IRES-protein complexes extremely challenging. Conventional strategies of IRES-protein analysis involve in vitro approaches with synthetic RNA, aptamer tags, and in vitro purification. These approaches have identified a handful of IRES trans-acting factors (ITAFs), which are overwhelmingly represented by hnRNPs and other abundant RNA binding proteins (1, 4, 7). IRES-protein complexes are dynamic structures and their compositions are subject to change depending on many factors including RNA folding, subcellular localization, and transcription/post-transcription processes that modify the RNP/mRNA complex as it moves from the nucleus through nuclear pores to sites of translation in the cytoplasm (1, 8, 9). Thus, although they are useful, in vitro purification strategies are not best suited for the capture of in vivo assembled IRES-protein complexes.

In order to preserve authentic RNA-protein complexes as they are isolated from living cells, several new methods have recently been developed (10–14). These strategies are protein-centric in that a specific RNA binding protein is tagged, expressed in vivo, and used as bait to capture its interacting RNAs for subsequent microarray analysis or deep sequencing. These strategies require a known RNA binding protein and do not allow effective identification of other components in RNA-protein complexes at a proteome scale. To circumvent this problem, we have developed an integrated proteomic strategy that is RNA-centric and uses MS2 in vivo Biotin Tagged RNA Affinity Purification (MS2-BioTRAP) and stable isotope labeling with amino acid in cell culture (SILAC)-based quantitative mass spectrometry. In this strategy, a specific RNA is tagged with a cluster of RNA stem-loops recognized by bacteriophage protein MS2, an RNA binding protein that binds to the single-stranded loop region with high nanomolar affinity (15–17). MS2 is HB-tagged and co-expressed for in vivo association with the stem-loop tagged RNA (18, 19). The HB tag consists of two hexahistidine tags, a TEV cleavage site, and a signal sequence for in vivo biotinylation (19). This enables rapid and effective one-step purification of MS2-HB, its associated stem-loop tagged RNA, and all other proteins bound to the tagged RNA. To maintain the integrity of protein-RNA complexes during the purification processes, in vivo UV cross-linking is carried out prior to cell lysis to freeze RNA-protein interactions in living cells. SILAC-based quantitative mass spectrometry is subsequently employed to quantitatively identify proteins associating with specific IRES RNAs in comparison with a non-IRES RNA (e.g. Cap) control sample. The results have been further validated by co-immunoprecipitation, quantitative Western blot, and siRNA knock-down experiments to demonstrate that MS2-BioTRAP captures bona fide interactors that regulate thelef1 IRES. The work presented here describes a general proteomic strategy that is valuable for studying in vivo RNA-protein complexes as they occur in living cells.

**EXPERIMENTAL PROCEDURES**

**Plasmid Construction for tagged-IREs and tagged-Cap—**Dicstronic reporter plasmids pRstF and pRstF-5'UTR (20) were used to generate tagged-Cap and tagged-IRES expression constructs, respectively. To generate a monocistronic reporter plasmid, the Nhel and EcoRI sites were used to remove the upstream Renilla luciferase open reading frame and bisect and destroy the subsequent stem-loop. The circular plasmid was regenerated by blunt end ligation. The monocistronic reporters were then linearized (XbaI site) between the Firefly luciferase stop codon and poly(A) signal sequence and a MS2 stem-loop fragment containing four tandem stem-loops was inserted by blunt end ligation (MS2 stem-loop template, SP73-β-globin-(MS2)4, was a gift from Klemens Herpel).

**Plasmid Construction for MS2-HB—**The MS2 coat protein sequence was amplified from pCT119-N55K (gift from David Peabody, University of New Mexico) using a three-piece ligation strategy. To generate a tandem-linked dimer of open reading frames, the first MS2 coat protein in the dimer was generated by PCR amplification of the MS2 coat protein plasmid sequence using a sense primer (5'-AACCTGAAGGCGCCGTAGCTTTTACCACTCCA-3') containing a NotI site (italicized) upstream of MS2 sequence and an antisense primer containing a BglII site downstream of the MS2 sequence (5'-ATTAGCGGCGGTAGCTTTTACCACTCCA-3'). The second MS2 coat protein was amplified using a sense primer (5'-CTCAAGACCTGGCAGATGGGATTTGACAGA-3') containing a BglII site upstream of the MS2 sequence and an antiserine primer (5'-CCATTATTAAAGGAGTTTGGTTCAGATT-3') containing a PacI site downstream of the MS2 sequence. BglII restriction enzyme sites at the 3' and 5' end of each PCR product allowed for blunt end ligation and creation of a tail-to-head tandem placement of two MS2 coat protein open reading frames. The two PCR products were digested with the indicated restriction enzymes and cloned into the HBTH (a derivative of HB tag) tag vector (pQCXIP backbone (18, 19)) linearized with NotI and PacI.

**Establishment of Stable Cell Lines (293MS2-HB)—**Two hundred and ninety-three stable cell lines expressing MS2-HB (293MS2-HB) were generated by retrovirus infection as described (19). Briefly, 293GFP packaging cells were plated at a density of 7 × 10^5 cells/10 cm-diameter tissue culture dish and transfected with pQCXIP-MS2-HB retroviral vector using a calcium phosphate protocol. Following 8 h, the medium was replaced with fresh Dulbecco’s modified Eagle’s medium (DMEM). Thirty-six hours post-transfection, the conditioned medium containing the retroviruses was collected every 8 h for 60 h. Two hundred and ninety-three cells were infected by incubation with equal volumes of fresh DMEM and retrovirus conditioned medium and 4 μg/ml of polybrene. Following 4–6 h, cells were washed and a second infection was performed. Thirty hours postinfection, cells were seeded at 2 × 10^5 cells/10-cm plate and cultured in DMEM containing the selection antibiotic puromycin (3 μg/ml). Following –5 days of selection, cells were seeded at a density of 2 × 10^4 to 2 × 10^5 cells per 10 cm-diameter tissue culture dish. Individual clones were picked from the plates and expanded to generate stable cell clones expressing MS2-HB.

**Cell Culture and Metabolic Stable Isotope Labeling Using SILAC (21)—**The stable cell line expressing MS2-HB, 293MS2-HB, was grown in SILAC DMEM (Thermo Scientific, #89985) supplemented with 28 μg/ml 13C6-lysine, 73 μg/ml 13C6-arginine, 12C6-lysine (Sigma) (light medium) or 13C6-lysine and 15N2-lysine (heavy medium) purchased from Cambridge Isotope Laboratories (Andover, MA), 10% fetal bovine serum, 1% penicillin/streptomycin, 3 μg/ml puromycin (stable cell selection), and 5 μM biotin (Sigma). Cell lines were grown...
Quantitative Profiling of In Vivo-assembled RNP Complexes

for more than seven cell doublings in the labeling media to ensure complete incorporation. The cells were then grown to confluence prior to cell lysis.

Cell Transfection, Harvest, and Lysis for Affinity Purification—Two hundred and ninety-three (MS2-HB) cells were cultured in 150-mm plates and transfected with the respective tagged-RNAs as described (21–23). Forty-eight hours post-transfection, cells were washed with ice-cold phosphate-buffered saline (PBS), cultures were then immersed in 7 ml 1× PBS, and UV cross-linked using 2400 Stratalinker to irradiate one time for 400 mJ/cm². Cells were harvested by scraping and collected by centrifugation. Cell pellets were lysed using either the native lysis buffer L (100 mM NaCl, 50 mM Tris, 5 mM MgCl₂, 10% glycerol, 0.5% NP-50, RNAsIN (Optzyme Ribonuclease Inhibitor, Fisher, 5000×), 1 mM PMSF, Protease Inhibitor (Sigma, 1000×), 50 mM NaF, 0.1 mM Na₄VO₄, 5 mM EDTA, 5 mM EGTA, 0.5 mM β-mercaptoethanol) or denaturing lysis buffer A-8 (6 mM Urea, 300 mM NaCl, 50 mM NaH₂PO₄/NaHPO₄, 0.5% Nonidet P-40) (lysis volumes were chosen to achieve 15–20 mg/ml total protein). Lysate was sonicated using the Branson Sonifier 450 (Setting Duty 50%, Output 4) three times for 15-s intervals. Lysate was centrifuged to remove large debris and clarified by filtration through 1.6 µm GD/X Glass Microfilter filters (Whatman). All steps were performed on ice, with ice-cold reagents.

