Published online 15 February 2019

Nucleic Acids Research, 2019, Vol. 47, No. 8 ***4181–4197***
doi: 10.1093/nar/gkz099

SAM68 interaction with U1A modulates U1 snRNP recruitment and regulates *mTor* pre-mRNA splicing

Suryasree Subramania[1,2], Laurence M. Gagné[1,2], Sébastien Campagne[3], Victoire Fort[1,2], Julia O'Sullivan[1,2], Karel Mocaer[1], Miki Feldmüller[3], Jean-Yves Masson[1,2], Frédéric H.T. Allain[3], Samer M. Hussein[1,2] and Marc-Étienne Huot[1,2,*]

[1]Centre de recherche du CHU de Québec-Université Laval (axe Oncologie), Québec, QC G1R 3S3, Canada, [2]CRCHU de Québec – Axe Oncologie, Québec, QC G1R 3S3, Canada and [3]Institute of Molecular Biology and Biophysics, Department of Biology, ETH Zurich, CH-8093 Zurich, Switzerland

Received November 17, 2017; Revised January 14, 2019; Editorial Decision February 04, 2019; Accepted February 05, 2019

ABSTRACT

Src associated in mitosis (SAM68) plays major roles in regulating RNA processing events, such as alternative splicing and mRNA translation, implicated in several developmental processes. It was previously shown that SAM68 regulates the alternative splicing of the mechanistic target of rapamycin (*mTor*), but the mechanism regulating this process remains elusive. Here, we report that SAM68 interacts with U1 small nuclear ribonucleoprotein (U1 snRNP) to promote splicing at the 5′ splice site in intron 5 of *mTor*. We also show that this direct interaction is mediated through U1A, a core-component of U1snRNP. SAM68 was found to bind the RRM1 domain of U1A through its C-terminal tyrosine rich region (YY domain). Deletion of the U1A-SAM68 interaction domain or mutation in SAM68-binding sites in intron 5 of *mTor* abrogates U1A recruitment and 5′ splice site recognition by the U1 snRNP, leading to premature intron 5 termination and polyadenylation. Taken together, our results provide the first mechanistic study by which SAM68 modulates alternative splicing decision, by affecting U1 snRNP recruitment at 5′ splice sites.

INTRODUCTION

Alternative splicing is a highly regulated event requiring an impressive amount of ribonucleoprotein complexes and associated factors (1,2). In this process, intervening sequences are excised out from nuclear pre-messenger RNA (pre-mRNA) by the macromolecular machinery called the spliceosome (3). The recognition of the 5′ splice sites by U1 small nuclear ribonucleoprotein (U1 snRNP) defines the initial stages of spliceosome assembly. U1 snRNP along with U2, U4, U5 and U6 snRNPs forms the major spliceosome, the core machinery that catalyzes splicing reactions in eukaryotes (4). Although core spliceosomal assembly and its catalytic activity are rather well defined, an increasing number of accessory spliceosomal proteins modulate its activity and specificity, thereby making alternative splicing a highly regulated process (5). The main challenge for efficient intron splicing is the recognition of the 5′ and 3′ splice sites. This is mainly achieved by U1 snRNP (6,7), U2 snRNP and U2AF (8,9). These spliceosome components drive the assembly of the formation of the early spliceosome called complex E (10,11). Now it is well known that regulatory factors can bind sequences neighboring the 5′ splice site to prevent or promote U1 snRNP binding (12). Increasing evidence highlight the importance of RNA-binding proteins in facilitating U1 snRNP recognition of 5′ splice sites and regulating alternative and constitutive splicing. These include FUS (13,14), SF2 (15,16), TIA-1 (17), RBM24 (18), hnRNPs (19,20) and SAM68 (21–24).

Src associated in mitosis of 68 kDa (SAM68), a 443-amino acid polypeptide, belongs to the signal transduction and activation of RNA family of RNA-binding proteins (RBPs) and was identified as a substrate of phosphorylation by c-SRC during mitosis and cellular transformation (25,26). SAM68 was shown to be able to bind mRNA (27), as well as DNA, upon its methylation (28). The multi-functionality of SAM68 can be rightly attributed to its modular organization. The RNA binding activity of SAM68 is confined to its highly conserved GSG (GRP33/SAM68/GLD-1) domain, comprising of hnRNP K homology (KH) domain flanked on its N terminus by 80 amino acids (NK) and its C-terminus of 30 amino acids (CK), respectively (29,30). It has been demonstrated by X-ray crystallography that the NK region is required for the RNA-dependent homodimerization of SAM68 (31). In addition, SAM68 has six proline rich sequences on either side of GSG domain along with a tyrosine rich C-terminus that were shown to be targeted by various signaling pathways (32–34). The tyrosine phosphorylation of SAM68 as well as its interaction with SH2 binding proteins has been shown to

*To whom correspondence should be addressed. Tel: +418 525 4444(Ext. 14552); Fax: +418 69 5439; Email: marc-etienne.huot@crchudequebec.ulaval.ca

impair its affinity for RNA (23,33). Thus, SAM68 is a versatile adaptor and nucleic acid docking protein whose activity is modulated by cell signaling.

