COPI transport complexes bind to specific RNAs in neuronal cells

Adrian G. Todd¹, Hai Lin², Allison D. Ebert³, Yunlong Liu² and Elliot J. Androphy¹,*

¹Department of Dermatology and ²Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, USA and ³Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI 53226, USA

Received September 13, 2012; Revised October 22, 2012; Accepted November 13, 2012

Our fundamental understanding of how several thousand diverse RNAs are recognized in the soma, sorted, packaged, transported and localized within the cell is fragmentary. The COPα and COPβ proteins of the coat-omer protein I (COPI) vesicle complex were reported to interact with specific RNAs and represent a candidate RNA sorting and transport system. To determine the RNA-binding profile of Golgi-derived COPI in neuronal cells, we performed formaldehyde-linked RNA immunoprecipitation, followed by high-throughput sequencing, a process we term FLRIP-Seq (FLRIP, formaldehyde-cross-linked immunoprecipitation). We demonstrate that COPα co-immunoprecipitates a specific set of RNAs that are enriched in G-quadruplex motifs and fragile X mental retardation protein-associated RNAs and that encode factors that predominantly localize to the plasma membrane and cytoskeleton and function within signaling pathways. These data support the novel function of COPI in inter-compartmental trafficking of RNA.

INTRODUCTION

Establishment of localized protein synthesis dependent on sub-cellular RNA transport is essential for organismal development and cellular polarity, particularly in neuronal cells in which the sites of translation span long distances (1). However, the proteins and mechanisms governing the spatiotemporal trafficking of large numbers of individual RNAs into distinct subcellular compartments are largely unresolved and the repertoire of RNAs ultimately destined for peripheral localization within the neurite remain to be defined (2–6). A variety of human pathologies have been described where a defect in RNA processing, transport or localized translation within the neuron is thought to trigger the disease state (7). The majority of research aimed at understanding the composition of ribonucleoprotein (RNP) particle complexes has focused on direct RNA–protein interactions, characterizing RNA-binding domains and their RNA recognition motifs and understanding how cis- and trans-elements coordinate active transport and regulate post-transcriptional processing (8). Given the spectrum of emerging RNA-binding proteins (RBPs) and their specific target motifs, localization likely involves coordination of diverse macromolecular complexes and transient interactions from proteins that do not interact directly with the RNA.

Vesicle complexes of the endoplasmic reticulum (ER) and Golgi apparatus are tantalizing mediators of localizing RNAs and are positioned to mobilize cytoplasmic RNAs destined for the peripheral neurite (9). COPα (alpha subunit) is the major subunit of coatamer protein I (COPI), which actively moves cargoes within the Golgi-ER in somatic cells (10), and unlike clathrin or COPII vesicles, it displays a highly metamorphic range of subcomplexes composed of vesicle subunits believed to be functional in directing a diverse range of cargoes within the cell (11). COPα directly binds to the survival motor neuron (SMN) protein, an RNP assembly factor, which co-traffic together within the neurite (12). COPα and COPβ are components of the post-synaptic Proteome (PSP) and post-synaptic density (PSD) and co-immunoprecipitate the β-actin mRNA (12–14). Disruption of COPI complex components or accessory proteins results in neurodegeneration and mis-localization of RNA (15–17). Mutations in the murine archain 1 gene encoding the delta-subunit of COPI lead to cerebellar atrophy and Purkinje neuron degeneration (17). Similarly, disruption of Scyl 1, a COPI complex accessory protein that binds to the beta-1 subunit of COPI, leads to motor neuron loss and muscle wasting in the mdf mutant mouse (15). Alteration of COPI function results in the mis-localization of RNA in yeast (16). Collectively, these
observations suggest that the COPI pathway may be an evolutionarily conserved mechanism of RNP trafficking.

To determine the RNA-binding profile of Golgi-derived COPI proteins in neuronal cells, we performed formaldehyde-linked RNA immunoprecipitation of COPI, followed by high-throughput sequencing, a process we refer to as FLRIP-Seq (FLRIP, formaldehyde-cross-linked immunoprecipitation). This methodology permits recovery of RNAs associated with COPI that may not bind directly to COPI. We demonstrate that COPI complexes incorporate a specific set of RNAs that harbor putative neurite-targeting motifs, display significant overlap with neuronal RBPs and are known to localize to the plasma membrane and cytoskeleton.