Affinity Purification—The purification procedure is similar to a previously reported procedure (19, 22). Briefly, Dynal streptavidin M-280 magnetic beads (Invitrogen) were prepared by washing three times with the same buffer used for cell lysis. Amount of beads used varied with amount of total protein in the lysate (~1 µl bead slurry/20 µg lystate). For native purification, we employed the MAP-SILAC method (mix after purification) to isolate RNA-protein complexes from light and heavy labeled lysates separately to prevent interaction exchange during purification (21). Each lysate was applied to beads in a 1.5-ml tube and rotated at 4 °C for 5 min. An ice-cold magnet was applied to the side of the tube to isolate the beads along with the captured RNP complexes. The flow-through lystate was removed and the beads were washed on ice 2× with ~50× bed volume of the same buffer used for cell lysis.

For on-bead trypsin digestion, beads were washed again in ice-cold 25 mM NH₄CO₃ (30× bed volume) and resuspended with 5–10 ng/µl Promega (Madison, WI) trypsin (diluted in 25 mM NH₄CO₃). Digestion of the sample proceeded at 37 °C for 8–12 h. For denaturing conditions, equal amounts of lysate were mixed from the light and heavy conditions, followed by addition of the magnetic beads (PA-MAP-APC, purification after mixing) (21). The steps thereafter were the same as in native purification. Optional TEV digestion and elution: For native purification, TEV (Tobacco Etch Virus Protease; Invitrogen) digestion can be used to release MS2-HB and captured RNA/protein complexes. Following the wash step with lysis buffer, complexes were washed again with 30× bed volume of ice-cold TEB buffer (50 mM Tris-HCl pH 8.0, 10% glycerol, 0.5 mM EDTA, 1% Triton X-100, 1 mM DTT), then incubated with two bed volumes of TEB (0.25 U/µl, diluted with TEB) for 1 h at room temperature. A magnet was applied to collect the elution and followed with trypsin digestion.

Mass Spectrometric Analysis-Liquid Chromatography (LC) MS/MS—Tryptic digests were first separated by strong cation exchange chromatography using a polysulfobethyl A column (2.1 mm i.d. × 10 cm long) (Nest Group) at a flow rate of 200 µl/min using AKTA Basic 10 (GE Healthcare) (21). Peptide elution was achieved by a salt gradient (0–350 mM KCl), and fractions were manually collected based on UV absorbance at 215 nm. The collected fractions were desalted using Vivapure C18 microspin columns (Vivascience, Auburn, MA) and analyzed by a linear ion trap (LTQ) Orbitrap XL mass spectrometer (Thermo-Electron Corp) (24). The LC analysis was performed using a capillary column (100 µm i.d. × 150 mm long) packed with Inertsil ODS-3 resin (GL Sciences, CA); the peptides were eluted using a linear gradient of 2% to 35% B in 85 min at a flow rate of 350 µl/min (solvent A, 100% H₂O-0.1% formic acid; solvent B, 100% acetonitrile-0.1% formic acid). A cycle of one full Fourier transform scan mass spectrum (350–1800 m/z, resolution of 60,000 at m/z 400) followed by 10 data-dependent MS/MS acquired in the linear ion trap with normalized collision energy (setting of 5%). Target ions already selected for MS/MS were dynamically excluded for 30 s.

Database Searching for Protein Identification and Quantification—The MS data was extracted and analyzed as described (24). Monoisotopic masses of parent ions and corresponding fragment ions, parent ion charge states, and ion intensities from LC-MS/MS spectra were extracted using in-house software based on Raw_Extract script from Xcalibur v2.4. Following automated data extraction, the resultant peak lists for each LC-MS/MS experiment were submitted to the development version (5.3.0) of Protein Prospector (UCSF) for database searching using a concatenated Swissprot database (837302 sequence entries) composed of a SwissProt database (3/24/2009) and its randomized version. Homo sapiens was selected as the restricted species. Trypsin was set as the enzyme with a maximum of two missed cleavage sites. The mass tolerances for parent and fragment ions were set as 20 ppm and 0.8 Da respectively. Chemical modifications such as protein N-terminal acetylation, methionine oxidation, N-terminal pyroglutaminomation, and deamidation of asparagine were selected as variable modifications. For SILAC experiments, 13C₆-15N₂-Arg and 13C₆-15N₂-Lys were also chosen as variable modifications. The Search Compare program in Protein Prospector was used for summarization, validation, and comparison of results. A false positive (% FP) rate of ≤0.2% was used for peptide identification calculated in Search Compare (24). At this false positive rate, false protein hits from the decoy database were not observed. General protein identification was based on at least two peptides. If a protein was identified by multiple peptides from more than two preparations, it was considered as a hit in a preparation identified by only one peptide. For SILAC experiments, the Search Compare program was also used to calculate the relative abundance ratios of Arg/Lys-containing peptides based on ion intensities of monoisotopic peaks observed in the LC MS spectra. The SILAC ratios were further validated by checking all of the raw spectra within the Search Compare program. The ratio outliers (with >30% standard deviations) were excluded for further calculation. If the peptide peaks were mixed with other peptide peaks, they were excluded for quantification. The SILAC ratios reported here were average values, and the accuracy and significance of the measurements were evaluated using standard deviations. The ratios obtained from replicate experiments were examined manually. The resulting SILAC ratios were normalized to the MS2-HB protein in each sample.

Western Blot Analysis—Cell lysates and purified samples were resolved by 10% SDS-PAGE, transferred to nitrocellulose membrane, and analyzed by immunoblotting. Blots were probed (1:1000) with antibodies purchased from Thermo Fisher Scientific (Streptavidin-HRP), Cell Signaling (#Ab414A2; #2013; #2924; #9722, LERF-1, #4777; Lamin A/C, #2286) (Beverly, MA) and Abcam (PSF, #ab38148; hNRP H/F, #ab10689; FUS, #ab23439; NonO, #ab45359; YB-1, #12148) (Cambridge, MA) and Zymed Laboratories Inc. (β-tubulin, #32–2600) (San Francisco, CA). PDB2P2 antisera was a gift from Bert Semler.

Transient Transfections and Reporter Assays—For luciferase reporter assays 2×10⁵ cells were seeded 18–24 h prior to transfection. FuGENE (Roche) transfection reagent was added according to the manufacturer’s protocol to transfect 500 ng of the respective Firefly luciferase reporter vector and 100 ng of a control CMV β-galactosidase reporter vector. Cells were washed with 1× PBS-EDTA and lysed
with 1× passive lysis buffer (Promega) 24 h post DNA transfection. Cell lysates were assayed for luciferase activities with luciferin substrate (Sigma) using a SIRIUS luminometer (Berthold Detection Systems, Oak Ridge, TN). β-galactosidase activities were determined using the Galacton-Plus substrate (Applied Biosystems, Foster City, CA); these activities served as an internal control for transfection efficiency. Firefly luciferase/β-galactosidase ratios were calculated to determine normalized luciferase values for each sample. Lysate for dicistronic reporter assays were prepared similarly and analyzed using the Dual Luciferase Reporter Assay System (Promega, Seattle, WA) according to the manufacturer’s protocol.

**siRNA Knock-down**—Monolayers of cells were seeded (2 × 10^5 cells) 24 h prior to transfection of siRNA (Qiagen, Valencia, CA, USA) using the Qiagen HiPerfect Transfection reagent. Two siRNA transfections (3 nM siRNA) were performed over a 48 h period. Following 48 h, fresh media was added and cells were transfected with dicistronic reporter vectors using FuGene 6 transfection reagent according to the manufacturer’s protocol. Cells were harvested for Western blot and luciferase activity analysis.

**RNA Isolation, Northern Blot Analysis**—Total RNA was isolated from cells and streptavidin bead purification using TriZol reagent (Invitrogen) according to the manufacturer’s instructions. RNA was denatured in Northern Max Formaldehyde Loading Dye (Ambion, Austin, TX) at 65 °C for 5 min and resolved on a 10.1074/mcp.M110.007385–4 4% agarose gel at 5 V/cm (100V) for 2 h. RNA was transferred to a Brightstar-Plus (Ambion) membrane by vacuum transfer (Bio-Rad Vacuum Blotter) with 10× SSC/10 nM NaOH for 2 h. Following transfer, membranes were briefly rinsed with ddH2O, UV-irradiated, stained with 0.04% methylene blue dye (in 0.5 m NaOAc) for 1 min, and rinsed with ddH2O to visualize the ribosomal 28S and 18S RNAs. Membranes were then prehybridized at 60 °C in ExpressHyb Hybridization Solution (BD Biosciences) with 0.2 mg/ml salmon sperm DNA (Stratagene, La Jolla, CA) for 2 h in a hybridization chamber. The membranes were hybridized in hybridization solution with 20 μg/ml yeast tRNA (Invitrogen) and random prime [32P]-labeled probes (3 × 10^6 cpm/ml) against the Firefly luciferase (Fluc) coding region sequence. After 24 h, membranes were washed twice with SSC/0.1% SDS, 60 °C and subsequently washed twice in SSC/0.5% SDS at 65 °C. Bands were visualized using a phosphorimager and Quantity One version 4.3.0 software (Bio-Rad).

