

Friend of Prmt1, FOP is a Novel Component of the Nuclear SMN Complex Isolated Using Biotin Affinity Purification

Keiichi Izumikawa^{1,2}, Hideaki Ishikawa^{1,2}, Harunori Yoshikawa^{1,2}, Goro Terukina^{1,4}, Naoki Miyazawa^{1,2}, Hiroshi Nakayama^{2,3}, Yuko Nobe^{2,4}, Masato Taoka^{2,4}, Yoshio Yamauchi⁴, Sjaak Philipssen⁵, Toshiaki Isebe^{2,4} and Nobuhiro Takahashi^{1,2*}

¹Department of Applied Biological Science, United Graduate School of Agriculture, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan

²Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), Sanbancho 5, Chiyoda-ku, Tokyo 102-0075, Japan

³Biomolecular Characterization Team, RIKEN Advanced Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198 Japan

⁴Department of Chemistry, Graduate School of Science, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji, Tokyo 192-0397, Japan

⁵Department of Cell Biology, Erasmus MC, 3015 GE Rotterdam, The Netherlands

Abstract

SMN (survival motor neuron protein) complexes are essential for the biogenesis of uridine-rich small nuclear ribonucleoproteins (UsnRNPs). During the biogenesis, the SMN complexes bound to UsnRNPs are transported from the cytoplasm to the nucleus, and moved to Cajal body (bodies)/Gems (Cajal/Gems) where the SMN complexes-UsnRNPs are subjected to additional chemical modifications and dissociated to the SMN complexes and the mature UsnRNPs. Although the mature UsnRNPs are assembled into spliceosome with newly transcribed pre-mRNA in the perichromatin fibrils at the chromatin, the role of the dissociated nuclear SMN complexes remains undetermined. In this study, we identified Friend of Prmt1 (FOP; chromatin target of Prmt1, CHTOP; C1orf77) as a novel component of the nuclear SMN complexes by the biotin affinity purification, coupled with the mass spectrometry-based protein identification. FOP was associated with SMN, Gemins 2, 3, 4, 6, and 8, unrip, and fragile X mental retardation 1 protein (FMR1), as well as with U5 and U6 snRNAs in the nucleus, but not with Sm proteins, Gemin5, coilin, and U1 and U2 snRNAs. Using the quantitative proteomic method with SILAC coupled with RNA interference, we also showed that FOP is required for the association of the SMN complexes with hnRNPs, histone proteins, and various RNA-binding proteins. It is reported that FOP localizes mainly in the nuclear speckles, binds chromatin, and plays a role in mRNA transcriptional regulation. Our present data suggest that the nuclear SMN complex containing FOP participates in the process of mRNA post-transcriptional regulation.

Keywords: SMN; Gemin; C1orf77; Chromatin target of PRMT1; CHTOP; SRAG; LC-MS/MS; SILAC

Abbreviation: CHTOP: Chromatin Target of Prmt1; Coilin: Coilin p80; FMR1: Fragile X Mental Retardation 1 Protein; FOP: Friend of Prmt1; GAR: Glycine-Arginine-Rich; LC: Liquid Chromatography; MS: Mass Spectrometry; Prmt: Protein Arginine Methyl Transferase; SILAC: Stable Isotope Labeling using Amino Acids in Cell Culture; SMA: Spinal Muscular Atrophy; SMN: Survival of Motor Neuron Protein; SRAG: Small Protein Rich in Arginine and Glycine; unrip: unInteracting Protein; UsnRNP: Uridine-rich Small Nuclear Ribonucleoprotein

Introduction

SMN is linked to one of the most common inheritable causes of childhood mortality, spinal muscular atrophy (SMA) [1], a neuromuscular disorder characterized by loss of motor neurons, which results from low levels of, or loss-of-function mutations in the SMN protein [2,3]. The *SMN* gene is duplicated as an inverted repeat on chromosome 5. Only mutations in the telomeric copy, *SMN1*, give rise to SMA [4]; another centromeric copy, *SMN2*, expresses a functionally defective form of the protein that lacks the carboxyl terminus encoded by exon 7, but its mutations are not related to SMA.

Human SMN protein is a 294-amino-acid polypeptide that forms complexes with at least eight proteins: Gemin2 (SIP1), Gemin3 (DEAD box RNA helicase 20, DDX20), Gemin4, Gemin5 (a WD-repeat protein), Gemin6, Gemin7, Gemin8 and unrip (unr interacting protein, STRAP). The SMN complexes promote assembly of UsnRNP core complex consisting seven different Sm proteins, and U1, U2, U4, U5, or U7 snRNAs [5]; i. e., the SMN complexes assist the ordered

assembly of the seven Sm proteins as three sub-complexes, B/B'-D3, D1-D2 and E-F-G Sm onto the Sm sequence of UsnRNAs [6]. Protein arginine methyl transferase (Prmt) 5 promotes this assembly, in which it catalyzes symmetrical arginine methylation of some of the Sm core proteins [7]. Indeed, SMN interacts with a dimethylarginine residue on some of the Sm core proteins present in the UsnRNP core complex [8]. Gemin5 and unrip also participate in this UsnRNP assembly, and are detached from the SMN complexes-UsnRNPs at later stages of the biogenesis in the nucleus [9-11]. The assembled cytoplasmic UsnRNPs are transported into the nucleus, and are targeted to Cajal/Gems (splicing snRNPs) or to the histone locus body (U7 snRNP) that constitutes a part of Cajal/Gems. Coilin p80 (coilin) is required for the localization of SMN complexes in the Cajal/Gems through direct interaction of SMN with a symmetrically dimethylated arginine in a RG domain of coilin [12,13]. Coilin also interacts with Sm proteins in UsnRNPs through its C-terminal domain, and the binding is not

***Corresponding author:** Nobuhiro Takahashi, Ph.D., Department of Applied Biological Science, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan, Tel./Fax: 81-042-367-5709; E-mail: ntakahas@cc.tuat.ac.jp

Received October 21, 2013; **Accepted** December 23, 2013; **Published** December 26, 2013

Citation: Izumikawa K, Ishikawa H, Yoshikawa H, Terukina G, Miyazawa N, et al. (2013) Friend of Prmt1, FOP is a Novel Component of the Nuclear SMN Complex Isolated Using Biotin Affinity Purification. J Proteomics Bioinform S7: 002. doi:10.4172/jpb.S7-002

Copyright: © 2013 Izumikawa K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

affected by dimethylation [14]. Sm proteins and UsnRNPs are probably released from the SMN complexes upon the binding with coilin in Cajal/Gems [15], and are then transferred into splicing speckles [16]. Dissociation of Gemin5 may also promote the release of UsnRNP from SMN complexes in Cajal/Gems, since Gemin5 is required for the initial binding of snRNA to the SMN complexes in the cytoplasm [17]. The UsnRNPs that are released from the Cajal/Gems then receive additional proteins from the nuclear speckles. The final assembly of the spliceosome from its component UsnRNPs (and other proteins) takes place *in situ* on nascent pre-mRNAs at the site of transcription that occurs at the outer edge of the nuclear speckle, where all or almost all introns are removed in conjunction with or soon after the transcription [18]. The nuclear speckles are rich in spliced mRNAs and various pre-mRNA metabolic factors, such as spliceosome assembly factor SC35, and thus, are believed to be the places that complete mRNA maturation, release for mRNA export with some recycling factors, and recycle splicing complex including UsnRNPs [19]. Although it has been well documented that improper process of mRNAs impedes transit of the mRNAs through the nuclear speckle [20,21], the mechanism how the mRNAs are retained in recycle splicing complex and impeded the transit remains undetermined. The splicing complexes including UsnRNPs that are recycled after each round of splicing into the cytoplasm, must pass again through the Cajal/Gems for re-assembly before they can take part in a further round of splicing [22-24].

In this study, we isolated SMN complexes present in the nucleus and identified Friend of Prmt1, FOP as a novel component of the nuclear SMN complexes. As noted by alternative names of FOP such as SRAG (small protein rich in arginine and glycine) [25] and CHTOP (Chromatin target of Prmt1) [26], human FOP has a central glycine-arginine-rich (GAR) domain containing 26 RG/GR repeats that is recognized and methylated with Prmt1, localized mainly in the nuclear speckles and tightly associated with facultative chromatin in interphase cells [26]. The C-terminus of FOP also harbors a duplication of the sequence LDXXLDAYM (where "X" is any amino acid) [26]. FOP is expressed widely among tissues and cell lines [25], and is highly conserved in all vertebrates, while no orthologs was identified so far in yeast, worm, and fly. FOP has a critical role in specific mRNA transcriptional regulation during fetal globin expression [27], and is also involved in mRNA export as a component of TREX complex [28].