SAM68 is known to bind single-stranded U/A-rich mRNA molecules, mainly through U(U/A)AA repeats (35). The RNA-binding activity of SAM68 was shown to be involved in various aspects of mRNA processing including alternative splicing (29). This was initially shown following ERK1/2 signaling pathway activation, which promoted a SAM68-induced inclusion of the variable exon5 in CD44 (24,33). SAM68 has been involved in the alternative splicing of mRNAs implicated in neurogenesis (36,37), adipogenesis (21,38–40), spermatogenesis (41,42) and epithelial-to-mesenchymal transition (43). SAM68 regulated alternative splicing was further highlighted with *SMN2* (44), *BCL-x* (22), *Cyclin D1* (22) and *mTor* (21) pre-mRNA transcripts. While the mechanisms underlying the splicing of SMN-2, BCL-x and Cyclin D1 are becoming clearer, the mechanism regulating SAM68-induced alternative splicing of *mTor* pre-mRNA remains elusive.

mTOR is a central regulator of cell homeostasis, growth, proliferation and survival (45). Its dysregulation occurs in many human diseases such as cancer, obesity, Type 2 diabetes and neurodegeneration (45,46). Hence, it is crucial to understand the mechanism of SAM68 regulated *mTor* pre-mRNA splicing. Using the *Sam68*$^{-/-}$ mouse as models, we had previously unveiled a novel role of SAM68 in driving alternative splicing of *mTor* pre-mRNA (21). We found that impairing SAM68 binding to its target elements found near the 5′ splice site of intron 5 decreases the expression of full-length *mTor* mRNA by increasing intron 5-induced premature termination leading to the production of a shorter mRNA termed *mTor*$_{i5}$, with no detectable protein product. The production of *mTor*$_{i5}$ is increased in *Sam68*$^{-/-}$ mouse tissues indicating that SAM68 mediates the balance between both isoforms. As a result, *Sam68*$^{-/-}$ mice have decreased mTOR protein levels and attenuated mTORC1 and mTORC2 activities. RNA-binding assays determined that SAM68 binds multiple U/A-rich sequences distributed throughout intron 5 and enhances splicing at the upstream exon/intron junction (21). These observations suggest that SAM68 has the ability to regulate an important *mTor* pre-mRNA alternative splicing checkpoint, though the underlying mechanism remains unknown.

Here, we investigated the mechanism by which SAM68 modulates *mTor* pre-mRNA splicing. First, we found that SAM68 was detected in the immunoprecipitates of the core components of U1 snRNP, namely U1A and U1–70K. Reciprocal immunoprecipitation with Flag-tagged SAM68 showed enrichment of U1 snRNP. Concomitantly, purified recombinant SAM68 can capture U1 snRNP through direct interaction with U1A. Domain mapping experiments revealed that the tyrosine rich C-terminal region of SAM68 (YY domain) was sufficient to interact with U1A. Using endogenous RNA immunoprecipitation assays, we found that SAM68 can recruit U1 snRNP to the 5′ splice site of *mTor* intron 5. Thus, these results provide the first mechanistic insight on how SAM68 regulates *mTor* pre-mRNA alternative splicing and could unveil a broader regulatory function of SAM68-mediated 5′ splice site recognition.

MATERIALS AND METHODS

Plasmid constructions

pGEX-6P3-*SAM68*-Flag and pGEX-6P3-*Sam68*-Flag were constructed by inserting full-length human and mouse *SAM68* coding sequence (cds) into pGEX-6P3 (GE Healthcare) with N-terminal GST tag, a PreScission protease cleavage site (see below 'Protein purification and GST pulldown' section) between GST and SAM68. Due to the high sequence homology between human and mouse *SAM68*, both constructs were produced using EcoRI-*SAM68*-F and NotI-*SAM68*-R. Flag tag was then inserted at the C-terminus of *SAM68* by annealing the oligos, SacI-Flag-F and SacI-Flag-R, and inserting the adaptor at SacI sites of the plasmid. pGEX-6P2-U1A-His was generated by inserting U1A cds, obtained by polymerase chain reaction (PCR) from HEK-293T total cDNA using EcoRI-U1A-F and SalI-U1A-R at EcoRI-SalI sites. cDNA was amplified from total RNA of HEK-293T using Superscript VILO Master mix (Invitrogen). pGEX-6P2-U1C-His was sub-cloned by PCR from pGEX-2TK-U1C using EcoRI-U1C-F and XhoI-U1C-R and inserted at EcoRI-XhoI sites. pGEX-6P2-U1–70K was sub-cloned by PCR from pINTO-NSA:hSNRNP70 using EcoRI-U1–70K-F and XhoI-U1–70K-R and inserted at EcoRI-XhoI sites. pcDNA-Flag-*SAM68* and pcDNA-Flag-*Sam68* were constructed by inserting corresponding cds, obtained by PCR from pGEX-6P3-*SAM68*-Flag and pGEX-6P3-*Sam68*-Flag using EcoRI- *SAM68*-F and NotI- *SAM68*-R at EcoRI-NotI sites.