**Semi-quantitative RT-PCR**—First-strand cDNA synthesis was performed using oligo(dT) primers and iScript cDNA synthesis kit as directed by the manufacturer (Bio-Rad). Quantitation of target transcripts was performed by real-time PCR using SYBR Green qPCR master mix (SA Biosciences) with an ABI Prism HT7000 cycler (Applied Biosystems). Primer sets were as follows: Firefly luciferase sense 5′-ACGCCAGGTTCGCAAGGTTTTC-3′; antisense 5′-TTCCGACTTCCGGCCCAAAACACACT-3′ and internal control UBA52 sense 5′-AGACAAGGAGGTATCC-3′; antisense 5′-GAAGGGCAAGGATGC-3′.

**Immunofluorescence**—To visualize MS2 coat protein localization 24 h post-transfection of the respective reporter plasmids, 293HS2-LEF1 cells were cross-linked with 3.7% formaldehyde, blocked with 5% normal goat serum (Vector Labs, Burlingame, CA) for 1 h at room temperature and hybridized with α-MS2 (1:500; Tetracore) at 4 °C overnight. Cells were washed with 1× PBS and incubated with phalloidin (F-actin staining) and secondary goat α-rabbit antibody conjugated to Alexa Fluor 594 (1:1000; Molecular Probes/Invitrogen) for 1 h at room temperature. Subsequently, cells were washed, stained with Dapi (nucleus), and mounted on glass slides. Images were captured using a Zeiss LSM 510 Meta confocal microscope and analyzed with LSM5 Image Browser (Zeiss).

**PANTHER and STRING Analysis**—PANTHER (Protein analysis through evolutionary relationships) Classification System is a database of annotated gene and gene functions based on their evolutionary relationships. This database was used to determine the types and general function of the proteins purified (see Fig. 3A). Gene IDs were submitted to the database and sorted by GO molecular function. Minor modifications were made to the PANTHER categorization to combine all nucleic acid associated protein category into a single nucleic acid binding category. The STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database displays predicted protein-protein interactions based on various sources (genomic context, high-throughput experiments, co-expression, and literature). To illustrate the possible interactions between the purified factors, gene IDs for the enriched subset were submitted for analysis.

**RESULTS**

**Design and Expression of MS2-HB and Stem-loop Tagged-RNA**—To capture a comprehensive set of IRES regulatory factors, we aimed to develop a new method that integrates an RNA/Protein tagging system with quantitative mass spectrometry. Our approach utilizes the bacteriophage RNA binding protein MS2 to capture RNAs (Fig. 1A) (15, 17, 27). MS2 dimers bind specifically and with high affinity to a RNA stem-loop target sequence (11, 15, 16, 28, 29), which enables MS2 to capture target RNAs in vitro from cell extract systems including target RNA assembled in spliceosome or exon junction complexes, or even to visualize mRNA target localization in living cells (12, 15, 17, 30–34). To create an in vivo target for MS2 in living cells, we constructed a LEFT1 IRES expression plasmid in which the 1.2 kb IRES region was cloned upstream of Firefly luciferase coding sequences (Fig. 1A). To tag the mRNA produced from this plasmid, four MS2-targeted stem-loops (tagged-IRES) were cloned at the 3′ end of the luciferase open reading frame (Fig. 1A). Similarly, a stem-loop tagged-Cap construct lacking the IRES element was engineered and used to analyze canonical cap-dependent translation. In addition, the tagged-Cap served as the control RNA for quantitative identification of specific IRES interacting proteins during SILAC-based mass spectrometry analysis. Each mRNA sequence was cloned into a modified pRL-SV40 luciferase reporter plasmid with the RNA polymerase II SV40 promoter/enhancer directing transcription (35). Other features include an intron flanked by a set of splice sites placed ~200 nucleotides downstream of the transcription start site, and a polyadenylation signal downstream of the luciferase stop codon. Splicing and polyadenylation signals ensure that the expressed mRNAs follow basic steps of RNA processing, so that each transcribed mRNA should be capped at the 5′ end with a 7-methylguanosine cap, marked at the site of splicing, and finished with a polyadenylated 3′ end.

In order to effectively isolate MS2-associated RNA-protein complexes, a eukaryotic expression vector for MS2 was constructed in which the MS2 protein was tagged at its C-terminal end with a HTBH tag (19), a derivative of the HB tag (18).
The HB tag is a novel and versatile tandem affinity tag that can be used for purification of protein complexes under native or fully denaturing conditions (18, 19). The HTBH tag consists of two 6xHis clusters separated by a TEV cleavage sequence and an in vivo biotinylation site (19) (Fig. 1A). The high affinity of biotin for streptavidin (Kd ~ 10^-15 M) enables efficient isolation of endogenously biotinylated HB-tagged proteins and allows for stringent washing conditions to eliminate nonspecific background binding (18). To significantly increase the affinity and target specificity for the RNA stem-loop sequence (16, 17), an HB-tagged MS2 mutant (N55K) dimer was generated by linking two copies of the N55K open reading frame followed by the HB sequence. Others have shown that a linked dimer has enhanced stem-loop recognition and suppressed nonspecific interactions with other proteins (16, 17). As shown in Fig. 1B, co-expression of the MS2-HB dimer with the stem-loop tagged RNA was not detrimental to Firefly luciferase production by either the IRES or Cap transcripts, instead, it may have provided mRNA stability and a modest increase in expression. To minimize overexpression, a 293 stable cell line expressing MS2-HB (i.e. 293MS2-HB) was generated by retroviral infection. The localization of MS2-HB was detected in both the nucleus and cytoplasm of all conditions. D, Evaluation of affinity purification efficiency by Western blot. Protein samples collected from streptavidin-bead purification of lysates from cells transfected with tagged-IRES in the presence (+) or (-) absence of MS2-HB. Input, flowthrough, and product (“bound”) isolated on the streptavidin magnetic beads were assayed by Western and Northern analysis (see panel E). MS2-HB is detected with overbiotin antibody. Input and flowthrough lanes demonstrate isolation of >90% of MS2-HB by streptavidin bead isolation. E, Northern blot analysis shows that intact stem-loop tagged target IRES RNA (3 kb) is captured by MS2-HB coat protein (bound). F, Semi-quantitative RT-PCR analysis of tagged-IRES and tagged-Cap RNA captured by MS2-HB.
tagged-RNAs in purified samples and their capture efficiency, Northern blot analysis and semi-quantitative RT-PCR was performed. As shown in Fig. 1E, full-length tagged-IRES mRNA (~3kb) was specifically copurified with MS2-HB, whereas no IRES RNA was isolated in the absence of MS2-HB (Fig. 1E, compare “bound” lanes). In addition, semi-quantitative RT-PCR analysis indicated that our method has captured ~1% of both tagged-RNAs from whole cell extracts (Fig. 1F, and data not shown). Together, the results have demonstrated that MS2-HB is effective for rapid isolation of functional stem-loop tagged RNAs.

**MS2-BioTRAP Strategy for Quantitative Profiling of RNA-Protein Complexes**—In order to isolate and identify specific RNA-protein complexes, we have developed an integrated proteomic approach coupling MS2-BioTRAP with SILAC-based quantitative mass spectrometry (21, 36). The general workflow is displayed in Fig. 2. As shown, one population of 293MS2-HB stable cells was grown in light-labeled (12C14N-Arg/Lys) media and another population was grown in heavy-labeled (13C15N-Arg/Lys) media (Fig. 2). Each culture was propagated extensively (at least seven doublings) to permit sufficient incorporation of stable isotope labeled amino acids.

To identify IRES specific interacting proteins, the stem-loop tagged IRES target RNA was transiently expressed in the light culture, whereas the stem-loop tagged Cap control RNA was transiently expressed in the heavy culture (Fig. 2). To obtain authentic RNA-protein complexes, UV irradiation was applied to living cells to freeze RNA-protein interactions prior to cell lysis. In this work, two purification strategies were employed as illustrated in Fig. 2 (Path I and II), in an effort to generate a comprehensive in vivo RNA-protein interaction profile.