Materials and Methods

Reagents and antibodies

Antibodies against SMN (2B1), Gemin2 (2E17), Gemin3 (12H12), Gemin4 (17D10), Gemin6 (20H8), Gemin8 (1F8), and Unrip (3G6) were generous gifts from Dr. Gideon Dreyfuss. Anti-Gemin5 was obtained from Bethyl Laboratories. Anti-Sm (Y-12) was obtained from Thermo Fisher Scientific. Antibody to fibrillarin, lamin B, p32, and FMR1 were obtained from Santacruz. Anti-FLAG-M2, anti-coilin p80, and anti-FLAG-M2-conjugated agarose beads were obtained from Sigma-Aldrich. Antibody to FOP (KT64) was prepared as described [26]. Anti-GAPDH was obtained from Ambion. Anti-Prmt1 was obtained from Upstate. HRP conjugated anti-mouse IgG was obtained from Cell Signaling Technology. HRP-conjugated anti-rabbit IgG were obtained from Sigma-Aldrich. All general reagents were purchased from Wako Pure Chemical, Kanto Chemical, or Nacalai Tesque.

Cell culture

Flp-In T-REx 293 (293TRex) cells, 293EBNA cells, HEK293 cells,

and 293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich, St. Louis, MO) (final concentration of glucose in medium was 4.5 mg/l for 293TRex cells, and 1.0 mg/l for HEK293, 293EBNA and 293T cells) supplemented with 10% heat-inactivated fetal bovine serum (Biowest LLC, Miami, FL), containing streptomycin (0.1 mg/ml, WAKO Pure Chemicals, Osaka, Japan), and penicillin G (100 U/ml, WAKO Pure Chemicals) at 37°C under 5% CO₂ in air.

Construction of epitope-tagged expression plasmids

RT-PCR was performed with the Super Script II kit (Invitrogen) using total mRNA prepared from HEK293 cells. cDNAs encoding SMN (NM_000344) was amplified by PCR using primer sets 5'-GAAGAAGGATCCGCGATGAGCAGCGCGGCAGT-3' and 5'-GAAGAACTCGAGATTTAAGGAATGTGAGCACCT-3'. The PCR products were cloned into the BamH I/Xho I site of pcDNA3.1(+). The cloned cDNAs were subcloned into pcDNA 3.1(+)-DAP (encoding biotinylation sequence and FLAG tags and an additional N-terminal epitope tag 6×histidine) [29]. The DAP tag-fused SMN (HB-SMN-F) coding sequence was excised with Bgl II/Apa I and ligated into the Bgl II/Apa I site of pcDNA5/FRT for inducible expression (HB-SMN1-F-pcDNA5/FRT). cDNA encoding FOP (NM_015607) was amplified by PCR using primer sets 5'-TATAGGATCCATGGCTGCACAGTCAGCGCCG-3' and 5'-TATACTCGAGTCAATCATTGTTTCGGGATCTGT-3' for the construction of HB-FOP-pcDNA5/FRT, or 5'-TATAGGTACCGCCACCATTGGACTACAAGGACGACGACGACAAGGGATCCATGGCTGCACAGTCAGCGCCG-3' and 5'-TATACTCGAGTCAATCATTGTTTCGGGATCTGT-3' for that of FLAG-FOP-pcDNA5/FRT/TO. The PCR products were cloned into the BamH I/Xho I site of HB-SMN1-F-pcDNA5/FRT for HB-FOP-pcDNA5/FRT, or Kpn I/Xho I site of pcDNA5/FRT/TO (Invitrogen) for FLAG-FOP-pcDNA5/FRT/TO. All cloned cDNAs were verified by DNA sequencing.

Construction of cell lines expressing HB-SMN-For HB-FOP

The cell lines expressing HB-SMN-For HB-FOP were established using an Flp-In T-REx Expression System (Invitrogen). Flp-In T-REx 293 (293TRex) cells were cultured in a well of a 24-well plate, and at an approximately 70% confluency, 293TRex was transfected with 1 µl Lipofectamine2000, containing 0.25 µg pOG44 (Invitrogen), and 0.25 µg of HB-SMN1-F-pcDNA5/FRT, HB-FOP-pcDNA5/FRT, or FLAG-FOP-pcDNA5/FRT/TO. The clonal cells were selected in the culture medium containing 100 µg/ml hygromycin B (Invitrogen).

Cellular fractionation and preparation of nuclear extract

Sub-confluent 293TRex cells expressing HB-SMN-F or HB-FOP stably, 293EBNA cells, or 293T cells were collected, lysed with buffer A (16.7 mM Tris-HCl (pH7.4), 50 mM NaCl, 1.67 mM MgCl₂, 1 mM PMSF, 2 µg/ml aprotinin, 2 µg/ml pepstatin A and 2 µg/ml leupeptin) containing 0.1% Triton X-100, incubated for 5 min on ice, and centrifuged at 1,000×g for 5 min. The supernatant was collected as the cytoplasmic extract (Cyto). To fully remove the cytoplasmic constituents, the pellet was suspended again with buffer A and centrifuged at 1,000×g for 5min. The supernatant was collected as the cytoplasmic wash fraction (Wash), and the pellet was collected as the nuclear pellet. To prepare the nuclear extract under low salt condition, the nuclear pellet was lysed with buffer B (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM MgCl₂, 1 mM PMSF, 2 µg/ml aprotinin, 2 µg/ml pepstatin A and 2 µg/ml leupeptin) containing 0.5% IGEPAL CA-630 for 10 min at 4°C and centrifuged for 20 min

at 20,000 ×g. The supernatant was used as the nuclear extract under the low salt condition (Nuc: Low salt). To prepare the nuclear extract under high salt condition, the nuclear pellet was lysed with buffer C (50 mM Tris-HCl pH 7.4, 400 mM NaCl, 5 mM MgCl₂, 1 mM PMSF, 2 µg/ml aprotinin, 2 µg/ml pepstatin A and 2 µg/ml leupeptin) containing 0.5% IGEPAL CA-630 for 10 min at 4°C and centrifuged for 20 min at 20,000×g. The supernatant was used as the nuclear extract under the high salt condition (Nuc: high salt). To prepare total cell extract, the collected cells were lysed with 1 x SDS sample buffer, sonicated twice for 30 sec using a Bioruptor (Cosmo Bio, Tokyo, Japan) at its highest setting, and then centrifuged at 20,000×g for 30 min. The supernatant was used as total cell extract (TCE). An equal aliquot of each extract was subjected to western blot analysis. The nuclear extract under the high salt condition was subjected to the following biotin affinity pull down or immune precipitation.

Western blot analysis

Proteins were separated by SDS-PAGE and electrophoretically transferred to a PVDF membrane (Millipore, Billerica, MA). The membrane was blocked with 5% non-fat dried skim milk in TBS for 1 h at 25°C, and incubated with the appropriate primary antibody at 4°C overnight. After washing with TBS containing 0.1% (w/v) Tween 20 (TBST) for 10min three times, the membranes were incubated with a secondary antibody conjugated with horseradish peroxidase (HRP) for 1 h, and then washed with TBST for 10 min three times. To detect biotin-tag of HB-SMN-F or HB-FOP, HRP-conjugated streptavidin (Thermo Scientific) was used. Chemiluminescent-stained proteins were detected by using LAS4000 image analyzer (Fujifilm).

Biotin affinity purification method

The nuclear extract under the high salt condition was prepared as described in Cellular fractionation and preparation of nuclear extract. 4 mg of the nuclear extract was incubated with 30 µl of Neutr Avidin Agarose Resins (Thermo Scientific) for 4 h at 4°C and washed five times with 1 ml of buffer C. HB-tagged protein complexes were eluted with 150 µl of Protein-RNA extraction buffer (7 M Urea, 350 mM NaCl, 1% SDS, 10 mM Tris-HCl(pH8.0), 10 mM EDTA, 2% 2-mercaptoethanol) for 5 min. The eluted protein complexes containing RNAs were separated from Neutr Avidin Agarose Resins by centrifugation at 1,000×g for 5 min at 4°C, and subjected to Phenol-chloroform extraction; namely, 150 µl of the eluates were added to 150 µl of Phenol-chloroform (pH8.0) (Ambion) and 15 µl of 3 M sodium acetate (pH5.2), and after vigorous mixing for 20 min at 25°C, the organic phase containing proteins and the aqueous phase containing RNAs were separated by centrifugation at 20,000×g for 20 min at 4°C. Proteins from organic phase were precipitated with 700 µl of isopropanol and following centrifugation at 20,000×g for 10 min at 4°C. RNAs from aqueous phase were collected by isopropanol precipitation. The precipitated proteins or RNAs were washed independently with 75% ethanol, dried up, and subjected to SDS-PAGE or denaturing urea PAGE, respectively. For liquid chromatography (LC)-tandem mass spectrometry (MS/MS) analysis of the HB-FOP complex, the precipitated proteins were dissolved with 5 µl of dissolution buffer (8M Urea, 50 mM Tris-HCl (pH 8.0), and 1% 2-mercaptoethanol), mixed well, diluted to ten times with 50 mM of ammonium hydrogen carbonate and 5 mM of calcium chloride, and subjected to the digestion with 200 ng of trypsin for at 37°C for 6 hr. The resulting trypsin digest solution containing peptides adding 0.1% formic acid was analyzed using a nanoscale LC-MS/MS system.