pcDNA-Flag-*SAM68*$_{I184N}$ and pcDNA-Flag-*SAM68*$_{V229F}$ were generated by swapping the 679 bp, AgeI-XbaI fragment from pcDNA mCherry-*SAM68*$_{I184N}$ and pcDNA mCherry-*SAM68*$_{V229F}$, respectively, to pcDNA-Flag-*SAM68* WT. pcDNA-Flag-*SAM68*-Nter and pEGFP-*Sam68*-Nter were generated by deletion PCR with *SAM68*-F and *SAM68*-Nter-R primers using pcDNA-Flag-*SAM68* and pEGFP-*SAM68* as templates, respectively. pcDNA-Flag-*SAM68*-Cter was generated from pcDNA-Flag-*SAM68* by PCR using EcoRI-*SAM68*-Cter-F and NotI-*SAM68* Cter-R and cloning the amplicon at EcoRI-SalI sites of pEGFP-C1. pEGFP-C1-*SAM68*-C1 was obtained by deletion PCR using *SAM68*-F and *SAM68*-C1-R, respectively. pEGFP-*SAM68*-C2 to C5 were generated by cloning the PCR amplicons obtained using the reverse primer, SalI-*SAM68*-R and forward primers namely EcoRI-*SAM68*-C2, C3, C4, C5 at EcoRI-SalI sites of pEGFP-C1. pEGFP-*SAM68*-NLS was obtained by deletion PCR using EcoRI-*SAM68*-NLS-F and EcoRI-*SAM68*-NLS-R. pLKO-sh*SAM68* was generated by annealing and inserting the oligos, sh*SAM68*-F and sh*SAM68*-R, at AgeI-EcoRI restriction sites of pLKO.1 (Addgene plasmid #8453). Primer sequences and generated plasmid are listed in supplementary Table S1.

Antibodies, western blotting and immunoprecipitation

The following antibodies were used in this study: anti-Flag (1:2000, 2368S, Cell Signaling Technology), anti-U1–70K (1:1000, 05–1588, EMD-Millipore), anti-U1A (1:1000, ab55751, abcam), anti-U1C (1:1000, ab192028, Abcam),

anti-SAM68 (1:2000, AD-1, gift from Dr Stéphane Richard), anti-GAPDH (1:2000, MM-0163-P, Médimabs), anti-mTOR (1:1000, 2983S, Cell Signaling Technology), beta-actin (1:1000, 8457L, Cell Signaling Technology), anti-GFP (1:2000, ab290, Abcam) and anti-His (1:1000, 12698S, Cell Signaling Technology). Endogenous immunoprecipitation was done by immobilizing anti-U1–70K, anti-U1A or control IgG-Mouse (Santa Cruz Biotechnology, sc-2025) on Protein A/G PLUS-Agarose beads (Santa Cruz Biotechnology, #sc-2003). HEK-293T cells were harvested, lysed for 10 min at 4°C in 1× phosphate-buffered saline (PBS), pH 7.4, 1% Triton-X-100, 150 mM NaCl, RNaseA (10 mg/ml; Sigma, #R5503) and 1× protease inhibitor complete ethylenediaminetetraacetic acid (EDTA)-free (Roche). The cell lysates were sonicated five times for 30 s with a Bioruptor ultrasonic cell disruptor and centrifuged at high speed for 30 min at 4°C to remove cell debris. For endogenous immunoprecipitation, the respective antibodies conjugated to Protein-A/G PLUS-Agarose beads (Santa Cruz Biotechnology, sc-2003) were added to pre-cleared cell lysates. Following 1 h at 4°C, the beads were washed several times with lysis buffer and the immunoprecipitates were eluted with Laemmlli buffer. Flag-tagged proteins were immunoprecipitated with Flag-M2 affinity beads (Sigma, #A2220). GFP-tagged proteins were immunoprecipitated with homemade GFP-Trap-A beads.

RNA immunoprecipitation and RT-PCR

RNA from 50% of the Flag-*Sam68* and Flag-YFP immunoprecipitates was isolated using TRIzol™ Reagent (Invitrogen, #15596–018) and reverse transcribed using SuperScript™ VILO™ MasterMix (Invitrogen, #11755–050) or M-MuLV reverse transcriptase, according to the manufacturer's instructions. Extracted RNA was incubated with random hexamers, oligo-dT or gene-specific primer for first strand cDNA synthesis at 25°C for 10 min, 42°C for 1 h and 85°C for 5 min. One-fifth of the reaction product was amplified for U1snRNA transcript using U1snRNA-F and U1snRNA-R primers and GAPDH mRNA using GAPDH-F and GAPDH-R primers, respectively. Primer sequences are listed in Supplementary Table S1.