The first strategy (Fig. 2, path I) involves the isolation of MS2-HB associated RNA-protein complexes under native conditions, in which the MAP-SILAC (mix after purification) method was employed to prevent the exchange of dynamically interacting proteins between their light and heavy labeled forms as previously reported for native purification in a standard SILAC protocol (21, 37). Using this method, cells in the light and heavy cultures were lysed separately. RNA-protein complexes were subsequently purified from the tagged-IRES (light) and tagged-Cap (heavy) containing lysates by binding to streptavidin magnetic beads. The resulting purified samples were then mixed equally for trypsin digestion on the beads. This strategy presumably allows the capture of
cross-linked RNA-protein as well as noncross-linked protein-protein interactions; however, nonspecific background might be high. Alternatively, the bound proteins can be directly eluted with TEV cleavage (see Experimental Procedures).

In comparison, the second strategy (Fig. 2, path II) employs the purification of MS2-HB associated RNA-protein complexes under fully denaturing conditions (e.g. 8 M urea), permitting the capture of only cross-linked RNA-protein interactions. With this strategy, the PAM-SILAC (purification after mixing) method was used because it is best suited for isolating cross-linked complexes under denaturing conditions (37). In this strategy, equal amounts of cell lysates from the two compared samples were mixed first and then HB-tag based purification was carried out. Although the HB tag allows two-step purification, we only performed one-step purification by binding to streptavidin beads. Single-step isolation enabled a fair comparison of the results from both strategies. As for the PAM purification under native conditions, denatured proteins bound to streptavidin beads were directly digested with trypsin on-bead for subsequent analysis.

Trypsin digests from both strategies were analyzed by LC MS/MS for protein identification and quantification. With SILAC-based quantitative mass spectrometry, the relative abundance of a given protein in the two compared samples can be determined by the relative abundance ratios (i.e. SILAC ratios) of its arginine/lysine containing peptides (i.e. light-labeled versus heavy labeled) measured by mass spectrometry. If proteins specifically interact with the tagged-IRES RNA, relative abundance ratios (i.e. SILAC ratios) of their light labeled tryptic peptides (from the tagged-IRES sample) to the corresponding heavy labeled peptides (from the tagged-Cap sample) should be different from 1 (17). However, if proteins bind at equal levels to both tagged-RNAs, relative abundance ratios of their tryptic peptides should be close to 1. Thus a large set of RNA binding factors that bind to the tagged-IRES versus the tagged-Cap control RNA can be quantitatively determined.

Quantitative Profiling of IRES RNA Interacting Proteins—With the strategies described above, we have carried out two native and two denaturing purifications for four independent pair-wise comparisons. As a result, the SILAC ratio of MS2-HB was determined as 1.2 ± 0.2 in all four experiments, demonstrating that MS2-HB is isolated uniformly using our purification strategies. Uniform isolation of MS2-HB thus enabled normalization of SILAC ratios of proteins in each experiment to SILAC ratios of MS2-HB for parallel comparison among multiple samples. Based on a threshold previously established by immunoblotting (38), proteins with normalized SILAC ratios greater than 1.5 were determined as specific, enriched interactors of the tagged IRES RNA.

From the four experiments, a comprehensive list of proteins was identified. This list is composed of only those proteins that were consistently detected in the two replicate experiments for each purification strategy. Proteins identified by a single peptide were not included in a list unless they were identified in both replicate purifications. We reproducibly captured a total of 535 proteins from the native purification and 326 proteins from the denaturing purification (Fig. 3A; supplemental Table S1 and S2). The majority of the proteins (75%) captured under denaturing conditions were also captured under native conditions. The large overlap suggested that the more stringent denaturing condition retained relevant factors whereas removing indirectly interacting (components of complexes) or nonspecific binding factors that are isolated by the less stringent native purification. Categorization of proteins in functional categories using PANTHER classification tools defined the overall spectrum of proteins isolated in each purification condition (Fig. 3A) (26). Overall, the classification profiles were similar for native and denaturing conditions, with nucleic acid binding proteins the most prominent category, and the category with the greatest difference between native and denaturing conditions.

Further analysis of nucleic acid binding proteins revealed the greatest number of common and IRES-enriched regulatory factors (Fig. 3B). Panther analysis identified five subcategories of nucleic acid binding proteins isolated in each condition: “mRNA processing,” “other RNA binding proteins,” “translation,” “hnRNPs,” and “DNA related factors.” The diverse functions represented by these subcategories suggest that RNA was captured from multiple subcellular compartments with the target RNA undergoing different processes from transcription to translation (Fig. 3B; supplemental Table S3 and S4). Immunofluorescent detection of MS2-HB in both the nuclear and cellular space correlated with this result (Fig. 1C). To understand the interaction specificity of the identified proteins with RNAs, we have categorized them into two groups: (1) IRES-enriched proteins (SILAC ratios >1.5), indicating that these proteins either preferentially interact with the IRES, or they are relatively more abundant in the IRES-RNA complex (Fig. 3B) and (2) commonly purified proteins (SILAC ratios ≤1.5), meaning that they bind equally well to the tagged-IRES and tagged-CAP matrix (38).

Common Interacting Proteins—Common interacting factors were predominantly translation and mRNA processing subunits of complexes predicted to be equivalently associated with both target-RNAs. These include components of poly(A) binding complexes, splicing complexes, and translation complexes (Fig. 3B). Some of these complexes were isolated more intact than others, and not surprisingly, were better represented in the native condition. For example, 53 of 81 known ribosomal subunits were isolated as common interacting proteins under native conditions as was all subunits of translation elongation factors EF-1 and EF-2. In the denaturing condition, 27 ribosomal subunits, but no translation elongation factors were purified as common interactors. Likewise, at least one subunit from each of the translation initiation factors (eIF 1, 2, 3, 4, 5, and 6), core components of the core exon junction complex (eIF4A-III, WIBG), and several of its outer shell and
Fig. 3. Categorization of proteins isolated under native or denaturing conditions. A, (inset) Venn diagram of the total number of proteins isolated under denaturing and native purification conditions. PANTHER categorization of total proteins isolated under each condition. B, Subcategorization of nucleic acid binding proteins. Graph displays number of proteins isolated and characterized as those that bind to both the tagged-IRES and tagged-Cap target RNAs (Common) or enriched on the tagged-IRES (Enriched) under each purification condition. IRES-enriched proteins have SILAC ratios \( \geq 1.5 \) and common to RNA proteins have SILAC ratios \( \leq 1.5 \). C, SILAC ratio analysis of background binding proteins and IRES-enriched nucleic acid binding proteins under denaturing (top panel) and native (bottom panel) conditions. Total number of peptides are plotted against SILAC ratios to compare range and overlap of ratios between background binding factors (●) and IRES-enriched factors (△). Enrichment ratio (SILAC ratio \( \geq 1.5 \)) and average ratios for each dataset are displayed with standard deviations. Venn diagram displays overlap between proteins isolated in each dataset (inset).
auxiliary factors were isolated as common interactors in the native condition (UAP56, Aly/REF, RBM8A, RNPS1, and ACIN1). With the denaturing purification, several translation initiation subunits and ribosomal subunits were enriched with the tagged-IRES (see below). Splicing and processing factors were also purified as common interactors, including subunits and auxiliary factors for each of the snRNPs (U1, U2, U4, U5, and U6), and poly(A) binding proteins PABP1, PABP2. Other complexes expected to interact equally well with tagged-IRES and tagged-CAP mRNAs were not detected, such as the nuclear and cytoplasmic cap binding complexes and proteins associated with sites of RNA decapping, degradation, and/or storage (stress granules or P-bodies). The observation that many of the commonly associated proteins are involved in splicing and translation and not mRNA degradation suggests that MS2-HB captured pools of newly processed RNA in the nucleus as well as actively translating RNA-protein complexes in the cytoplasm.

HnRNP and DNA binding proteins from other nucleic acid binding protein categories were also identified as common interacting proteins in both native and denaturing conditions. The hnRNP category represents a group of abundant RNA binding proteins that shuttle between the nucleus and cytoplasm. Even though many hnRNPs have been characterized as IRES interacting proteins, the majority of the hnRNPs we isolated were common to both tagged-IRES and tagged-Cap RNAs (17 hnRNPs isolated in the native- and 12 hnRNPs isolated in the denaturing-conditions). Likewise, proteins in the DNA binding protein category were isolated in all conditions and were the largest subcategory of nucleic acid binding protein commonly bound to RNAs in the native purification. Much of this binding is most likely to be nonspecific and/or indirect because denaturing purification eliminated most of these interactions. Taken together, the types of common interacting proteins detected in our analysis suggest that MS2-HB captured pools of newly processed RNA in the nucleus as well as actively translating RNA-protein complexes in the cytoplasm.