Protein identification by LC-MS/MS

Pulled-down proteins were separated on a gel by SDS-PAGE; the gel was cut into some pieces, each of which was subjected to in-gel digestion with trypsin [30], and the resulting peptides were analyzed using a nanoflow LC-MS/MS system with quadrupole-time-of-flight 2 hybrid mass spectrometer (Q-Tof2, Micromass, Wythenshawe, UK) as described [31,32]. The peptide mixture was applied to a Mightysil-RP-18 (3 µm particle, Kanto Chemical, Osaka, Japan) fritless column (45 mm×0.150 mm i.d.) and separated using a 0 - 40% gradient of acetonitrile containing 0.1% formic acid over 80 min at a flow rate of 50 nl/min. Eluted peptides were sprayed directly into Q-Tof2. The peptides were detected in the MS mode to select a set of precursor ions for a data-dependent, collision-induced dissociation (CID) MS/MS analysis, and every 4 s the four ions having the greatest signal intensity were subjected to the MS/MS analysis. The MS/MS signals were acquired by MassLynx version 3.5 (Micromass, UK) and converted to text files by ProteinLynx software (Micromass). The database search was performed by MASCOT version 2.3.2 (Matrix Science Ltd., London, UK) against the UniProt human databases version SwissProt_2013_08, 540732 sequences, with the following parameters: fixed modification: carbamidemethyl (Cys), variable modifications: oxidation (Met), N-acetylation, pyroglutamine; maximum missed cleavages: 2; peptide mass tolerance: 200 ppm; MS/MS tolerance: 0.5 Da. The criteria were based on the vendor's definitions (Matrix Science, Ltd.). Furthermore, we set more strict criteria for protein assignment: more than two peptides with p value<0.05 were considered as 'hit'.

Denaturing urea PAGE of RNAs

The dried up RNA from the pulled-down protein complexes was resolved well in 100% formamide, and heated at 65°C for 5 min and cooled immediately into ice. Then one-fourth of volume of 50 mM EDTA (pH8.0) containing loading dye was added to RNAs in formamide and used as a RNA sample. RNA was separated on 8% (w/v) polyacrylamide gels containing 8 M urea and 0.5×TBE (45 mM Tris, 32.3 mM boric acid, 1.25 mM EDTA, pH 8.3) with 0.5×TBE as a running buffer, stained with SYBR Gold (Invitrogen, Carlsbad, CA) for 10 min, and visualized with LAS4000 image analyzer.

In gel RNase digestion for LC-MS/MS analysis

In gel RNase digestion was done principally by the method described [33]. Briefly, gel pieces containing RNA bands were excised from gels, cut into small pieces, and dried under vacuum. The gel pieces were digested with 15 µl of 2 ng/µl RNase T1, with incubation at 37°C for 1 hr. The nucleolytic fragments were extracted from the gel using 100 µl of RNase-free water, passed through a centrifugal filter unit with a polyvinylidene fluoride membrane (Ultrafree-MC, Millipore, Billerica, MA), and then 5 µl of 2 M triethylammonium acetate (pH 7.0) was added before LC-MS analysis.

LC-MS/MS Apparatus for RNA Analysis

The LC system used was essentially as described [34], consisting of a nanoflow pump (LC Assist, Tokyo, Japan) that delivers solvent to the fritless spray tip electrospray ionization column and a ReNCon gradient device. The column was prepared with a fused-silica capillary (150 µm i.d.×375 µm.o.d.) using a laser puller (Sutter Instruments Co., Novato, CA) and was slurry-packed with reversed-phase material (Develosil C30-UG-3, particle size 3 µm, Nomura Chemical, Aichi, Japan) to a length of 50 mm. High voltage for ionization in negative mode (-1.4 kV) was applied to the metal union, and the eluate from RPLC was

sprayed on-line to an LTQ-Orbitrap hybrid mass spectrometer (model XL, Thermo Fisher Scientific, San Jose, CA). Reverse phase separation of oligoribonucleotides was performed at flow rate of 100 nl/min using a 30-min linear gradient from 10 to 40% methanol in 10 mM triethylammonium acetate (pH 7.0). The mass spectrometer was operated in a data-dependent mode to automatically switch between Orbitrap-MS and linear ion trap-MS/MS acquisition as described [35]. Survey full scan MS spectra (from m/z 500 to 1,500) were acquired in the Orbitrap with resolution R=30,000 (after accumulation to a target value of 500,000 ions in the linear ion trap). The most intense ions (up to four, depending on signal intensity) were sequentially isolated for fragmentation in the linear ion trap using CID at a target value of 30,000 ions. An MS scan was accumulated for 2 s and an MS/MS scan for 3 s. The resulting fragment ions were recorded in the linear ion trap with a high-scan rate 'Enhanced' mode. Target ions selected for MS/MS were dynamically excluded for 60 s. General mass spectrometric conditions were as follows: electrospray voltage, 1.4 kV; no lock mass option; with sheath and auxiliary gas flow; normalized collision energy, 35% for MS/MS. Ion selection threshold was 10,000 counts for MS/MS. An activation q-value of 0.25 and activation time of 30 ms were applied for MS/MS acquisitions. LC was performed at a flow rate of 100 nl/min using a 30-min linear gradient from 10% to 40% methanol in 10 mM triethylammonium acetate (pH 7.0). The mass spectrometer was operated in a mode to automatically switch between Orbitrap-MS and linear ion trap-MS/MS acquisition as described [35].

Database search and interpretation of MS/MS Spectra of RNA

We used the software, Ariadne [36], for database searches for RNA identification. Ariadne searched DNA/RNA sequence database using MS/MS data of an RNase T1 digest of an RNA sample. Databases used were an in-house small RNA database, which was constructed from the cDNA of less than 1000 bases in NCBI GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), and EMBL sequence database (<http://www.ebi.ac.uk/embl/>), which contained more than 1,500 entries, and the genome database of Homo Sapiens (reference assembly GRCh37 obtained from ftp://ftp.ncbi.nlm.nih.gov/genomes/H_sapiens/). The following search parameters were used: the maximum number of missed cleavages was set at 1; the maximum number of variable modifications for a nucleotide was 1; The precursor mass tolerance of ± 20 ppm and MS/MS tolerance of ± 750 ppm were allowed; The significance thresholds for both the nucleotide and RNA identifications were set at 5%; The nucleotide lengths considered by the software were from 4 to 19. The original MS/MS spectra for identified nucleotides by Ariadne searches were carefully inspected by at least two MS specialists; Nucleotides with less than 50% sequence coverage were eliminated from the identifications; the length of an identified RNA estimated by PAGE and that from the database was used to confirm RNA identification. After the RNA identification, unidentified MS/MS spectra were manually inspected to identify unidentified nucleotides of the RNA such as those having lengths shorter than 3 nt or longer than 20 nt. The manual identification criteria were as follows: mass differences for nucleotides were within 20 ppm; product ions should cover more than half of the sequence of a candidate nucleotide.

Immunoprecipitation of endogenous SMN complexes

Mouse monoclonal anti-SMN antibody bound to Dynabeads were prepared by incubating 3 μ g of the antibody with 15 μ l of protein G Dynabeads (Invitrogen, Carlsbad, CA) in PBS containing 0.01% Triton X-100 for 30 min at 25°C, and were washed with PBS containing

0.01% Triton X-100 three times. The antibody-bound Dynabeads were incubated with 2 mg of the nuclear extract from 293EBNA cells for 4 h at 4°C and washed five times with 1 ml of buffer C. SMN binding proteins were eluted by 1X SDS sample buffer, and subjected to the following western blot analysis. For SILAC analysis, anti-SMN bound beads was made using Dynabeads Co-Immunoprecipitation Kit (Invitrogen) according to the manufacturer's instructions. Briefly, 20 μ g of anti-SMN antibody was bound to 5 mg of Dynabeads, Epoxy M-270. FLAG-tagged protein was pulled down using 15 μ l of anti-FLAG M2 agarose beads (SIGMA). The cytoplasmic extract or nuclear extract was incubated with FLAG-M2 affinity Agarose Resins (SIGMA) for 2 h at 4°C. After washing five times with buffer B or buffer C, FLAG-protein complexes were eluted with 500 μ g/ml of FLAG peptide or 1X SDS sample buffer, and subjected to the following SDS-PAGE and silver staining or western blot analysis.

Northern blot analysis

The separated RNAs by denaturing urea PAGE were transferred to a Hybond N+ membrane. The membrane was dried and UV cross-linked using UV-cross linker (FUNAKOSHI) at 120 mJ/cm², and subsequently hybridized to biotin-labeled DNA probes at 42°C overnight in Perfect Hyb Plus hybridization buffer (Sigma) according to the manufacturer's instructions. The hybridized membrane was washed sequentially with 2X SSC containing 0.1% SDS for 5 min at 5°C, 0.5X SSC containing 0.1% SDS for 20 min at 42°C, and 0.1X SSC containing 0.1% SDS for 20 min at 25°C. The hybridized RNA was detected using a Chemiluminescent Nucleic Acid Detection Module kit (Thermo Fisher Scientific) according to the manufacturer's instructions. RNAs were detected by using LAS4000 image analyzer and quantified by MultiGauge software. The sequence of oligonucleotides used as probes were noted; 5'-GTATCTCCCTGCCAGGTAAGTAT-3' for U1 probe, 5'-TACTGCAATACCAGGTCGATGCGT-3' for U2 probe, 5'-ATCATCAATGGCTGACGGCAGTTG-3' for U3 probe, 5'-GACTATATTGCAAGTCGTCACGGC-3' for U4 probe, 5'-GACTCAGAGTTATTCCTCTCCACG-3' for U5 probe, and 5'-ACGAATTTGCGTGTATCCTTGCG-3' for U6 probe.