RNA-binding assay

Purified SAM68 or U1A (stored in 20 mM Tris–HCl, pH 7.4, 200 mM NaCl, 10% Glycerol, 1 mM dithiothreitol) was added at the indicated concentration to a mix containing 10 nM of ^{32}P-labeled U1 snRNA in 1× RNA binding buffer (50 mM Tris–Cl, pH 7.6, 200 mM potassium acetate, 5 mM $MgCl_2$, 2.5 mM dithiothreitol) supplemented with 0.25 μl of RNasin Ribonuclease Inhibitor. The mix was incubated at room temperature for 15 min, then complexes were fixed with 0.5% glutaraldehyde for 10 min at room temperature. The samples were loaded onto a 5.5% Tris/Borate/EDTA (TBE) 1× acrylamide (29:1) gel and run at 150 V for 2 h and 30 min at 4°C, dried onto DE81 filter paper, then visualized by autoradiography. Quantifications were performed on a FLA-5100 phosphorimager system (Fujifilm), and statistics were analyzed with Prism.

Protein purification and GST pulldown

Competent *Escherichia coli* BL21 DE3 Codon plus RP strain (Stratagene, Agilent technologies # 230255) were transformed with pGEX-6p3-*SAM68*-flag, pGEX-6p3-*Sam68*-flag, pGEX-6p3-*SAM68*-Nter, pGEX-6p3-Cter, pGEX6p2-U1A-His, pGEX6p2-U1–70k-His, pGEX6p2-U1C-His and empty vectors, namely pGEX6p3-flag and pGEX-6p2-His, respectively. Single colony of each construct was then grown in LB media at 37°C until desired density and then induced with 0.3 mM isopropyl B-D_1-thiogalactopyranoside (IPTG) at 30°C overnight. Bacterial pellets were collected and lysed with 50 mM Tris–Cl, pH 7.5, 1% Triton-X-100, 200 mM NaCl, 20% glycerol, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT) for 30 min, sonicated and centrifuged at 17 000 *g* for 30 min 4°C. GST fusion proteins were bound to 500 μl of washed and equilibrated Glutathione agarose beads (Sigma, #G4510). For GST-SAM68-flag, the GST-tag was removed by 20 units of PreScission Protease (2U/μl; GE Healthcare, #27-0843-01) at 4°C overnight in 50 mM Tris-HCl, pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT. Following this step, SAM68-flag was immunoprecipitated with flag-M2 affinity beads (Sigma, #A2220) for 1 h at 4°C, eluted with 9 μg of 3X FLAG® Peptide (4 mg/ml**;** Sigma, #F4799), and pooled elutes were dialyzed in 1× PBS, 200 mM NaCl, 0.05% Triton-X-100, 20% glycerol. Aliquots of the preparation were run on sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Brilliant Blue to validate preparation purity (Supplementary Figure S1).

For the NMR part of this study, GB1-hSAM68 (C2), GB1-U1A (1-282), U1A linkerRRM2 (156–282) and U1A RRM1 (1-126) were expressed under the control of lactose inducible promoters in *E-coli* BL21 DE3 at 37°C during 4 h in presence of 1 mM IPTG. To achieve isotope labeling, cells were grown in M9 minimal medium complemented with ^{15}N-NH_4Cl and/or ^{13}C-glucose. Bacterial pellets were re-suspended in buffer A (10 mM HEPES, pH 7.6, 500 mM NaCl, 0.5 M Urea, β-mercapto-ethanol 2.8 mM) in the presence of lysozyme and DNAse I (0.01 mg/ml each). Cells were lysed using a microfluidizer by five cycles at 15 000 psi and the lysates were clarified by centrifugation (30 000 *g*, 40 min). Proteins were then purified using Ni^{2+}-affinity chromatography and eluted with a gradient of imidazole. The C-terminal histidine tag of U1A RRM1 and U1A RRM2 were cleaved by thrombin (10 U/mg of purified protein, 6 h at room temperature in buffer A), while GB1-hSAM68 (C2) and GB1-U1A were kept as fusion proteins. All the proteins were further purified by size exclusion chromatography in 10 mM sodium phosphate, pH 6.8, NaCl 50 mM.