IRES Interacting Proteins—Compared with the number of proteins common to both tagged-RNAs, a much smaller number of nucleic acid binding proteins were enriched on the tagged-IRES (Fig. 3B). These proteins are components of chromatin, mRNA transcription, processsing, export, and translation. Some of these proteins are multifunctional nuclear proteins found in several complexes associated with various RNA processes (SFPQ/PSF and NONO/p54nrb) or even with IRES activities (PCBP2 and ELAVL1/HuR) (1, 7, 8, 39, 40). Others are uncharacterized RNA binding proteins and helicases not previously associated with eukaryotic or viral IRESs (EWS and FUS) (41, 42). Interestingly, several proteins involved in general translation were highly enriched with the IRES template. These include translation factors (eIF4A-1, eIF-2A, and eIF3g) and ribosomal subunits (RPS7, RPS19, and RPL26) with known canonical as well as noncanonical functions (43–47). That selected subunits of the translation machinery might be specifically enriched with an IRES template has significance for how IRES and cap translation might differ (see Discussion). Under stringent denaturing conditions, 36 IRES-enriched nucleic acid binding proteins were isolated (Fig. 3B) and subjected for further analysis using STRING to generate an interaction network based on known and predicted protein interactions from multiple databases (www.string-db.org). The STRING results show that the IRES-enriched proteins are highly interactive and that the central core of this interaction network is comprised of RNA binding proteins, some of which are known ITAFs (ELAV1/HuR and SFPQ/PSF) (Fig. 4) (1, 25). In summary, use of the MS2-BioTRAP system with complex extracts from whole cells appears to provide a global snapshot of putative LEF1 IRES regulatory factors that assemble onto mRNA.

Comparison of IRES Binding Specificity by SILAC Analysis—In this work, we have used protein SILAC ratios (>1.5) as the quantitative measure to determine specific IRES interacting proteins purified under either native or denaturing conditions. During the analysis, we noticed that the SILAC ratios of the specific IRES interactors indicated a moderate level of enrichment in our purifications. This may be attributed to the nature of proteins associated with RNA as many of them are abundant proteins. In addition, it is possible that the purification methods retained sufficiently high levels of indirect or background binding proteins such that SILAC ratios were masked. For example, a protein that binds specifically to the IRES and nonspecifically to other features of the RNA or the bead matrix will be detected as a protein with a modestly enriched SILAC ratio. To assess how many enriched IRES binding proteins identified in the stringent denaturing or gentle native purification condition also engage in background binding, we performed an additional SILAC analysis. This analysis was performed in the absence of any RNA target to reveal background binding of proteins to MS2-HB protein and streptavidin beads. Under these conditions, any proteins purified were expected to bind equivalently and be detected as peptides with SILAC ratios close to 1.0. We compared the list of proteins isolated as background with the list of IRES-enriched nucleic acid binding proteins (excluding proteins identified by a single peptide; supplemental Table S6 and S7). One out of the 35 IRES-enriched proteins identified in the native purifications could be purified without the RNA target. This suggests that the majority of proteins isolated with a tagged-RNA target are specific interactors. It should be noted however, that the average SILAC ratio of background binding proteins was 1.42 ± 0.29, a range of values that comes close to the range of SILAC ratios for IRES-enriched proteins (1.84 ± 0.32). Under denaturing conditions, 16 of 36 IRES proteins were detected in the background binding experiment. Although this represents a significant degree of overlap in protein identification, the SILAC ratios of the 16 shared proteins were significantly lower in the background binding conditions.
experiment (1.06 ± 0.16) than their enriched ratios in the presence of tagged-IRES (Fig. 5C). Therefore, these 16 factors are true IRES binding candidates. These results demonstrate that although our procedure will isolate “sticky” or abundant proteins that interact nonspecifically with the bead matrix, SILAC ratios can reliably distinguish true interactions—even for abundant proteins that exhibit both modes of binding (supplemental Table S6). In fact, when we consider the overall spread of average SILAC ratio for IRES-enriched proteins versus their background counterparts, we find that the ratios are separated by three standard deviations (1.86 ± 0.50 versus 1.06 ± 0.16) (Fig. 3C). Graphical representation of the SILAC ratios in the native and denaturing purifications highlight how the denaturing condition reduces background and ensures a more reliable SILAC ratio (Fig. 3C). We conclude that either purification condition reliably detects RNA-associated factors above background, and that the SILAC ratios derived from the denaturing purification condition enable a more reliable distinction between enriched and common sets of proteins.

Validation of Selected IRES Binding Factors—For a validation test of the results from the native and denaturing isolations, eight proteins were chosen for RNA-protein complex purification and Western blot analysis (Fig. 5A). Target factors were selected based on reproducible enrichment in the denaturing purifications (SFPQ/PSF, hnRNP H, hnRNP F, and FUS), and/or for their known or potential connections to other IRESs and each other (SFPQ/PSF and NONO/p54nrb, paraspeckle factors; FUS, TEF family member; hnRNP H, hnRNP F, and PCBP2, hnRNP proteins) (39, 41, 48). We also included two translation factors identified as IRES-enriched in the native purification (eIF-2A, eIF4A-1) to test the validity of enrichment under less stringent conditions. UV cross-linking and native purification were performed for both tagged-IRES and tagged-Cap RNA and the relative levels of capture analyzed by Western blot. Preferential enrichment with the IRES target was detected for all eight factors regardless of whether they were identified as enriched in native or denaturing purifications. Interestingly, the extent of IRES enrichment detected by immunoblotting ranged from 1.5- to greater than 30-fold where eIF4A-1, an RNA heli-

Fig. 4. STRING protein network map of the 36 nucleic acid binding proteins enriched on the IRES under denaturing condition. Colored nodes indicate the individual protein identified. Lines between nodes represent direct and indirect association of proteins. The strength of association is indicated by the thickness of the line. See supplemental Table S5 for protein description.
Quantitative Profiling of In Vivo-assembled RNP Complexes

A

| Gene Name | Protein Name | IRES Cap | SILAC Ratio* |
|-----------|--------------|----------|--------------|
| eIF4A†    | Eukaryotic initiation factor 4A-I | ![Western blot](image) | 1.63 ±0.13 |
| eIF2A†    | Eukaryotic initiation factor 2 subunit 1 | ![Western blot](image) | 2.12 ±0.43 |
| SFPQ‡     | Splicing factor, proline- and glutamine-rich | ![Western blot](image) | 1.69 ±0.15 |
| hnRNPH‡   | Heterogeneous nuclear ribonucleoprotein H | ![Western blot](image) | 1.68 ±0.09 |
| hnRNPF‡   | Heterogeneous nuclear ribonucleoprotein F | ![Western blot](image) | 1.74 ±0.12 |
| FUS‡      | RNA-binding protein FUS | ![Western blot](image) | 1.59 ±0.09 |
| p54nr1ο   | Non-POU domain-containing octamer binding protein | ![Western blot](image) | 1.79 ±0.20 |
| PCBP2ο    | Poly(rC)-binding protein 2 | ![Western blot](image) | 1.75 ±0.16 |
| MS2 Coat protein dimer | ![Western blot](image) | 1.00 |

B

C

D

![Western blots](image)

Fig. 5. **A**, Validation of candidate IRES binding factors. Western blot analysis of selected proteins isolated under native purification conditions. This subset represents factors that were originally identified by SILAC-based mass spectrometry as reproducibly enriched in native condition (†), reproducibly enriched in denaturing condition (‡), or enriched in one native and one denaturing condition (°). SILAC ratios represent the average values from ≥2 individual datasets. B, SiRNA knockdown of candidate IRES-enriched factors in 293 cells. Cells were transiently transfected with FUS, SFPQ, or YB-1 siRNA (Qiagen) twice over a 48 h period. Top panel displays Western blot analysis of knockdown efficiency with β-tubulin as the loading control. Knockdown of FUS, SFPQ, or YB-1 was 30%, 50%, and 85%, respectively. The LEF1 IRES dicistronic reporter vector (bottom schematic), which reports cap-dependent translation (CAP, Renilla luciferase) and IRES-dependent translation (IRES, Firefly luciferase) was transiently transfected following 48 h of siRNA knockdown. Results are reported as the fold change of luciferase activity in the presence of target siRNA (FUS, SFPQ, YB-1) over luciferase activity of scrambled siRNA (scr). C, Inhibition of eIF4A-1 by Hippuristanol treatment. H293 cell lines are transiently transfected with tagged-RNA and CMV reporter vectors, treated with dimethylsulfoxide or 100 nM Hippuristanol, and harvested and assayed 12 h following treatment for luciferase or β-galactosidase assay.
case component of the cytoplasmic cap binding complex, showed the greatest level of enrichment (34-fold).