Ribonuclease treatment of the pulled-down protein complexes

The pulled-down complexes bound to the agarose resins described above were incubated with 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM MgCl₂ containing 10 μ g/ml RNase A for 30 min at 37°C, washed three times with 1 ml of buffer C and eluted with Protein-RNA extraction buffer. The eluted proteins and RNAs were separated as described above, and subjected to the following western blot analysis and northern blot analysis, respectively.

Immunofluorescence staining

The FLAG-FOP inducibly expressed cells were grown on collagen-coated culture slides and FLAG-FOP was expressed by the treatment with 100 ng/ml of doxycycline for 24 hr. To visualize nuclei, the cells were washed with PBS followed by permeabilization with PBS containing 0.1% (w/v) Triton X-100 for 10 min at room temperature. After washing twice with PBST, the cells were fixed with 4% paraformaldehyde in PBS for 10 min at room temperature and washed twice with PBST. The cells were then blocked with 3% (w/v) nonfat dried milk in PBS for 1 h, and incubated with anti-FLAG rabbit polyclonal antibody (SIGMA) and anti-SMN antibody (2B1) as the primary antibodies for more than 2 h. The cells were washed three

times in PBST for 10 min and then incubated with FITC-conjugated anti-rabbit IgG goat antibody (American Qualex) and Cy-conjugated anti-mouse IgG sheep antibody (SIGMA) as the secondary antibodies for 1 h at room temperature. The cells were washed three times in PBST for 10 min and then mounted with VECTASHIELD (Vector Laboratories) containing 4',6-diamidino-2-phenylindole (DAPI). The resulting cells were examined with an Axiovert 200 M microscope (Carl Zeiss, Germany).

RNA interference

293T (2×10^6 cells) were cultured in 60-mm Petri dishes until they reached 70% confluency. 293T cells were transfected with 15 μ l of Lipofectamine RNAiMax and 300 pmol of stealth small interfering RNA (siRNA) or control scramble RNA (scRNA).

The following stealth RNAs were used to knockdown the transcripts of FOP as noted: 5'-AUACCUACGAACCGGUUUGUAGC-3' and 5'-GCUAACAACCGGUUCGUAAGGUAU-3' for scRNA, and 5'-AAUCAUUGGUUCGGAUCUGUCUG-3' and 5'-AGACAGAUCCCGAAACCAAUGAUU-3' for siRNA.

Quantitative MS with SILAC

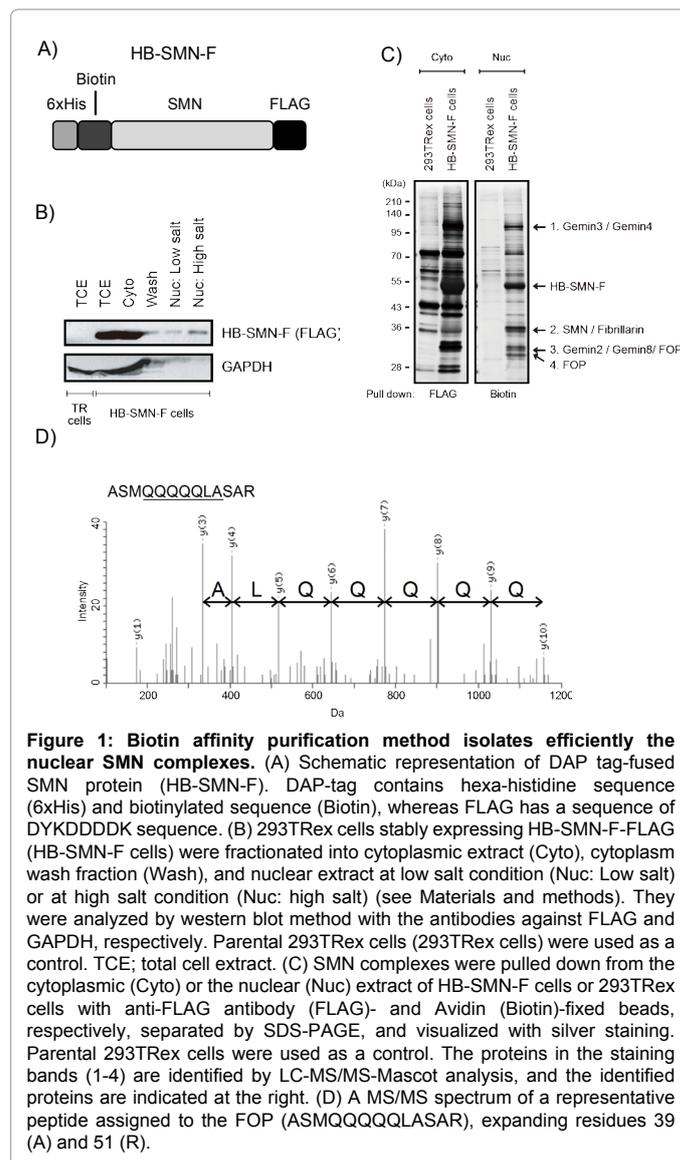
293T cells were grown in SILAC DMEM (Thermo Scientific) supplemented with either "light" $L\text{-}^{12}\text{C}_6$ lysine-2HCl (L-Lys) and $L\text{-}^{12}\text{C}_6$, $^{14}\text{N}_4$ arginine-HCl (L-Arg) or "heavy" $L\text{-}^{13}\text{C}_6$ lysine-2HCl (H-Lys) and $L\text{-}^{13}\text{C}_6$, $^{15}\text{N}_4$ arginine-HCl (H-Arg). For full incorporation of the "light" or "heavy" amino acids, cells were grown for at least six passages. Subsequently, 2×10^6 cells cultured in DMEM with L-Lys and L-Arg, or H-Lys and H-Arg were transfected with FOP scRNA and siRNAs described above, respectively. After being cultured in 96 hour, each of the si/scRNA transfected cells was harvested and immunoprecipitated with anti-SMN antibody as described above. Their immunoprecipitates were eluted with 1x SDS sample buffer, respectively, combined and subjected to SDS-PAGE followed by staining with colloidal blue staining kit (Invitrogen). For quantitative MS, the gel stained with colloidal blue was cut into 16 pieces, each of which was subjected to in-gel digestion with trypsin and the resulting peptides were analyzed using a nanoflow LC-MS/MS system with LTQ-Orbitrap hybrid mass spectrometer (model XL, Thermo Fisher Scientific, CA, USA). The mass spectrometer was operated in a data-dependent mode to automatically switch between Orbitrap-MS and linear ion trap-MS/MS acquisition. Survey full scan mass spectra (from m/z 450 to 1500) acquired in the Orbitrap with resolution $r=15,000$ (an AGC target value of 500,000 counts). The most intense ions (top ten, depending on signal intensity) were sequentially isolated for fragmentation. The resulting fragment ions were recorded in the linear ion trap with a normal mode (Scan Rate, 16666Da/sec). General mass spectrometric conditions were as follows: electrospray voltage, 1.6 kV; normalized collision energy, 35% for MS/MS. Ion selection threshold was 10,000 counts for MS/MS. The resulting raw files were processed and analyzed using Proteome Discoverer version 1.3 (Thermo Fisher Scientific) with default parameters, except for the labels where lysine 6 and arginine 10 were selected. Searches were conducted within MASCOT version 2.3.2 against UniProt human database version SwissProt_2013_08, 540732 sequences and a reverse database, with the following parameters: fixed modification; carbamidemethyl (Cys), variable modifications: oxidation (Met), N-acetylation, pyroglutamine, K+6 and R+10; maximum missed cleavages: 2; peptide mass tolerance 15 ppm; MS/MS tolerance: 0.8 Da. We selected the candidate peptides with probability-based Mowse scores that exceed its threshold indicating a significant

homology ($p < 0.05$), and referred to them as "hits." The criteria were based on the vendor's definitions. Furthermore, we set more strict criteria for protein assignment: more than two peptides with p value < 0.05 were considered as 'hit'. Protein ratios were calculated based on vendor's default definition. Only protein groups with at least three SILAC counts were considered for further analysis.