For GST pulldown, 300 ng of purified recombinant hSAM68-Flag was incubated with 150 ng of glutathione bound GST-tagged U1–70K-His, U1A-His, U1-C-His or with negative control GST-His in binding buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 10% glycerol, 0.5% Triton-X-100, 10 mg/ml RNaseA), supplemented with 1× protease inhibitor complete EDTA-free (Roche). The reaction volume was made upto 300 μl in total and incubated for 1 h at 4°C. Beads were washed and bound proteins were resolved by SDS-PAGE and analyzed by western blot.

Biotinylated probes synthesis and RNP pulldown

Intron 5 baits with and without 5′ splice sites and either Wild-Type (WT) or Mutated (Mut) SAM68-binding sequence were cloned in pcDNA-Neo. The baits were *in vitro* transcribed from EcoRI linearized plasmids using T7-Hi-Scribe kit (Invitrogen) according to manufacturer's instructions. After DNaseI treatment, the reaction mixtures were column purified and 3′ end labeled with UTP-Biotin (Roche) with poly-U-polymerase according to manufacturer's instructions. Baits were immobilized on streptavidin agarose beads and were washed three times with binding buffer (20 mM HEPES, pH 7.9, 200 mM NaCl, 10% glycerol, 0.5% Triton-X-100, RNase inhibitor, 1× protease inhibitor). The beads were blocked with 100 μg/ml of bovine serum albumin for 30 min at 4°C and washed again in binding buffer. Immobilized baits were incubated with 1 μg of either purified SAM68-Flag or GST-Flag for 30 min at 4°C, then sh*SAM68* HEK-293T cell lysates were added for 1-h incubation at 4°C. Beads were washed in binding buffer and the retained proteins were eluted in Laemmli buffer, run on SDS-PAGE gels and blotted using Flag and U1A antibodies.

NMR spectroscopy

NMR data were recorded at 313 K using the 500 MHz Avance III or the 600 MHz Avance III (Bruker), both equipped with cryo-probe. Data were processed with Topspin (Bruker) and analyzed with CARA. Sequential assignment of hSAM68(C2) was deduced from the analysis of classical triple resonance experiments (3D HNCACB, 3D CACB(CO)NH, 3D HNCO and 3D HN(CA)CO). Uniformly ^{15}N labeled GB1-hSAM68 (C2) was titrated with unlabeled GB1-U1A, U1A RRM1 (1-126) or U1A linker-RRM2 (156–282) and the formation of the binary complexes was monitored by measuring 2D ^{15}N-^{1}H HSQC spectra after each addition. Reverse NMR titrations of uniformly ^{15}N labeled U1A versions by unlabeled GB1-SAM68 (C2) were performed using a similar strategy. Chemical shift perturbations were plotted onto the surface representation of the structure of the free form of the RRM1 of U1A (47).

In vitro Xrn-1 protection assay

XbaI linearized pcDNA 3.1-mTor minigene plasmid was used for *in vitro* transcription. Plasmid DNA templates were eliminated by DNaseI treatment followed by column purification of the RNA template. RNA templates were either used directly for 5′ Xrn-1 exonuclease assay or 3′ end labeled with poly-UTP-biotin using poly-U-polymerase and bound to streptavidin agarose beads. Streptavidin-bound RNA templates were incubated with WT MEFs cell extracts or Sam68$^{-/-}$ MEFs cell extracts supplemented with either *in vitro* purified GST-Flag or mSAM68-Flag for 30 min at 4°C in binding buffer (10 mM Tris–HCl, pH 7.9, 10 mM $MgCl_2$, 50 mM NaCl, 0.5% Triton-X-100, 1 mM DTT). 5′ monophosphate RNA templates were generated by treating with RNA 5′ pyrophosphohydrolase or RppH (5000 U/ml, New England Biolabs, M0356S) for 1 h at 37°C in reaction buffer (50 mM NaCl, 10 mM Tris–HCl, pH 7.9, 10 mM $MgCl_2$, 1 mM DTT). The 5′ monophosphate RNA transcripts were then treated with 1U of Xrn-1 (1000 U/ml, NEB, M0338L) for 1 h at 37°C in reaction buffer. The reaction was stopped by heating at 70°C for 10 min. cDNA was amplified using primer RRT and amplicon corresponding to full-length bait was produced using forward (FSS) and reverse (RSB) primers while mSAM68 protected amplicon was produced using forward (FSB) and reverse (RSB) primers.

RNA immunoprecipitation (RIP)

RNA immunoprecipitation (RIP) assays were performed using a modified CLIP protocol (48). Nuclear fraction was isolated and lysed in 1× PBS, pH 7.4, 150 mM NaCl, 1% Triton-X-100, 0.5% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail for 30 min on ice. The extracts were treated with DNase I (10 U/μl, Roche, Cat. No. 04716728001) and RNase A (10 mg/ml, Sigma, R-4875) for 1 h at 4°C under rotation. The samples were diluted in binding buffer (1× PBS, pH 7.4, 150 mM NaCl, 1% Triton-X-100) and supplemented with RNase inhibitor. Samples were precleared and 10% of were saved for inputs, while the rest were used for immunoprecipitation using anti-U1A antibody (Abcam, mab55751) or rabbit IgGs (Jackson Immuno Research, 011–000-003). The inputs and immunoprecipitated samples were treated with 50 μg Proteinase K for 30 min at 65°C. RNA was then isolated and used for RT-qPCR analyses that were performed using primers ei5-F and ei5-R for exon5–intron 5 junction, ei4-F and ei4-R for exon4–intron4 junction and ei37-F and ei37-R for exon37–intron37 junction.