To test for a functional link between validated proteins and IRES activity, we interfered with ITAF function using either siRNA knockdown methods or a small molecule inhibitor (Fig. 5B). For siRNA knockdown, a dicistronic vector, which assesses cap-dependent translation of the first open reading frame (Renilla Luciferase) and IRES-dependent translation of the second open reading frame (Firefly Luciferase), was used to compare effects on both modes of translation. We targeted proteins that were enriched on the IRES as well a factor common to both tagged-IRES and tagged-Cap (YB-1, SILAC ratio <1.5). YB-1 is a known eukaryotic ITAF for a heterologous IRES in the cMYC messenger RNA (7). Reduced levels of YB-1 by siRNA (85% knockdown efficiency) did not significantly affect either cap- or IRES-dependent translation (Fig. 5B) compared with the control scrambled siRNA. Knockdown of some of the more abundant validated proteins was difficult to achieve (NONO/p54nrb, EWS), but was successful for FUS and SFPQ (Fig. 5B). Transient transcription of the dicistronic vector into cells with lower levels of FUS (30% knockdown efficiency) did not have an effect on IRES-dependent translation, but increased expression of cap-dependent Renilla Luciferase almost twofold. Knockdown efficiency of SFPQ was better, with proteins levels reduced to 50% that of the control siRNA knockdown. Under these conditions translation of both cap- and IRES-dependent open reading frames was increased, with the strongest effects observed for the cap-dependent open reading frame. SFPQ is a multifunctional protein, but its role as a paraspeckle protein might have been detected as knockdown could have relieved nuclear retention of the large, highly folded dicistronic luciferase reporter. These results show that, as opposed to a nonenriched protein, which has no effect on the dicistronic reporter mRNA (YB-1), knockdown of two of the IRES-enriched factors show specific effects on expression.

Finally, we tested for the functional relevance of the strongly associating eIF4A-1 RNA helicase (Fig. 5C). We used the small molecular inhibitor Hippuristanol, which binds specifically and directly to eIF4A-1, to inhibit RNA binding. Hippuristanol has previously been shown to interfere with the function of several viral IRESs and IRES activity in cMYC mRNA (43, 44). We used a low dose of Hippuristanol that has no effect on cap-dependent translation (100 nM) and is eightfold lower than the IC50 dose for inhibition of HeLa cell growth (43). We observed a selective, significant inhibition of the tagged-IRES reporter, reducing gene expression by as much as 50% (Fig. 5C). No significant inhibition of the tagged CAP reporter was detected (Fig. 5C). In addition, no effect was detected for expression from the cotransfected control plasmid—a cap-dependent production of β-galactosidase. Thus, the IRES is much more sensitive to inhibition of eIF4A-1 RNA binding activities than conventional cap-dependent translation mechanisms, an observation that may apply to other cellular IRESs (44). To test whether this sensitivity extends to full-length, endogenous LEF1 mRNA, we treated 293 cells with 100 nM Hippuristanol for 12 h and harvested cells for Western blot analysis of endogenous LEF-1 protein (Fig. 5D). We observed a threefold decrease in LEF-1 protein levels, suggesting that as for the tagged-IRES reporter, LEF-1 translation is highly dependent on eIF4A-1 activity (Fig 5D).

**DISCUSSION**

We present a novel integration of MS2-BioTRAP with SILAC-based quantitative mass spectrometry for a rapid, efficient, and versatile strategy to isolate and identify proteins recruited to RNA targets of interest. MS2-BioTRAP converts a method of RNA capture with MS2 protein for in vitro analysis into an in vivo system. We have optimized MS2 for high affinity binding by expressing a linked dimer of MS2 and incorporating a mutation that improves RNA binding affinity into the subnanomolar range (15–17). Other methods for capture of in vivo assembled RNA/protein complexes have used RNA aptamers embedded in the target RNA (14, 49). These methods have been successful in identifying RNA binding proteins associated with a specific target RNA, and they have the advantage that they are not constrained by the need to co-express a protein for target RNA capture. However, they are limited by their relatively low affinity for matrices, and the need for native purification conditions to retain aptamer structure (13, 14, 49). Other methods for RNA-protein complex investigation are protein-centric (CLIP and PAR-CLIP), in that the capture of RNP complexes relies on epitopes present in a specific RNA binding protein of interest and UV cross-linking to capture RNAs associated with that specific protein for subsequent purification and sequencing (11, 23). MS2-BioTRAP uses advantageous features of the CLIP approaches to develop an RNA-centric strategy.

We have demonstrated that tagging RNA with the MS2 stem-loop cluster does not hamper its normal processing and translation (Fig. 1B) and at the same time, enables its rapid and efficient capture by biotinylated MS2-HB. An added strength of MS2-BioTRAP is the ease and versatility of adapting additional techniques depending on the target of interest. First and foremost, the stable 293MS2-HB cell line can be used for any stem-loop-tagged RNA. Additional stable MS2-expressing cell lines can be created for analysis of cell specific proteins and regulation. The purification strategy can also be

---

Average luciferase and β-galactosidase activity values are averages from three independent experiments and reported with standard error bars. D, Western blot analysis of endogenous LEF-1 protein expression following 12 h of 100 nM Hippuristanol (Hipp) treatment. Membranes were probed with LEF-1 and Lamin A/C-specific antibodies. Relative density of LEF-1 signals indicated are normalized with Lamin A/C signals.

10.1074/mcp.M110.007385–12 Molecular & Cellular Proteomics 10.4
customized and enhanced by including additional steps for isolation from specific subcellular compartments or complexes (e.g. nucleus, paraspeckles, polysomes, P-bodies, etc.), or other purification steps (nickel matrices and conventional biochemical isolation strategies). Additional fractionation will reduce sample complexity and might serve to increase SILAC enrichment ratios of target-interacting proteins.

The HB tag contains histidine and biotinylation tags, as well as a TEV cleavage site that makes it versatile for both single and two-step affinity purification (18, 19). Taking into account that speed may be a critical parameter for isolation of dynamically assembled RNP complexes, our selective use of the biotin moiety of HB for single-step capture with streptavidin beads allows for a very short (5min) and efficient isolation (>95%). In comparison to other tandem affinity tags that are available, the HB tag is the only one that can be used to isolate either noncross-linked or cross-linked proteins under both native and denaturing conditions. We note however, whereas Panther categorization of IRES-enriched proteins showed very similar patterns for the native versus denaturing isolations, only four of the 36 proteins identified as enriched in the denaturing isolation were reproducibly enriched in the native purifications (Fig. 3B). This is a significant variation between the two isolations. It is possible, as others have reported, that RNA-protein complexes reorganize following cell lysis in native conditions and denaturing conditions prevent this re-assembly (48). Indeed, there was no difference in the subcategories of nucleic acid binding proteins that were IRES-enriched in the two conditions, only specific proteins within these subcategories. Also as shown in Fig. 3C, SILAC ratios generated under denaturing conditions better distinguish common and background binding proteins from true IRES interactors. Therefore, whereas purification of RNP complexes under native conditions is feasible, purification of UV-cross-linked samples under denaturing purification appear to be more attractive because of reduced background and more reliable and distinguishable SILAC ratios.

In this work, we have analyzed the proteins that bind to a complex, highly structured, 1.2 kb IRES versus a control RNA that relies on cap-dependent mechanisms of translation. Equal capture of most ribosomal subunits and translation elongation factors demonstrates that ribosome binding was equivalent between the two RNA targets, a quantitation that matches well with equivalent production of luciferase protein from both target RNAs. We identified known ITAFs and other RNA binding proteins as enriched on the IRES target RNA, but it is notable that we did not detect any highly enriched RNA binding protein specific for the LEFT1 IRES (SILAC ratio > 3). This might be because of the technical considerations listed above, but it is also possible that IRES motifs are not recognized by one or two highly specific ITAFs but rather an assortment of RNA regulatory factors with moderate specificity that collectively function to promote ribosome association and translation.