Results

The nuclear SMN associates with Friend of Prmt1 (FOP)

Although the roles and the components of SMN complexes present in the cytoplasm have been well established, those in the nucleus are unknown except that coilin and fibrillarin were associated with SMN in the nucleus [12,13,37]. We, therefore, tried to isolate the SMN complexes present in the nucleus, and first constructed 293TRex cell line, which has a single-copy transgene encoding HB-SMN-F [29] (Figure 1A). The cytomegalovirus promoter was used to drive stable expression of the transgene, which was integrated at a common locus in the 293TRex genome [38]. As most SMN complex population



existed in the cytoplasm, we first separated the HB-SMN-F expressing 293TRex cells into the cytoplasmic and the nuclear fractions by cell fractionation method with the lysis buffer containing 150 mM NaCl (low salt buffer). Under this low salt condition, about 80% of the SMN population in the total cell extract was recovered as the SMN complexes from the cytoplasmic fraction by a pull down method using anti-FLAG M2 antibody beads (Figures 1B and 1C) [30-32], while we detected only a small amount of the nuclear HB-SMN-F (Figure 1B). To improve the yield of the nuclear SMN, we used lysis buffer containing 400 mM NaCl (high salt buffer), and found that this high salt buffer increased the yield of the nuclear HB-SMN-F about three folds of that extracted with low salt buffer (Figure 1B). Our initial attempts to extract the nuclear SMN complex from anti-FLAG M2 antibody beads with the high salt buffer were unsuccessful (data not shown), probably because the interaction of anti-FLAG M2 antibody and the FLAG epitope was inhibited under high salt condition. Therefore, the biotin affinity purification procedure presented here is essential for efficient recoveries of the nuclear HB-SMN-F and its associated proteins (Figures 1B and 1C). The subsequent MS-based protein identification showed that the isolated nuclear SMN complexes contained Gemins 2, 3 (DDX20), 4, and 8, endogenous SMN, and fibrillarlin that is reported to associate with SMN (Figure 1C and Supplementary Table 1) [39]. In addition, the MS-based method identified Friend of Prmt1, FOP, with the protein score 113 and matched sequence 2, as a potential component of the nuclear SMN complexes (Figure 1D and supplementary Table 1).

FOP is a component of the nuclear SMN complexes that lackSmB/B' and SmD1

We validated the results of MS-based analyses by western blotting; the method clearly detected Gemins 2, 3, 4, and 8, endogenous SMN, fibrillarlin, and FOP (Figure 2A). The method also detected Gemin6 and unrip, the known components of SMN complexes, as well as coilin, which is a marker of Cajal/Gems in the nuclear HB-SMN-F complexes (Figure 2A). Interestingly, SmB/B' and SmD were absent in the isolated nuclear SMN complexes (Figure 2A). We considered that FOP is a novel *de facto* component of the SMN complexes present specifically in the nucleus for the additional evidences: (1) the endogenous nuclear SMN complexes prepared by immunoprecipitation using anti-SMN antibody contained FOP, as well as the several known components, Gemins 2, 3, 4, and endogenous SMN (Figure 2B); (2) the SMN complexes isolated from the nuclear fraction prepared by centrifugal cell fractionation technique contained FOP, whereas the equivalent complexes isolated from the cytoplasmic fraction did not (data not shown); (3) the pulled-down protein mixture prepared from the nuclear fraction using HB-SMN-F as affinity bait contained FOP, whereas the equivalent mixture prepared from the cytoplasmic fraction did not (Figure 2C).

We also recovered several small RNA components from the nuclear SMN complexes isolated under the high salt condition (Figure 2D), and identified the RNAs by the LC-MS/MS system equipped with a genome-oriented database searching engine Ariadne [33,35,36]. This analysis identified a number of known RNA components of the SMN complexes, U1, U3, and U4 snRNAs (Figure 2D, supplementary Table 2), showing that our preparation of the nuclear SMN complexes in fact contained UsnRNAs. However, the result suggest that the UsnRNAs are present in only small population of the isolated nuclear SMN complexes because the nuclear SMN complexes lack SmB/B' and SmD (Figure 2A).

The FOP-associating SMN complexes do not contain U1 and U2 snRNAs

To confirm the association of FOP with the nuclear SMN complexes further, we performed a reverse pulldown-assay using FOP associating protein(s) as bait(s). We prepared the high salt nuclear extract of 293TRex cells expressing HB-FOP, where the DNA fragment-encoding FOP fused to HB-tag towards amino-terminus was integrated into the 293TRex genome (Figure 3A). Like the nuclear SMN, the expressed HB-FOP was extracted more efficiently under the high salt condition than the low salt condition (Figure 3B). This is consistent with the report of van Dijk et al. [26] which demonstrated that FOP was associated with chromatin, localized mainly in facultative heterochromatin region, and thus was extracted mostly under high salt conditions. Biotin affinity

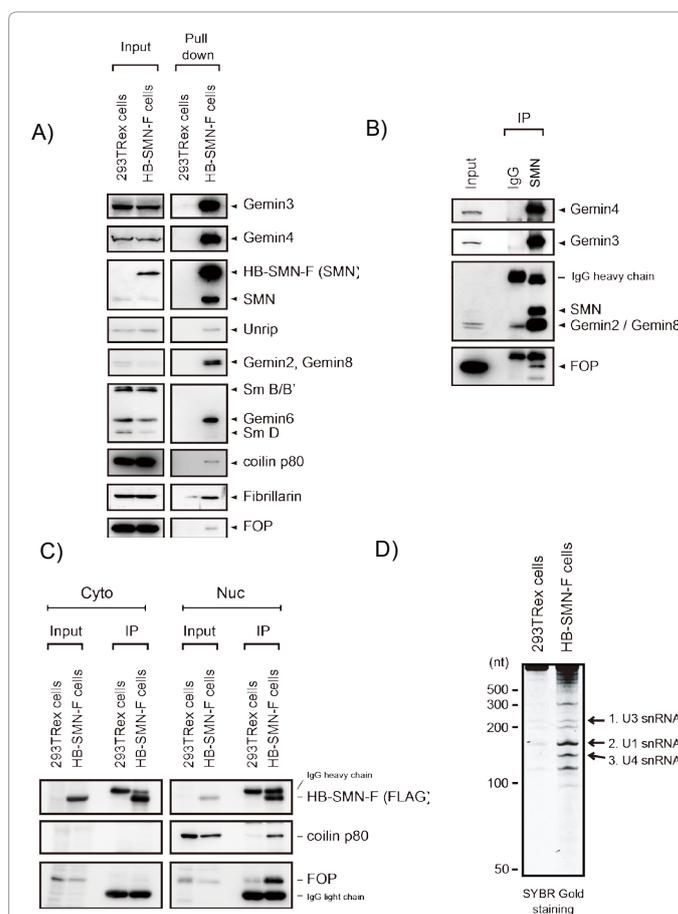


Figure 2: The nuclear SMN complexes isolated contain FOP but not SmB/B' and SmD. (A) The nuclear SMN complexes were pulled down from the nuclear extract of parental 293TRex cells or from 293TRex cells expressing HB-SMN-F (HB-SMN-F cells) by biotin affinity method, and analyzed by western blotting with the antibodies indicated at the right. (B) Endogenous SMN complexes were immunoprecipitated from the nuclear extract of 293TRex cells with anti-SMN antibody or control IgG(IgG), and analyzed by western blotting with the antibodies indicated at the right. A staining band corresponding to that of IgG heavy chain is also indicated. (C) SMN complexes were immunoprecipitated with anti-FLAG M2 antibody fixed beads from cytoplasmic or nuclear extract of HB-SMN-F or parental 293TRex cells, and analyzed by western blotting with the antibodies indicated at the right. A staining band corresponding to that of IgG light chain is also indicated. (D) RNAs extracted from the SMN complexes isolated from the nuclear extract of HB-SMN-F cells or 293TRex cells were separated by denaturing urea PAGE and stained with SYBR gold. The RNAs in the staining bands (1-3) are identified by LC-MS/MS-Ariadne analysis, and the identified RNAs are indicated at the right.