RESULTS

RNA interaction map of COPI complexes in NSC34 cells

As the association between COPI and RNA is likely facilitated through its interactions with authentic RBPs, we applied formaldehyde to a clonal population of motor neuron-like NSC34 cells to covalently cross-link COPI protein complexes to RNA. After capture using a COPI antibody (12), high-stringency washes (18) and formaldehyde cross-link reversal were performed to optimize signal-to-noise ratios and ensure maintenance of RNA integrity throughout the protocol.

Isolation of RNA for transcriptome and FLRIP samples was performed in triplicate using serum starvation to induce differentiation and neurite formation (19) (Fig. 1A). The input of these three independent isolates defines the total transcriptome and was determined from 25–28 million 50 bp sequences (Fig. 1C), of which 9–14 million were mapped to the mouse genome, generating a quantitative representation of the NSC34 transcriptome, following serum starvation. Unique reads with no more than two mismatches identified transcripts from 12,168 genes. Mapping summaries are presented in Figure 1A and Supplementary Material, Table S1. Approximately 37% of COPI-FLRIP reads mapped to coding regions of the genome, similar to the proportion observed with the input NSC34 transcriptome (Fig. 1B). A total of 1,519 transcripts were significantly enriched in COPI-impregnated samples (Fig. 1C and D, false discovery rate (FDR) < 0.05, Supplementary Material, Table S2).

Enrichment of specific RNAs in the COPI transcriptome

To confirm association, several RNAs were selected, with enrichment ranging between 2- and 16-fold to validate the data set based on mean reads per kilobase per million (RPKM) scores of 100 or more and mapping to multiple exons or to the 5′- or 3′-UTR (Fig. 2A), including leucyl-tRNA synthetase 2 (Lars2), Synaptojanin 2 (Synj2), Fibroin (Fbrs), activating transcription factor 5 (ATF5), 2-oxyglutarate iron-dependent oxygenase domain containing 2 (Ogfod2) and IMP (inosine 5′-monophosphate) dehydrogenase 1 (IMPDH1) (Fig. 2B). Methylsterol monooxygenase 1 (Sc4mol) and phosphatidylinositol glycan anchor biosynthesis, class F (Pigf) showed no enrichment in FLRIP samples and were used for comparison (Supplementary Material, Table S3). FLRIP from fresh NSC34 cells was performed to confirm COPI enrichment of these RNAs (Figs 2C and 3C). Specificity of interaction was evidenced by comparing FLRIP with antibodies to COPI and Growth-Associated Protein 43 (GAP43), a prominent neuronal protein locally translated in growth cones (20). This target was preferred to an immunoglobulin control as its antibody efficiently recognized its specific target via immunoprecipitation (data not shown). Lars2, Synj2, Fbrs, ATF5 and Ogfod2 showed specific association with the COPI, whereas IMPDH1 associated with both COPI and GAP43 (Fig. 2D). To determine whether COPI associates with RNA in other neuronal systems, COPI-FLRIP RNAs from motor neurons derived from human induced pluripotent stem cell (iPSC) cultures were screened, and detection of Synj2 but not Pigf associations with COPI was evident (Fig. 2E).

COPI subunits display differential association with COPI-associated RNAs

To support the implication that it was the COPI coatamer complex that recognized specific RNAs, we tested other COPI factors for ability to immunoprecipitate the identified RNAs (Fig. 3). Antibodies to the COPγ and COPγ2 subunits displayed high affinities for their endogenous targets (Fig. 3A) and we were used to screen FLRIP libraries for Lars2, Synj2, ATF5, Fbrs, Ogfod2 and Sc4mol (Fig. 3B). Interestingly, only Ogfod2 RNA was detectable in both COPγ and COPγ2 samples, whereas other COPI targets were detected with COPγ (Synj2), or COPγ2 (Lars2, ATF5) or neither (Fbrs) (Fig. 3B). Sc4mol was not amplified in COPγ and COPγ2 samples (Fig. 3B).

COPI-associated RNAs are candidates for peripheral localization

Enrichment of a large number of RNAs associated with COPI prompted us to determine whether any were common to annotated gene databases from neuronal tissue. Comparison of COPI-FLRIP RNAs with the Genes to Cognition database showed 8% of COPI-associated RNAs were common to the PSP and PSD data sets (125 and 130 genes, respectively, Supplementary Material, Table S4–S6) (14). Eight to 10% of COPI-associated RNAs comprise the axonal transcriptome (Supplementary Material, Table S4) (21), supporting a role for COPI peripheral RNA targeting. The DAVID Gene Ontology (GO) database was queried with all COPI-associated RNAs against the NSC34 transcriptome to elucidate cellular compartments and molecular functions that may be regulated by COPI–RNA interaction (22). Highest scoring cellular compartments (CC) were the cytoskeleton (9.5% of annotated RNAs, $P = 6 \times 10^{-16}$) and plasma membrane (14.5% of annotated RNAs, $P = 3 \times 10^{-14}$) (Fig. 4A). Significant pathways associated with COPI-interacting RNAs included MAPK (2.8%, $P = 9 \times 10^{-42}$) and cancer signaling (2.6%, $P = 4.1 \times 10^{-35}$).