Fourteen proteins known to be IRES-interacting ITAF proteins were isolated with denaturing purifications (e.g. YB-1, PCBP2, SFPQ/PSF, NONO/p54nr, PTB, hnRNPK, ELAV1/HuR, hnRNP D, hnRNP A1, SFRS3, hnRNP C1/C2, ILF2/NF45, hnRNPU, hnRNP A/B) (1, 3). However, some were IRES-enriched in only one of two replicate purifications (e.g. enriched in 1 of the 2 native purifications, or 1 of the 2 denaturing purifications; YB-1, PCBP2, NONO/p54nr, PTB, hnRNPK, hnRNPD, whereas others were purified equally well with IRES and Cap RNAs and thus designated common interactors (hnRNPC1/C2, ILF2/NF45, hnRNBU, hnRNPA/B) (3). Importantly, only a small set were reproducibly enriched with IRES target RNA (SFPQ/PSF, ELAV1/HU, and hnRNPD). In addition to known ITAFs, potential regulatory factors without documented connections to IRESs were enriched with the tagged-IRES mRNA. Several ribosomal subunits with noncanonical functions were IRES-enriched under the denaturing condition. For example, ribosomal subunit RPS19 and its tightly associating partner RPS7 were highly enriched. RPS19 is the most common mutated subunit in Diamond-Blackfan anemia and this protein binds directly to a specific sequence/structure in the 5’ UTR of its own mRNA (46, 50). Similarly, ribosomal subunit RPL26 acts noncanonically by binding to the 5’ UTR of p53 mRNA to activate translation (47, 51). Finally, elf3g, a subunit of the largest translation initiation factor was consistently IRES-enriched in all four purifications. elf3g is dispensable for cap-dependent translation and it has a recently discovered role in translation re-initiation, a form of translation regulation in the 5’ UTR of specific mRNAs (45, 52, 53). This is particularly interesting because the LEF1 5’ UTR contains a highly conserved upstream open reading frame (20). These proteins are potentially important new ITAF discoveries and a more detailed study of their structure and function with LEF1 mRNA should lead to a better understanding of ITAFs and the activities they build on IRES elements.

Two additional subgroups of IRES-enriched factors were reproducibly isolated with denaturing purification. One group is comprised of nucleosomal histones and chromatin-associated proteins (Fig. 4). Because UV irradiation only cross-links closely juxtaposed protein and nucleic acid, these proteins might be factors intimately associated with nascently transcribed target RNA when cotranscriptional processing brings mRNA and nucleosome complexes into close proximity (Fig. 3B) (54, 55). Also, the IRES-enriched factor DEK is known to associate with both exon junction complexes as well as nucleosomes and as the STRING network indicates, it could be a central linking factor that bridges chromatin and mRNA processing (56). Finally, a second IRES-enriched subgroup is notable in that two of the three known components of paraspeckles were consistently enriched with tagged-IRES purification (MATR3 and SFPQ) (39, 57, 58). The third protein component of paraspeckles, NONO/p54nr, was enriched in only one dataset (57). Paraspeckles are nuclear bodies built upon the NEAT1 noncoding RNA and they appear to function...
as nuclear retention bodies for highly structured RNAs and/or defective RNAs (39, 57). It is an interesting possibility that long, highly structured IRES mRNAs are retained in these structures and thus experience a different route from transcription to nuclear export compared with most mRNAs.

Although some of the IRES-enriched proteins have not been previously linked to IRESs and their activity, they have known, documented associations with one another (STRING analysis, Fig. 4), suggesting that the identified proteins might work together as a large IRES RNA/protein complex. In addition, several ribosomal subunits and translation initiation factors were enriched on the IRES target; their quantitative difference in association with tagged-CAP versus tagged-IRES RNAs hints that there could be differences in the way that conventional translation machinery associates with IRESs. Furthermore, as demonstrated by validation studies with Hupipistanol, inhibition of eIF4A-1 shows differential effect on canonical cap-dependent translation and LEFT1 IRES mediated translation. All of these identified associations derive from an in vivo assembly of the IRES target RNA and thus represent the first and most authentic snapshot of LEFT IRES regulatory factors. Thus, our strategy has a notable advantage over other methods in addition to its potential for broad application in studying in vivo RNA-protein complexes.

Acknowledgments—We thank Dr. Klemens Hertel (UC Irvine, Irvine, CA) for the gift of MS2 stem-loop plasmid, Dr. David Peabody (University of New Mexico, Albuquerque, NM) for the N55K MS2 mutant clone and for advice, Maxfield Tang for generating the tagged-CAP plasmid, Dr. Harry J. Mangalam (UC Irvine, Irvine, CA) for custom script for data compilation and sorting, Dr.s Ruslan Aphasizhev, Yongsong Shi, Bert Semler and members of the Semler laboratory (UC Irvine, Irvine, CA) for advice, reagents, and critique, Noriko Yokoyama, Kira Pate, and Rachel Cinco for their contribution to the purification process and Stephanie Sprowill for her assistance with the microscopy images. We would like to thank Waterman and Huang lab members for their help and support during this study.

This work was supported by National Institutes of Health (NIH) R21CA133275 to M. W. and NIH GM074830 to L. H.

© This article contains supplemental Fig. S1 and Supplemental Tables S1 to S6.

† To whom correspondence should be addressed: Medical Science I, B260, Department of Microbiology and Molecular Genetics.

REFERENCES

1. Fitzgerald, K. D., and Semler, B. L. (2009) Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. Biochim. Biophys. Acta. 1789, 518–528

2. Jackson, R. J., Hellen, C. U., and Pestova, T. V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. 11, 113–127

3. Pacheco, A., and Martinez-Salas, E. (2010) Insights into the biology of IRES elements through riboproteomic approaches. J. Biomed. Biotechnol. 2010, 458927

4. Stoneley, M., and Willis, A. E. (2004) Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. Oncogene. 23, 3200–3207

5. Balvay, L., Soto Rifo, R., Ricci, E. P., Decimo, D., and Ohlmann, T. (2009) Structural and functional diversity of viral IRESs. Biochim. Biophys. Acta. 1789, 542–557

6. Gilbert, W. V. (2010) Alternative ways to think about cellular internal ribosome entry. J. Biol. Chem. 285, 29033–29038

7. Cobbold, L. C., Spriggs, K. A., Haines, S. J., Dormann, H. C., Hayes, C., de Moor, C. H., Lilley, K. S., Bushnell, M., and Willis, A. E. (2008) Identification of internal ribosome entry segment (IRES)-trans-acting factor for the Myc family of IRESs. Mol. Cell. Biol. 28, 40–49

8. Fujimura, K., Kano, F., and Murata, M. (2008) Identification of PCBP2, a facilitator of IRES-mediated translation, as a novel constituent of stress granules and processing bodies. RNA. 14, 425–431

9. Lewis, S. M., Veyrier, A., Hosszu Ungureanu, N., Bonnal, S., Vagner, S., and Holm, M. (2007) Subcellular relocalization of a trans-acting factor regulating IAP IRES-dependent translation. Mol. Biol. Cell. 18, 1302–1311

10. Nonne, N., Ameyar-Zazoua, M., Souidi, M., and Harel-Bellan, A. (2010) Tandem affinity purification of miRNA target mRNAs (TAP-Tar). Nucleic Acids Res. 38, e20

11. Haefner, M., Landthaler, M., Burger, L., Khorshid, M., Haussler, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A. C., Munschauer, M., Ulbrich, A., Wardle, G. S., De, Zavolan, M., and Tuschi, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell. 141, 129–141

12. Tyagi, S. (2009) Imaging intracellular RNA distribution and dynamics in living cells. Nat. Methods. 6, 331–338

13. Butter, F., Scheibe, M., Mörl, M., and Mann, M. (2009) Unbiased RNA-protein interaction screen by quantitative proteomics. Proc. Natl. Acad. Sci. U.S.A. 106, 10626–10631

14. Hogg, J. R., and Collins, K. (2007) RNA-based affinity purification reveals 75K RNPs with distinct composition and regulation. RNA. 13, 868–880

15. Keryer-Bibens, C., Barreau, C., and Osborne, H. B. (2008) Tethering of proteins to RNAs by bacteriophage proteins. Biol. Cell. 100, 125–138

16. Lim, F., Spingola, M., and Peabody, D. S. (1994) Altering the RNA binding specificity of a translational repressor. J. Biol. Chem. 269, 9006–9010

17. Hook, B., Bernstein, D., Zhang, B., and Wickens, M. (2005) RNA-protein interactions in the yeast three-way hybrid: synthesis, affinity, sensitivity, and enhanced library screening. RNA. 11, 227–233

18. Tagwerker, C., Flick, K., Cui, M., Guerrero, C., Dou, Y., Auer, B., Baldi, P., Huang, L., and Kaiser, P. (2006) A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivisco-linking. Mol. Cell Proteomics. 5, 737–748

19. Wang, X., Chen, C. F., Baker, P. R., Chen, P. L., Kaiser, P., and Huang, L. (2007) Mass spectrometric characterization of the affinity-purified human 26S proteosome complex. Biochemistry. 46, 3553–3565