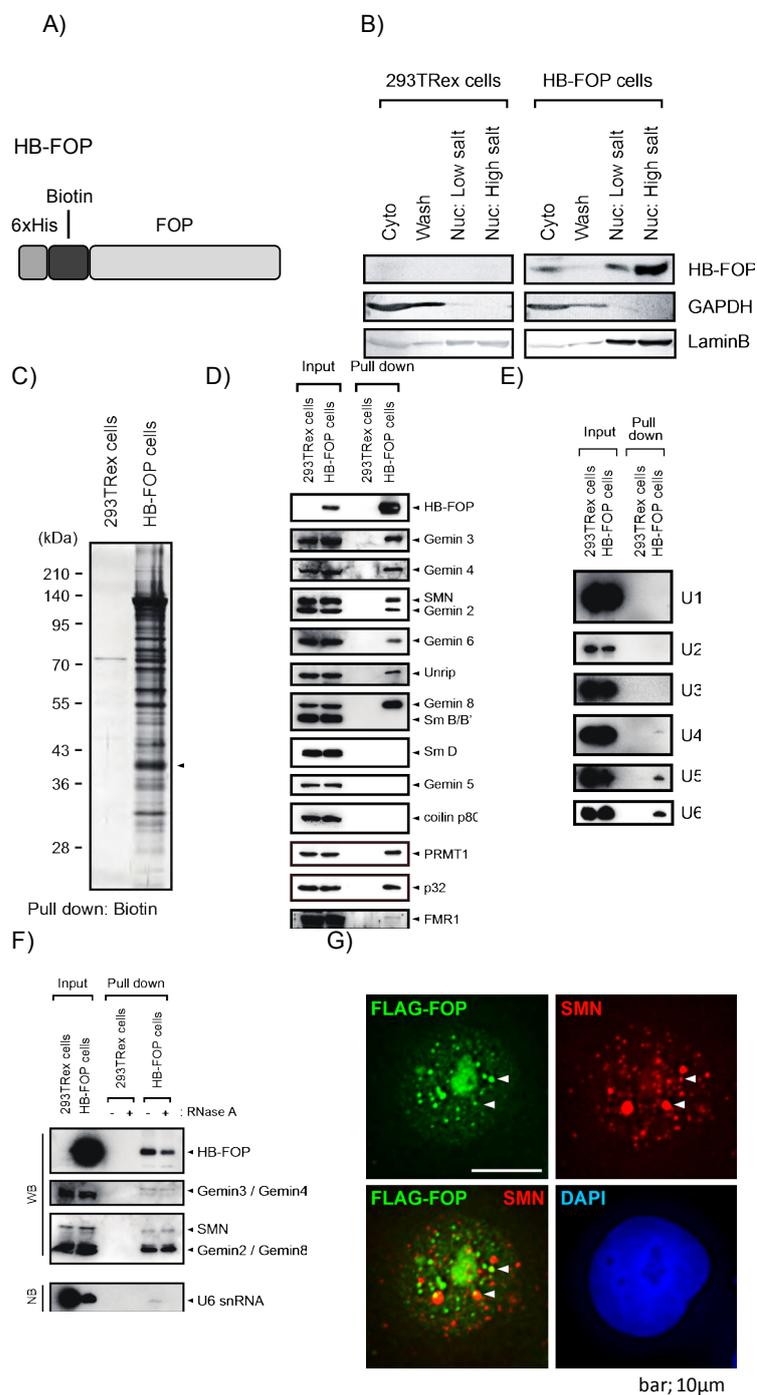


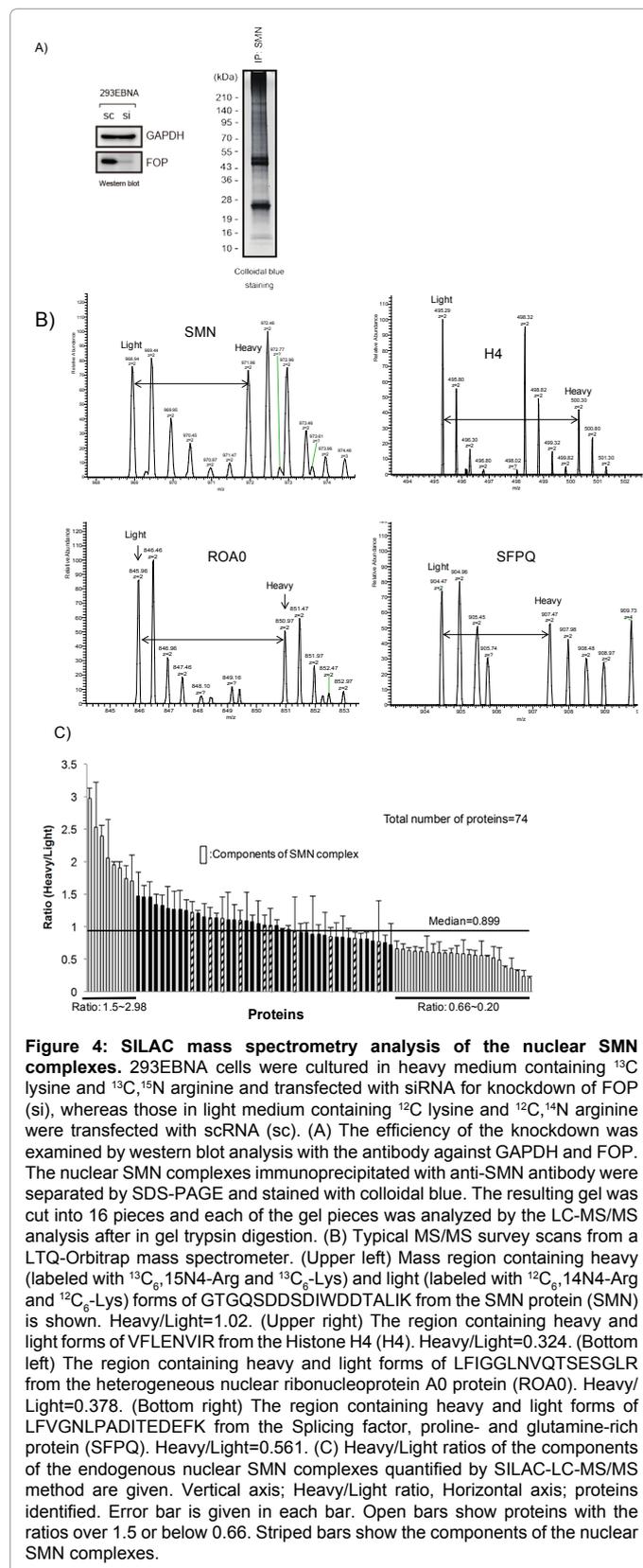
Figure 3: FOP-associated nuclear SMN complex lacks Gemin5 and coilin as well as U1 and U2 snRNAs. (A) Schematic representation of FOP tagged with DAP lacking FLAG (HB-FOP). (B) 293TRex cells stably expressing HB-FOP (HB-FOP cells) were fractionated into cytoplasmic extract (Cyto), cytoplasm wash fraction (Wash), and nuclear extract at low salt condition (Nuc: Low salt) or at high salt condition (Nuc: high salt) (see Material and methods). They were analyzed by western blot method with the antibodies against Lamin B and GAPDH, respectively. HB-FOP was detected with HRP-conjugated streptavidin. Parental 293TRex cells were used as a control (293TRex cells). (C) HB-FOP complex was pulled down from the nuclear extract of HB-FOP cells or 293TRex cells with avidin-fixed beads (Biotin), separated by SDS-PAGE and visualized by silver staining. The protein band of HB-FOP was assigned based on its mobility expected from the amino acid sequence of HB-FOP on SDS-PAGE gel. Molecular weights of the marker proteins are indicated at left. (D) The pulled-down HB-FOP complex were analyzed by western blot with the antibodies indicated at the right. HB-FOP was detected with HRP-conjugated streptavidin. Total RNA (1 μg) was used as input. (E) The RNAs extracted from the HB-FOP complex were analyzed by northern blot with the biotin labeled DNA probes complementary to U1-U6 sn(0)RNAs, respectively. (F) The pulled-down HB-FOP complex was analyzed by western blotting with the antibodies against proteins indicated at the right with (+) or without (-) RNase A treatment. U6 snRNA was detected by northern blot analysis. (G) 293TRex cells inducibly expressing FLAG-fused-FOP (FLAG-FOP cells) were subjected to immunocytochemical analysis with the antibody against FLAG (green) or endogenous SMN (red) after the removal of the cytoplasmic proteins by permeabilization with Triton. DAPI staining (blue) shows the nucleus. Scale bar: 10 μm.

purification and MS-based analyses showed that HB-FOP was associated with SMN, Gemins 2, 3, 4, and 8, and unrip (Figure 3C and Table 1 and Supplementary Table 3). The analyses also identified Fragile X mental retardation 1 protein (FMR1), which was reported to associate with SMN [40] (Table 1, and supplementary Table 3). The other proteins identified include protein arginine N-methyltransferases 1 (Prmt1) and 5 (Prmt5) both of which were reported to be the FOP binding proteins (26), the complement component 1 Q subcomponent-binding protein (p32), lamin-B receptor, protein arginine U5 snRNP 200 kDa helicase, 116 kDa U5 snRNP component, heterogeneous nuclear RNP (hnRNP) C1/C2, hnRNP H, nucleolar RNA helicase 2, ATP-dependent RNA helicase DDX39 (BAT1-associated protein), Cip1-interacting zinc finger protein, zinc finger CCCH domain-containing protein 14 and others (supplementary Table 1 and Supplementary Table 3). Western blot analysis using available antibodies confirmed minimally the presence of SMN, Gemins 2, 3, 4, 6, and 8, unrip, p32, Prmt1, and FMR1 in the HB-FOP associating complexes (Figure 3D). The isolated HB-FOP associating complexes contained neither Sm B/B' and Sm D, nor Gemin5 and coilin (Figure 3D). HB-FOP was associated with U4, U5, and U6 snRNAs but not with U1, U2, and U3 snRNAs as revealed by northern blot analysis (Figure 3E). These data suggest that FOP is associated with SMN complexes that were separated from the major U1 and U2 snRNPs containing Sm proteins in Cajal/Gems. We treated the HB-FOP associating complexes with RNase A and found that SMN and FOP were associated with each other independently of RNA, whereas the RNase treatment reduced U6 snRNA in the complexes (Figure 3F). In addition, we examined whether SMN and FOP were co-localized in the nucleus by immunocytochemical analysis. Because SMN mainly localizes in the cytoplasm, we stained SMN and FOP in the nucleus with their corresponding antibodies after the removal of the cytoplasmic proteins by permeabilization with 0.1% Triton, and showed that the two proteins were co-localized in several foci in the nucleus (Figure 3G), supporting their specific *in vivo* association.