In vivo splicing assays

Endogenous *mTor* transcript premature termination and polyadenylation in wild-type (WT) mouse embryonic fibroblasts (MEFs) and *Sam68*$^{-/-}$ MEFs were rescued using lentivirus-mediated transduction of mouse *Sam68(*WT) or mouse *Sam68*(ΔARM). Total RNA was extracted using TRIzol® reagent and digested with DNase I to get rid of contaminating genomic DNA. After column purification, cDNA was amplified by reverse transcription using oligo-dT and M-MuLV reverse transcriptase. The following pairs of primers were used F(e4) and R(i5) to amplify $mTor_{i5}$ and F(e4) and R(e6) to amplify $mTor_{Exon4–6}$. A portion of cell lysates were resolved in SDS-PAGE and blotted with mTOR, SAM68 and GAPDH antibodies.

mTor genomic fragment spanning from exon4 to exon6 with corresponding introns 4 and 5 was cloned in pcDNA Neo and referred to as $mTor_{Exon4–6}$ minigene, hereafter. sh*SAM68* HEK-293T cells were co-transfected with $mTor_{Exon4–6}$ minigene and either Flag-YFP, Flag-h*SAM68*-WT or Flag-h*SAM68*ΔARM. Forty-eight hours post-transfection, total RNA was extracted using Trizol® reagent and digested with DNase I to get rid of contaminating plasmid DNA. After column purification, cDNA was amplified by reverse transcription with random hexamers using Superscript VILO according to manufacturer's instructions. Cell lysates were resolved in SDS-PAGE and

blotted with Flag and β-actin antibodies. To identify splicing patterns derived from the reporter ($mTor_{Exon4–6}$ minigene), we used the following primers: F2; T7 Forward and R2; exon4 Reverse to amplify the total RNA produced from $mTor_{Exon4–6}$ minigene. F1; exon4 Forward and R1; intron 5 Reverse were used to amplify intron 5 retained minigene transcripts. Primer pairs GAPDH-F and GAPDH-R were used to amplify *GAPDH* mRNA. Primer sequences are listed in Supplementary Table S1.

RESULTS

SAM68 interacts with U1 snRNP

As aforementioned, mSAM68 depletion causes an increase in the rate of *mTor* intron 5-induced premature termination, resulting in the production of a non-translated smaller mRNA, termed $mTor_{i5}$ (21,40). Analysis of the intron 5 shows the presence of two stretches of high-affinity SAM68 U/A binding elements located near the upstream 5′ splice site, one of which is embedded in a cryptic polyadenylation signal (Figure 1A). This suggests that SAM68 association with these U/A rich sequences could facilitate 5′ splice site recognition and promote normal splicing of intron 5. Hence, we reasoned that SAM68 could be interacting with the U1 snRNP, the spliceosomal component that recognizes the 5′ splice site, and thus be a determining factor for early spliceosome assembly within *mTor* intron 5 (4). While the initial discovery was observed in mice, we reasoned that any SAM68–U1 interaction would also be observed in human cells, as human SAM68 (hSAM68) is almost identical to mouse SAM68 (mSAM68).

To validate if SAM68 could interact with U1 snRNP, flag-tagged hSAM68 (flag-hSAM68) was expressed in HEK-293T cells depleted of endogenous *SAM68* using lentiviral-shRNA targeting the 3'UTR of *SAM68* mRNA (sh*SAM68* HEK-293T) (Supplementary Figure S2A). Flag-tagged yellow fluorescent protein (flag-YFP) was used as negative control. Western blot analyses showed that the three core protein components of U1 snRNP, namely U1–70K, U1A and U1C, efficiently co-immunoprecipitated with Flag-hSAM68, while undetected in Flag-YFP controls (Figure 1B). RIP followed by RT-PCR using primers specific to U1 snRNA and GAPDH mRNA showed that U1 snRNA only co-immunoprecipitated in the presence of Flag-hSAM68, confirming that the U1 snRNP could be associated with hSAM68 (Figure 1B). This interaction was further validated endogenously, since U1 snRNP component U1A co-immunoprecipitated, with SAM68 (Supplementary Figure S2B). Concomitantly, hSAM68 was efficiently immunoprecipitated with endogenous U1–70K and U1A as compared to control immunoglobulin G (IgG) (Figure 1C and D).