20. Jimenez, J., Jang, G. M., Semler, B. L., and Waterman, M. L. (2005) An internal ribosome entry site mediates translation of lymphoid enhancer factor-1. RNA. 11, 1385–1399

21. Wang, X., and Huang, L. (2008) Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. Mol. Cell Proteomics. 7, 46–57

22. Oeffinger, M., Wei, K. E., Rogers, R., DeGrasse, J. A., Chait, B. T., Altschison, J. D., and Rout, M. P. (2007) Comprehensive analysis of diverse ribonucleoprotein complexes. Nat. Methods. 4, 951–956

23. Ule, J., Jensen, K., Mele, A., and Darnell, R. B. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. Mol. Cell. Biol. 25, 1376–1386

24. Kaake, R. M., Milenkovic, T., Przulj, N., Kaiser, P., and Huang, L. (2010) Characterization of cell cycle specific protein interaction networks of the yeast 26S proteosome complex by the QTAX strategy. J. Proteome Res. 9, 2016–2029

25. Jensen, L. J., Kuhn, M., Stark, M., Chaﬀron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., and von Mering, C. (2009) STRING 8—a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 37, D412–416

26. Thomas, P. D., Campbell, M. J., Kejarwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003) PANTHER: a library of protein families and subfamilies indexed by function. Genome Res. 13, 2129–2141

27. Bordwell, V. J., and Wickens, M. (1990) Purification of RNA and RNA-
protein complexes by an R17 coat protein affinity method. *Nucleic Acids Res.* **18**, 6587–6594

28. Peabody, D. S. (1993) The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.* **12**, 595–600

29. Carey, J., Cameron, V., de Haseth, P. L., and Uhlenbeck, O. C. (1983) Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. *Biochemistry.* **22**, 2601–2610

30. Gravelle, B. R., and Maniatis, T. (1998) Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol. Cell.* **1**, 765–771

31. Zhou, Z., Licklider, L. J., Gygi, S. P., and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature.* **419**, 182–185

32. Chubb, J. R., Bloomfield, G., Xu, Q., Kaller, M., Ivens, A., Skelton, J., Turner, B. M., Nellen, W., Shaulsky, G., Kay, R. R., Bickmore, W. A., and Singer, R. H. (2006) Developmental timing in Dictyostelium is regulated by the Set1 histone methytransferase. *Dev. Biol.* **292**, 519–532

33. Golding, I., Paulsson, J., Zawilski, S. M., and Cox, E. C. (2005) Real-time kinetics of gene activity in individual bacteria. *Cell.* **123**, 1025–1036

34. Bertrand, E., Chartrand, P., Schaefer, M., ShenoY, S. M., Singer, R. H., and Long, R. M. (1998) Localization of ASH1 mRNA particles in living yeast. *Mol. Cell.* **2**, 437–445

35. Greentzmann, G., Ingram, J. A., Kelly, P. J., Gesteland, R. F., and Atkins, J. F. (1998) A dual-luciferase reporter system for studying recoding signals. *Rna.* **4**, 479–486

36. Ong, S. E., Blagove, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics.* **1**, 376–386

37. Kaake, R. M., Wang, X., and Huang, L. (2010) Profiling of protein interaction networks of protein complexes Using affinity purification and quantitative mass spectrometry. *Mol. Cell Proteomics.*

38. Guerrero, C., Tagwerker, C., Kaiser, P., and Huang, L. (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol. Cell Proteomics.* **5**, 366–378

39. Clemson, C. M., Hutchinson, J. N., Sara, S. A., Ensminger, A. W., Fox, A. H., Chess, A., and Lawrence, J. B. (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell.* **33**, 717–726

40. Shav-Tal, Y., and Zippori, D. (2002) PSF and p54(nrb)/NonO–multi-functional nuclear proteins. *FEBS Lett.* **531**, 109–114

41. Bertolotti, A., Lutz, Y., Heard, D. J., Chambon, P., and Tora, L. (1996) hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* **15**, 5022–5031

42. Law, W. J., Cann, K. L., and Hicks, G. G. (2006) TLS, EWS and TAF15: a model for transcriptional integration of gene expression. *Brief Funct. Genomic Proteomic.* **5**, 8–14

43. Bordeleau, M. E., Mori, A., Oberer, M., Lindqvist, L., Chard, L. S., Higa, T., Belsham, G. J., Wagner, G., Tanaka, J., and Pelletier, J. (2006) Functional characterization of IRE5es by an inhibitor of the RNA helicase elF4A. *Nat. Chem. Biol.* **2**, 213–220

44. Spring, K. A., Cobbold, L. C., Jopling, C. L., Cooper, R. E., Wilson, L. A., Stoneley, M., Coldwell, M. J., Poncet, D., Shen, Y. C., Morley, S. J., Bushell, M., and Willis, A. E. (2009) Canonical initiation factor requirements of the Myc family of internal ribosome entry segments. *Mol. Cell Biol.* **29**, 1565–1574

45. Cuchalova, L., Koubá, T., Herrmannová, A., Daný, I., Chiu, W. L., and Valášek, L. (2010) The RNA recognition motif of eukaryotic translation initiation factor 3g (elF3g) is required for resumption of posttermination ribosomes for reinitiation on GCN4 and together with elF3l stimulates linear scanning. *Mol. Cell. Biol.* **30**, 4671–4686

46. Campagnoli, M. F., Ramenghi, U., Armiragli, M., Quarello, P., Garelli, E., Carando, A., Avondo, F., Pavesi, E., Frisbeurgh, S., Gleizes, P. E., Lorenzi, F., and Dianzani, I. (2008) RPS19 mutations in patients with Diamond-Blackfan anemia. *Hum. Mutat.* **29**, 911–920

47. Chen, J., and Kastan, M. B. (2010) 5′-3′UTR interactions regulate p53 mRNA translation and provide a target for modulating p53 induction after DNA damage. *Genes Dev.* **24**, 2146–2156

48. Kreic, A. M., and Swanson, M. S. (1999) hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* **11**, 363–371

49. Vasudevan, S., and Steitz, J. A. (2007) AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell.* **128**, 1105–1118

50. Schuster, J., Fröjmark, A. S., Nilsson, P., Badhai, J., Virtanen, A., and Dahl, N. (2010) Ribosomal protein S19 binds to its own mRNA with reduced affinity in Diamond-Blackfan anemia. *Blood Cells Mol. Dis.* **45**, 23–28

51. Boggs, K., Henderson, B., and Reisman, D. (2009) RBP-Jkappa binds to and represses transcription of the p53 tumor suppressor gene. *Cell Biol. Int.* **33**, 318–324

52. Phan, L., Schoenfeld, L. W., Valaske, L., Nielsen, K. H., and Hinnebusch, A. G. (2001) A subcomplex of three elF3 subunits binds elF1 and elF5 and stimulates ribosome binding of mRNA and tRNA(met). *EMBO J.* **20**, 2954–2965

53. Masutani, M., Sonenberg, N., Yokoyama, S., and Imataka, H. (2007) Reconstitution reveals the functional core of mammalian elF3. *EMBO J.* **26**, 3373–3383

54. Levchenko, V., and Jackson, V. (2004) Histone release during transcription: NAP1 forms a complex with H2A and H2B and facilitates a topologically dependent release of H3 and H4 from the nucleosome. *Biochemistry.* **43**, 2359–2372

55. Moore, M. J., and Proudfoot, N. J. (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell.* **136**, 688–700

56. Kappes, F., Burger, K., Baack, M., Fackelmayer, F. O., and Gruss, C. (2001) Subcellular localization of the human proto-oncogene protein DEK. *J. Biol. Chem.* **276**, 26317–26323

57. Fox, A. H., Lam, Y. W., Leung, A. K., Lyon, C. E., Andersen, J., Mann, M., and Lamond, A. I. (2002) Paraspeckles: a novel nuclear domain. *Curr. Biol.* **12**, 13–15

58. Salton, M., Lerenthal, Y., Wang, S. Y., Chen, D. J., and Shiloh, Y. (2010) Involvement of matrin 3 and SFPQ/NONO in the DNA damage response. *Cell Cycle.* **9**

In order to cite this article properly, please include all of the following information: Tsai, B. P., Wang, X., Huang, L., and Waterman, M. L. (2011) Quantitative Profiling of In Vivo-assembled RNA-Protein Complexes Using a Novel Integrated Proteomic Approach. *Mol. Cell. Proteomics* **10**(4):M110.007385. DOI: 10.1074/mcp.M110.007385.