FOP is required for the association of the nuclear SMN complexes with hnRNPs, histones and RNA-binding proteins

To gain insight into the functional relationship between FOP and the nuclear SMN complexes, we examined the effect of FOP knockdown on protein components associated with the nuclear SMN by SILAC MS analysis. We cultured scRNA- and siRNA-treated cells with light and heavy amino acids, and fractionated into the cytoplasmic and nuclear extracts by cell fractionation method, respectively. Each of the nuclear extracts was then subjected to immunoprecipitation with anti-SMN antibody to isolate the SMN-associating complexes (Figure 4A). The 96 h-siRNA treatment reduced FOP by about 10% relative to GAPDH when compared with scRNA treatment (Figure 4B). The immunoprecipitates from scRNA- and siRNA-treated cells were combined, separated by SDS-PAGE, stained with colloidal blue (Figure 4B), and subjected to in-gel protease digestion. Relative ratios of the protein components pulled down with endogenous SMN with and without FOP knockdown were obtained by LC-MS/MS analysis of the digest as the ratios of the mass intensities of the corresponding peptides with light and heavy amino acids [heavy (siRNA)/light (scRNA)] (Figures 4B and 4C and Supplementary Table 4). This quantitative analysis showed that the knockdown of FOP reduced the bindings of SMN to histones, hnRNPs and a series of RNA binding proteins by about 20%~60% of those without knockdown, whereas the contents of the known components of SMN complex remained unchanged (Figures 4B and 4C and Supplementary Table 4). These data suggest

that FOP is required for the association of the nuclear SMN complexes notably with the histones, presumably within the chromatin. They also



UniProt Accession	Protein Name	Gene symbol	MW [Da]	Protein Score	Match of sequences
Bait					
CHTOP_HUMAN	Chromatin target of PRMT1 protein	CHTOP	26380	709	8
SMN binding proteins					
SMN_HUMAN	Survival motor neuron protein	SMN	32285	255	3
GEMI2_HUMAN	Gem-associated protein 2	GEMIN2	32021	102	2
DDX20_HUMAN	Probable ATP-dependent RNA helicase DDX20	DDX20	92981	293	9
GEMI4_HUMAN	Gem-associated protein 4	GEMIN4	121728	207	5
GEMI6_HUMAN	Gem-associated protein 6	GEMIN6	18983	41	1
GEMI8_HUMAN	Gem-associated protein 8	GEMIN8	28904	54	2
STRAP_HUMAN	Serine-threonine kinase receptor-associated protein	STRAP	38756	121	4
FMR1_HUMAN	Fragile X mental retardation protein 1	FMR1	71473	265	9
Protein arginine N-methyltransferases (PRMTs) and related proteins					
ANM1_HUMAN	Protein arginine N-methyltransferase 1	PRMT1	42059	648	8
ANM5_HUMAN	Protein arginine N-methyltransferase 5	PRMT5	73322	90	4
Zinc finger proteins					
KHDR1_HUMAN	KH domain-containing, RNA-binding, signal transduction-associated protein 1	KHDR1	48311	290	7
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14	ZC3HE	83793	210	5
ZC11A_HUMAN	Zinc finger CCCH domain-containing protein 11A	ZC11A	89931	66	2
Transcription related proteins					
CIZ1_HUMAN	Cip1-interacting zinc finger protein	CIZ1	101066	396	8
P53_HUMAN	Cellular tumor antigen p53	TP53	44196	216	6
ERH_HUMAN	Enhancer of rudimentary homolog	ERH	12422	180	3
TR150_HUMAN	Thyroid hormone receptor-associated protein 3	THRAP3	108658	85	2
SAFB2_HUMAN	Scaffold attachment factor B2	SAFB2	107921	78	3
YLPM1_HUMAN	YLP motif-containing protein 1	YLPM1	220077	53	2
MATR3_HUMAN	Matrin-3	MATR3	95078	52	2
snRNA binding proteins					
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component	EFTUD2	110336	105	3
RU17_HUMAN	U1 small nuclear ribonucleoprotein 70 kDa	SNRNP70	51583	74	2
mRNA splicing factors					
C1QBP_HUMAN	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP	31742	591	5
IF4A3_HUMAN	Eukaryotic initiation factor 4A-III	EIF4A3	47126	227	2
PABP2_HUMAN	Polyadenylate-binding protein 2	PABPN1	32843	161	2
SRSF6_HUMAN	Serine/arginine-rich splicing factor 6	SRSF6	39677	74	2
hnRNPs and RNA helicases					
HNRH1_HUMAN	Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	49484	142	3
HNRPC_HUMAN	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRNPC	33707	143	3
HNRPF_HUMAN	Heterogeneous nuclear ribonucleoprotein F	HNRNPF	45985	113	2
DX39A_HUMAN	ATP-dependent RNA helicase DDX39A	DDX39A	49611	78	2
Chaperons & Cytoskeltons					
LBR_HUMAN	Lamin-B receptor	LBR	71057	1134	14
ACTA_HUMAN	Actin, aortic smooth muscle	ACTA2	42381	97	4
HSP71_HUMAN	Heat shock 70 kDa protein 1A/1B	HSPA1A	70294	743	13
HSP7C_HUMAN	Heat shock cognate 71 kDa protein	HSPA8	71082	371	9
Others					
NPM_HUMAN	Nucleophosmin	NPM1	32726	127	3
PNMA2_HUMAN	Paraneoplastic antigen Ma2	PNMA2	41711	117	3
VIME_HUMAN	Vimentin	VIM	53676	54	2

Table 1: Identification of proteins associated with the nuclear HB-FOP by LC-MS/MS analysis. The Nuclear HB-FOP -associated proteins containing SMN complex are indicated. Proteins are classified into functional groups based on SwissProt classification. Proteins were identified by Q-ToF2 as described in Supplementary Tables II. Proteins detected in mock are not listed in this table. UniProt accession, protein name, gene symbol as well as molecular weight are shown. we added GEMIN6 to the identification list by the confirmation of the association of GEMIN6 with HB-FOP using western blot analysis.

suggest that FOP recruits many hnRNP complexes or RNA binding proteins onto the nuclear SMN complexes. On the other hand, the knockdown increased the association of SMN complexes with ATP or GTP binding proteins including Guanine nucleotide-binding protein subunit beta-2-like 1, ADP/ATP translocase 1 etc. (Supplementary Table 4), suggesting that FOP suppresses the binding of those proteins to the SMN complexes.

Discussion

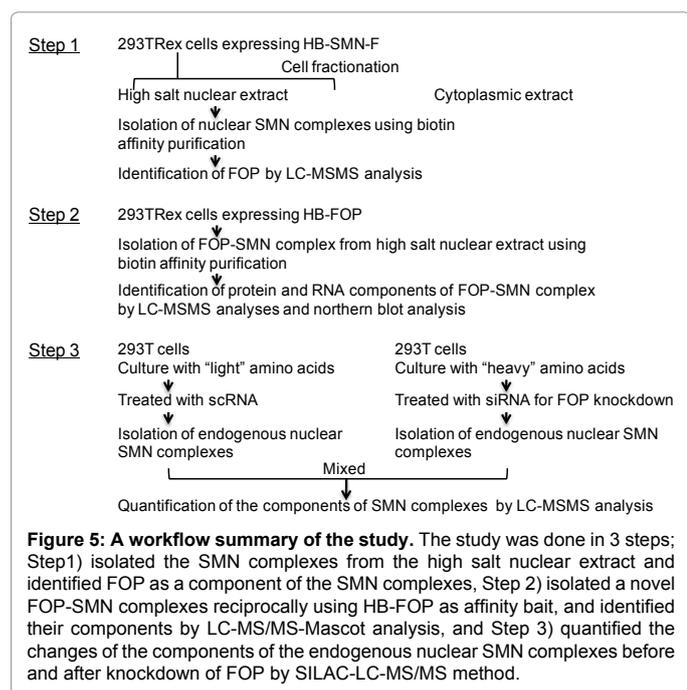
We summarized the strategy of the present study in Figure 5. In this study, we successfully isolated the nuclear SMN complexes by a combination of high salt extraction and biotin affinity purification methods, and found that FOP is a novel component of the nuclear SMN complexes (Step 1 in Figure 5). Our data suggest that the nuclear SMN complexes isolated using HB-SMN-F as affinity bait contain at least two distinct SMN complexes; one contains coilin in addition to Gemins and unrip (Figure 2A), and another contains FOP but does not contain coilin, U1 and U2 snRNAs (Figure 3D and 3E). The former nuclear SMN complex probably corresponds to that present in Cajal/Gems, because the complex contains coilin, which is presumably required to dissociate the SMN complex from U1 and U2 snRNPs in the Cajal/Gems [12-15,17]. Our results also suggest that the use of HB-FOP as affinity bait isolates preferentially the latter SMN complex (FOP-SMN complex) (Step 2 in Figure 5). In addition, the quantitative analysis of the components of the endogenous nuclear SMN complexes before and after the knockdown of FOP suggests that the FOP-SMN complex is retained on chromatin [26] possibly via the association with histones (Step 3 in Figure 5). FOP is also required for the association of SMN complexes with a number of hnRNPs and RNA binding proteins (Step 3 in Figure 5). Based on these data and the previous reports [26], we propose that FOP-SMN complex has a role in retaining spliced and/or unspliced mRNAs in near the transcription sites on chromosome. This does not conflict to the previous notion that FOP has a critical role in specific mRNA transcriptional regulation and mRNA export [27,28]. We do not exclude, however, other roles of FOP-SMN complex, such

that it may prevent improperly processed mRNAs from transit through the nuclear speckle [20,21], or that FOP may keep the nuclear SMN complexes near the transcription sites on chromosome until it is used for recycling UsnRNPs after splicing [22-24].