Interaction of SAM68 with U1 snRNP is RNA independent

Given that SAM68 is an RNA-binding protein, we first assessed if the SAM68-U1 snRNP interaction was mediated through U1 snRNA. Sequence analyses showed the presence of a potential SAM68-binding site (AUAAUUU) upstream and partially within the Sm domain (49). We tested and compared the affinity of both Sam68 and U1A to the U1 snRNA. While SAM68 can indeed bind U1 snRNA, it showed minimal affinity for this RNA when compared to U1A, a *bona fide* U1 associated protein (Figure 1F and G). Considering that SAM68 has numerous preferential RNAs targets (21,27,50) and its low affinity for the U1 snRNA (Figure 1F), it would be highly unlikely that its association with the U1 snRNP is mediated by U1 snRNA.

To identify which protein component of the U1 snRNP mediates this interaction, recombinant hSAM68-Flag were incubated with cell extracts from sh*SAM68* HEK-293T cells in the presence or absence of RNase A. Treatment with RNAse A had no effect on SAM68 association with U1 snRNP core components even though total RNA was completely digested, suggesting that SAM68 captures U1 snRNP independent of RNA (Figure 2A). However, U1 snRNPs were previously shown to resist RNaseA treatment, due to a compacted conformation (51). Therefore, these results do not exclude the possibility that tight association between U1 snRNP and SAM68 could also protect RNA mediating their interaction from enzymatic digestion. To demonstrate that this interaction was indeed RNA independent, we transfected sh*SAM68* HEK-293T with an RNA-binding deficient mutant of SAM68 (*Flag-hSAM68*$_{I184N}$) (36,52). This mutant was still able to co-immunoprecipitate U1 snRNP, further confirming that this interaction is not mediated through RNA but through protein–protein interactions (Figure 2B). This result was corroborated with another RNA-binding deficient mutant of SAM68 (Flag-hSAM68$_{V229F}$) (23) (Supplementary Figure S2C).

SAM68 interact with the U1 snRNP component U1A

To identify which of the core components of U1 snRNP had more affinity for SAM68, we first performed a salt sensitivity test strategy previously shown to disrupt U1 core proteins from the U1 snRNA, therefore disrupting the complex (53). Flag pulldown of purified hSAM68-Flag incubated with sh*SAM68* HEK-293T cell lysate were washed with buffer containing increasing salt concentration. Residual association between U1 snRNP proteins and Sam68 was then detected by Western blot (Figure 2C). We found that only U1A remained bound to hSAM68-Flag at salt concentration exceeding 300–400 mM of NaCl, while both U1–70K and U1C dissociate from the complex at a concentration of 200–250 mM NaCl. This result suggested that the interaction between SAM68 and U1 snRNP may be mediated by U1A. To address this, we incubated purified recombinant GST tagged U1A-His, and as controls U1C-His and U1–70K-His with the aforementioned purified hSAM68-Flag. Using these *in vitro* purified proteins, we confirmed that only GST-U1A-His could efficiently pulldown hSAM68-Flag, while both GST-U1C-His and GST-U1–70K-His as well as the negative control (GST-His) failed to pull down hSAM68-Flag (Figure 2D). Furthermore, U1A-His was still bound to hSAM68-Flag in the presence of micrococcal nuclease (MNAse), a non-specific endo-exonuclease that degrades both DNA and RNA, further demonstrating a direct interaction between both proteins, which is not mediated by nucleic acids (Supplementary Figure S2D).