COPI pathway disruption reduces RNA localization in yeast; therefore, the 3′ UTRs of COPI-associated RNAs were queried for common motifs (Fig. 4B) (23). A total of 393 (25.9%) COPI-associated RNAs harbor G-quadruplex motifs (Fig. 4B). G-quadruplexes may act as neurite-targeting elements when present in 3′ UTRs (24). Two G-quadruplex
motifs were analyzed for enrichment in COPa-associated 3′ UTRs. Both Motif 1 (GGG-N(0–6)-GGG) (25) and a more complex G-quadruplex Motif 2 (DWGG(N)(0–2)DWGG(N)(0–2)DWGG(N)(0–2)DWGG) (26), where D denotes any nucleotide except C and W denotes A or C nucleotides, were enriched in COPa-interacting RNAs ( \( P = 8.2 \times 10^{-71} \) and \( P = 1.5 \times 10^{-69} \), respectively, Fig. 4C).

Our analysis suggested a mechanism by which COPa associates with a distinct subset of RNAs to mediate their distal regulation. To determine whether COPa functions in combination with specific neuronal RBPs to facilitate RNA processing, we compared COPa-associated RNAs with those of two neuronal RBPs, neuro-oncological ventral antigen 1 (NOVA1) and fragile X mental retardation protein (FMRP) (27,28). Significant overlap was observed between COPa and FMRP ( \( P = 7.2 \times 10^{-56} \), Fig. 4D) and NOVA1 ( \( P = 0.02 \))-associated RNAs. Taken together, these data suggest that COPa interacts with a subset of RNAs that are localized peripherally and contain neurite-targeting motifs. The apparent non-exclusive nature of these RNAs as determined by pathway analysis suggests that these interactions may be indirect and mediated by additional RBPs such as the G-quadruplex-binding protein FMRP (24,29).

**DISCUSSION**

An intricate and precise series of events must be necessary to facilitate the subcellular distribution of RNAs. This is of particular importance in neurons, where RNAs must be sorted and selectively transported over long distances. RNA signatures or localization elements must be recognized by specific RNA recognition proteins, which are mobilized by transport carriers connected to molecular motors (1). The identities of the components encompassing these processes remain largely unknown. COPI and COPII are major cargo carrier vesicles derived from the remodeling and budding of the Golgi membrane, with the latter supporting trafficking in dendrites (9,30). Nuclear membrane remodeling and budding in muscle has been recently reported as a mechanism to facilitate RNP transport of Par6 mRNA to the neuromuscular junction in *Drosophila*, providing evidence that RNP trafficking is...
facilitated through membrane-derived components (31). The evolutionarily conserved COPa protein is the major subunit of Golgi-derived COPI coatomer and has been shown to move within neurites together with its binding partner SMN (12). Golgi positioning is a prerequisite of neuronal polarity and axonal specification, whereas disruption of COPI coat formation by brefeldin A and depletion of COPa suppresses neurite outgrowth (12,32,33 and our unpublished observations). It is tempting to speculate that these Golgi-derived vesicles provide an additional mechanism to facilitate RNP trafficking within the neuron.

A recent publication found that COPI-coated complexes assume a range of conformations, in contrast to COPII or clathrin-coated vesicles, which may facilitate interactions with different cargoes and cellular addresses (11). COPI was previously reported to be in complex with the β-actin and kor RNAs, and in conjunction with other reports (15–17,34), the data suggest that components of COPI vesicles act as authentic transport vehicles within the axon and dendrite (3,6,9,11,35–37). Although the alpha-subunit of coat-omer reportedly binds asialoglycoprotein receptor (ASGR) RNA in vitro (34) and contains multiple WD40 domains that have been reported to interact directly with RNA (38), we have been unable to detect specific RNA binding to purified COPa (data not shown). The beta subunit of COPI has also been reported to mediate transport of the kor mRNA via an indirect interaction (13). Although COPa immunoprecipitates with β-actin RNA (12), by RPMK reads β-actin was present in the COPI FLRIP reactions, but we were unable to detect enrichment of β-actin or other classical axonal RNAs, such as MAP2 or CaMKIIa by COPa-FLRIP, when normalized to the total pool of RNA. As only a small fraction of β-actin RNA localizes to the axon (39), if the interaction between COPa and RNA is restricted to a specific subcellular