Thus, for the first time in our knowledge, our present study identifies the FOP-SMN complexes that are separated from the mature U1 and U2 snRNPs in the nucleus, and provides its potential role in post-transcriptional gene regulation; i.e., FOP bridges between the nuclear SMN complexes and the region near the transcription sites on chromosome *via* the association with histones, and possibly retains spliced and/or unspliced mRNAs *via* the association with a number of hnRNPs and RNA-binding proteins. It has been reported that loss of Cajal/Gems is a cellular hallmark of SMA and amyotrophic lateral sclerosis (ALS) [41,42]. Since the formation of FOP-SMN complex is probably related to or coupled with the events taken place in Cajal/Gems, the detailed analysis of FOP-SMN complex may open up a novel molecular mechanism by which a common biochemical pathway leads to SMA and ALS.

References

- Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, et al. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80: 155-165.
- Lefebvre S, Burlet P, Liu Q, Bertrand S, Clermont O, et al. (1997) Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* 16: 265-269.
- Monani UR, Sendtner M, Coovert DD, Parsons DW, Andreassi C, et al. (2000) The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in *Smn*^{-/-} mice and results in a mouse with spinal muscular atrophy. *Hum Mol Genet* 9: 333-339.
- Hahnen E, Forkert R, Marke C, Rudnik-Schöneborn S, Schönling J, et al. (1995) Molecular analysis of candidate genes on chromosome 5q13 in autosomal recessive spinal muscular atrophy: evidence of homozygous deletions of the SMN gene in unaffected individuals. *Hum Mol Genet* 4: 1927-1933.
- Will CL, Lührmann R (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* 13: 290-301.
- Raker VA, Hartmuth K, Kastner B, Lührmann R (1999) Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol Cell Biol* 19: 6554-6565.
- Meister G, Fischer U (2002) Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J* 21: 5853-5863.
- Friesen WJ, Massenet S, Paushkin S, Wyce A, Dreyfuss G (2001) SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol Cell* 7: 1111-1117.
- Hao le T, Fuller HR, Lam le T, Le TT, Burghes AH, et al. (2007) Absence of gemin5 from SMN complexes in nuclear Cajal bodies. *BMC Cell Biol* 8: 28.
- Grimmler M, Otter S, Peter C, Müller F, Chari A, et al. (2005) Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Hum Mol Genet* 14: 3099-3111.
- Carissimi C, Baccon J, Straccia M, Chiarella P, Maiolica A, et al. (2005) Unrip is a component of SMN complexes active in snRNP assembly. *FEBS Lett* 579: 2348-2354.
- Hebert MD, Szymczyk PW, Shpargel KB, Matera AG (2001) Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes Dev* 15: 2720-2729.
- Hebert MD, Shpargel KB, Ospina JK, Tucker KE, Matera AG (2002) Coilin methylation regulates nuclear body formation. *Dev Cell* 3: 329-337.
- Xu H, Pillai RS, Azzouz TN, Shpargel KB, Kambach C, et al. (2005) The C-terminal domain of coilin interacts with Sm proteins and U snRNPs. *Chromosoma* 114: 155-166.



15. Toyota CG, Davis MD, Cosman AM, Hebert MD (2010) Coilin phosphorylation mediates interaction with SMN and SmB'. *Chromosoma* 119: 205-215.
16. Sleeman JE, Lamond AI (1999) Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *CurrBiol* 9: 1065-1074.
17. Yong J, Kasim M, Bachorik JL, Wan L, Dreyfuss G (2010) Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* 38: 551-562.
18. Patel SB, Bellini M (2008) The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res* 36: 6482-6493.
19. Hall LL, Smith KP, Byron M, Lawrence JB (2006) Molecular anatomy of a speckle. *Anat Rec A DiscovMol Cell EvolBiol* 288: 664-675.
20. Johnson C, Primorac D, McKinstry M, McNeil J, Rowe D, et al. (2000) Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. *J Cell Biol* 150: 417-432.
21. Smith KP, Byron M, Johnson C, Xing Y, Lawrence JB (2007) Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type I is blocked at entry into SC-35 domains. *J Cell Biol* 178: 951-964.
22. Morris GE (2008) The Cajal body. *BiochimBiophysActa* 1783: 2108-2115.
23. Stanek D, Pridalová-Hnilicová J, Novotný I, Huranová M, Blazíková M, et al. (2008) Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies. *MolBiol Cell* 19: 2534-2543.
24. Fourmann JB, Schmitzová J, Christian H, Urlaub H, Ficner R, et al. (2013) Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev* 27: 413-428.
25. Zullo AJ, Michaud M, Zhang W, Grusby MJ (2009) Identification of the small protein rich in arginine and glycine (SRAG): a newly identified nucleolar protein that can regulate cell proliferation. *J BiolChem* 284: 12504-12511.
26. vanDijk TB, Gillemans N, Stein C, Fanis P, Demmers J, et al. (2010) Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol* 30: 260-272.
27. vanDijk TB, Gillemans N, Pourfarzad F, van Lom K, von Lindern M, et al. (2010) Fetal globin expression is regulated by Friend of Prmt1. *Blood* 116: 4349-4352.
28. Chang CT, Hautbergue GM, Walsh MJ, Viphakone N, van Dijk TB, et al. (2013) Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J* 32: 473-486.
29. Hayano T, Yamauchi Y, Asano K, Tsujimura T, Hashimoto S, et al. (2008) Automated SPR-LC-MS/MS system for protein interaction analysis. *J Proteome Res* 7: 4183-4190.
30. Izumikawa K, Yanagida M, Hayano T, Tachikawa H, Komatsu W, et al. (2008) Association of human DNA helicase RecQ5beta with RNA polymerase II and its possible role in transcription. *Biochem J* 413: 505-516.
31. Fujiyama-Nakamura S, Yoshikawa H, Homma K, Hayano T, Tsujimura-Takahashi T, et al. (2009) Parvulin (Par14), a peptidyl-prolyl cis-trans isomerase, is a novel rRNA processing factor that evolved in the metazoan lineage. *Mol Cell Proteomics* 8: 1552-1565.
32. Yoshikawa H, Komatsu W, Hayano T, Miura Y, Homma K, et al. (2011) Splicing factor 2-associated protein p32 participates in ribosome biogenesis by regulating the binding of Nop52 and fibrillarin to preribosome particles. *Mol Cell Proteomics* 10:M110.006148.
33. Taoka M, Ikumi M, Nakayama H, Masaki S, Matsuda R, et al. (2010) In-gel digestion for mass spectrometric characterization of RNA from fluorescently stained polyacrylamide gels. *Anal Chem* 82: 7795-7803.
34. Natsume T, Yamauchi Y, Nakayama H, Shinkawa T, Yanagida M, et al. (2002) A direct nanoflow liquid chromatography-tandem mass spectrometry system for interaction proteomics. *Anal Chem* 74: 4725-4733.
35. Taoka M, Yamauchi Y, Nobe Y, Masaki S, Nakayama H, et al. (2009) An analytical platform for mass spectrometry-based identification and chemical analysis of RNA in ribonucleoprotein complexes. *Nucleic Acids Res* 37: e140.
36. Nakayama H, Akiyama M, Taoka M, Yamauchi Y, Nobe Y, et al. (2009) Ariadne: a database search engine for identification and chemical analysis of RNA using tandem mass spectrometry data. *Nucleic Acids Res* 37: e47.
37. Jones KW, Gorzynski K, Hales CM, Fischer U, Badbanchi F, et al. (2001) Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J BiolChem* 276: 38645-38651.
38. Sauer B, Henderson N (1988) Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *ProcNatlAcadSci U S A* 85: 5166-5170.
39. Wehner KA, Ayala L, Kim Y, Young PJ, Hosler BA, et al. (2002) Survival motor neuron protein in the nucleolus of mammalian neurons. *Brain Res* 945: 160-173.
40. Piazzon N, Rage F, Schlotter F, Moine H, Branlant C, et al. (2008) In vitro and in cellulo evidences for association of the survival of motor neuron complex with the fragile X mental retardation protein. *J BiolChem* 283: 5598-5610.
41. Yamazaki T, Chen S, Yu Y, Yan B, Haertlein TC, et al. (2012) FUS-SMN protein interactions link the motor neuron diseases ALS and SMA. *Cell Rep* 2: 799-806.
42. Shan X, Chiang PM, Price DL, Wong PC (2010) Altered distributions of Gemini of coiled bodies and mitochondria in motor neurons of TDP-43 transgenic mice. *ProcNatlAcadSci U S A* 107: 16325-16330.

This article was originally published in a special issue, **Affinity Proteomics** handled by Editor(s). Dr. Peter Nilsson, KTH-Royal Institute of Technology, Sweden