RNPs is surprisingly low, in the 1–2 transcript range (40). Whether both mechanisms cooperate in regulating cargo dependent upon sub-complex conformations (11), and our observations that the COPI pathway and associated COPI transport pathway could act as a ‘go between’ coupling RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspending RNA–protein complexes was performed by resuspecting RNAs between COPI and FMRP, we propose that COPI complexes interact with RBPs such as SMN and FMRP to aid in the identification and translational regulation of a specific subset of RNAs (12,27,44). We are currently investigating the dependence of COPI-RNA complexes on SMN and FMRP.

MATERIALS AND METHODS

Cell culture

Murine neuroblastoma x spinal cord NSC-34 cells (19) were grown in DMEM (Gibco) supplemented with 10% FCS (Clontech) and 1% penicillin/streptomycin (PS) at 37°C. Cells were grown to 80% confluence in a 150 mm dish and then split into two 150 mm dishes at a density of 2.36 × 10⁴ cells/ml in DMEM:F12 (Gibco) supplemented with 1% FCS and 1% PS to induce differentiation and promote neurite outgrowth. Cells were grown for 72 h until ∼70% confluent with media changed after 48 h.

Formaldehyde-cross-linked immunoprecipitation

Approximately 4 × 10⁶ cells were used per FLRIP. Cross-linking of RNA protein complexes was performed by removing media and replacing it with fresh media supplemented with methanol-free paraformaldehyde (Pierce) to a final concentration of 1%. After rocking at room temperature for 10 min, 1 ml of 1 M glycine was added to terminate cross-linking. Cells were washed twice with phosphate buffered saline, trypsinized, washed and pelleted.

Cross-linked pellets were lysed in polysome lysis buffer (PLB) modified slightly from that described (18) [100 mM KCl, 5 mM MgCl₂, 10 mM HEPES (pH 7.0), 0.5% NP-40, 1 mM DTT (Invitrogen), 100 U/ml RNaseOUT (Invitrogen), 2 mM vanadyl ribonucleoside complexes solution (NEB), 1× EDTA-free protease inhibitor cocktail (Roche)]. Cell pellets were lysed in 500 μL of PLB containing inhibitors and incubated on ice for 10 min. Lysates were cleared by centrifugation at 15 000g at 4°C for 15 min. For input analysis, 50 μL of lysate was removed, and the remaining 450 μL of extract split equally and used for isolation of cross-linked COPI complexes in duplicate. Total lysate volumes were made up to 500 μL by addition of PLB and immunoprecipitation performed by incubation with affinity-purified COPI antibody (12) pre-coupled to M-280 sheep anti-rabbit Dynabeads (Invitrogen).

RNA isolation

Following overnight incubation, beads and isolated complexes were separated from the lysate by use of a magnet. Unbound lysate was aspirated and beads washed four times with 0.5 ml of PLB (without inhibitors), followed by PLB, including 1 mM urea according to Peritz et al. (18). Dissociation of RNA–protein complexes was performed by resuspending
beads in 100 \mu l of buffer containing 100 mM DTT, 5 mM EDTA, 10 mM Tris, pH 7.0, and 1% SDS and heating to 70°C for 45 min to reverse cross-links followed by the addition of 1.5 \mu l (20 mg/ml) of proteinase K (RNA grade; Invitrogen) and further incubation at 50°C for 1 h.

Nucleic acid was isolated from input and immunoprecipitated samples using mirVana RNA Isolation Kit (Ambion) according to manufacturer’s instructions before DNase treatment, using Turbo DNase Kit (Ambion).

RNA QC and library preparation
RNA quality and abundance was determined using Bioanalyzer 2100 (Aglient), following manufacturer’s protocol.

Sample preparation and high-throughput sequencing
The input RNAs were treated with Ribominus (Life Technologies) to reduce ribosomal RNA; the FLRIP samples were not treated. For the input RNA, concentration was measured using the Agilent Bioanalyzer 2100, and 250–700 ng of the input RNA was fragmented with RNase III. The amount of FLRIP RNA was too low to accurately measure. Both input and FLRIP samples were then processed through the Life Technologies’ SOLiD4 RNA Sequencing Protocol; each individual sample was barcoded as part of the process. The barcoded libraries were pooled in equal amounts and processed through SOLiD4 EZ Bead preparation (Life Technologies), using 0.5 pm total in the final preparation. One full slide of 700 million template beads were used for 50 base reads forward sequencing on the Life Technologies SOLiD 4 Sequencer.

Bioinformatics and statistics analysis
We used a customized pipeline for RNA-seq data analysis. The analysis includes three steps: QC filtering, sequence alignment and gene differential expression analysis.

QC filtering
We first used the SOLiDTM Instrument Control Software (ICS) and the SOLiDTM Experiment Tracking System (SETS) software for the recalibration of the read quality. Sequences containing more than two ‘N’ or wildcards were discarded. Each sequence was scanned for low-quality regions, and if a 5 base sliding window has an average quality score of <20, the read was truncated at that position. Any read with a length of <35 bases was discarded. Our experience suggests that such strategy effectively eliminated low-quality reads while retaining high-quality regions.
Alignment

We used BFAST (47) as our primary alignment algorithm because it has high sensitivity for color space data. We used a TopHat-like strategy (48) to align the sequencing reads to the mouse genome mm9. After aligning the sequence reads to a filtering index including repeats—rRNAs (ribosomal RNA) and other sequences are not of interest—we conducted sequence alignment for three levels: genomic, known junctions and novel junctions [based on all the combination of all known exons (UCSC known gene)]. Two nucleotide mismatches were allowed.

Differential gene expression analysis

edgeR was used to calculate differences between FLRIP and whole transcriptome (input) samples by assuming the RNA-sequencing counts, following a negative binomial distribution (49). In addition, Benjamini and Hochberg’s algorithm is used to control the FDR.

For motif analysis, we searched for motif enrichment in the 3’ UTR of 1467 IP-abundant transcripts. The remaining 52 transcripts did not have a 3’ UTR. Of the 1550 input-abundant transcripts, 1413 had 3’ UTRs and were used as background controls. Two classic motifs were used in our analysis: G(3+)N(0–6)G(3+) (25) and DWGG(N)(0–2)DWGG(N) (0–2)DWGG(N)(0–2)DWGG (26). Motifs were searched by in-house Perl script with regular expression. χ² test was performed to measure whether the occurrences of these motifs are enriched in FLRIP samples comparing with input-associated RNA populations.

COPa-FLRIP validation

For the validation of FLRIPs, protein G Dynabeads (Invitrogen) were used for the coupling of 1 μg of COPa, GAP43 (B-5 sc-17790; Santa Cruz Biotechnology, Inc.), COPy (C-19 sc-14167; Santa Cruz Biotechnology, Inc.) or COPy2 (C-20 sc14165; Santa Cruz Biotechnology, Inc.) antibodies under conditions described above. Quality control steps performed prior to sequencing quantified the yield of RNA isolated from FLRIP complexes in the region of a thousand times lower relative to input samples. Input samples were diluted to approximate the RNA concentration of FLRIP samples. Given the low concentrations and fold enrichment of RNAs in FLRIP samples (Supplementary Material, Table S1), we performed end-point PCR to determine the RPKM detectability threshold of RNA targets.

RT–PCR

DNase-treated RNA isolated from FLRIPs and inputs was used as template for cDNA library production, using Superscript III and manufacturer’s instructions for oligo(dT) primer (Invitrogen). PCR analysis was performed using GoTaq polymerase (Promega). Primer sequences used for analysis are presented in Supplementary Material, Table S7. To enhance amplification for Syn2, PigF, Sc4mol, ATF5 and Impdh1, 0.3 M Betaine (Sigma-Aldrich) was used.

iPSC screen

Formaldehyde-cross-linked iPSC-derived motor neuronal cultures (50) were processed for FLRIP and RNA extraction as described for NSC34 cells above. Prior to RT–PCR and cDNA library production, RNA was amplified using MessageBOOSTER™ cDNA synthesis kit for qPCR (Epiconcept) following manufacturer’s protocol. Resulting FLRIP and input cDNA was screened using primers against Syn2 (For 5’-GGAGACGCCTCGATGCTTC-3’, Rev 5’-GACGTGCCACCGATTTAGGC-3’) and PigF (For 5’-AGACCCTGGCAGTTAGGCCC-3’, Rev 5’-ACCATACACTTCTGCCCAGGCC-3’).

Gene Ontology

GO analysis was performed using COPa-associated RNAs as a query against the NSC34 transcriptome (22).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

Sequencing was carried out at the Center for Medical Genomics at Indiana University School of Medicine.

Conflict of Interest statement. None declared.

FUNDING

This work was supported in part by a grant from the Muscular Dystrophy Association USA to E.J.A.

REFERENCES