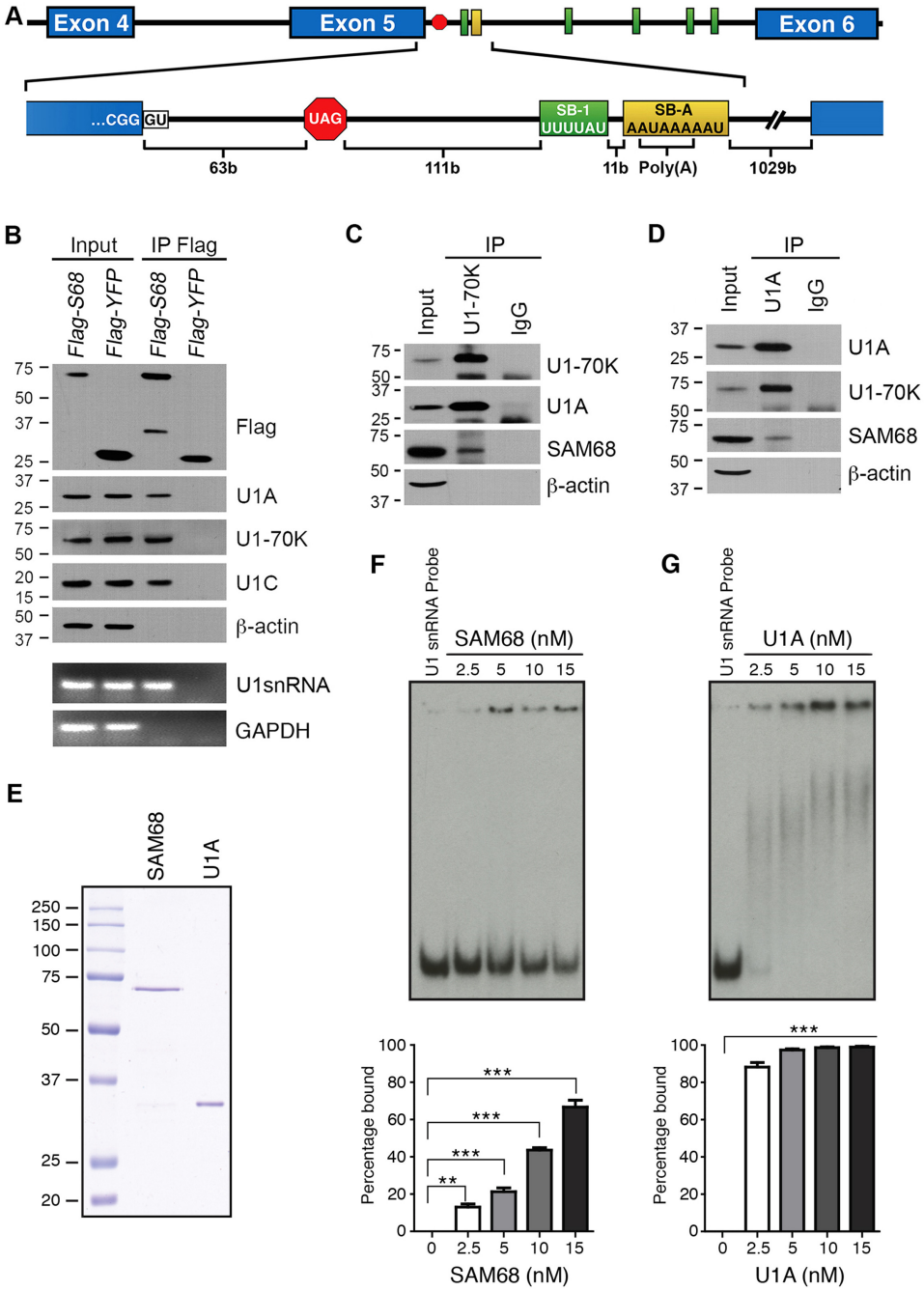

Figure 1. *In vivo* association of SAM68 with U1 snRNP. (**A**) Schematic representation of a portion of *mTor* pre-mRNA spanning from exon4 to exon6 (upper panel), with a close-up of the 5′ splice site and the subsequent SAM68-binding site (SB-1), as well as the cryptic polyadenylation signal that harbor SAM68-binding site (SB-A). (**B**) Co-immunoprecipitation of U1 snRNP with Flag-hSAM68. HEK-293T cells depleted of endogenous SAM68 (sh*SAM68* HEK-293T) were transiently transfected with Flag-hSAM68 or Flag-YFP (yellow-fluorescent protein), the latter serving as negative control. Flag-tagged proteins were immunoprecipitated using anti-Flag M2 agarose beads and immunoprecipitated proteins were detected with antibodies specific to U1–70K, U1A and U1C. β-Actin was used as negative control. Portion of the Flag-immunoprecipitates was used for RNA isolation and RT-PCR using U1 snRNA specific primers. GAPDH (glyceraldehyde 3-phosphate dehydrogenase) RNA was used as negative control of the RT-PCR made from the RNA immunoprecipitation. (**C**) Co-immunoprecipitation of endogenous hSAM68 with U1–70K. Immunoprecipitated proteins were detected with antibodies directed against SAM68 and U1–70K. β-Actin was used as negative control of immunoprecipitated proteins. (**D**) Co-immunoprecipitation of endogenous hSAM68 with U1A. Immunoprecipitated proteins were detected with antibodies directed against SAM68 and U1A. β-Actin was used as negative control of immunoprecipitated proteins. (**E**) Coomassie staining of purified human SAM68 and U1A. (**F**) RNA binding assay with purified SAM68 and labeled U1snRNA. Reactions contained 10 nM γ-p32 labeled U1snRNA in buffer with no protein (lane 1) or with purified SAM68 (lanes 2–5). Bottom panel: quantification from three independent binding experiments. Error bars represent the corresponding standard error. Unpaired two-tailed *t*-tests were used to compare the different concentrations of purified protein to the RNA only control. SAM68 *P*-values are 0.0014, 0.0005, <0.0001, <0.0001 in increasing order of SAM68 concentration. (**G**) RNA binding assay with purified U1A and labeled U1snRNA. Reactions contained 10 nM γ-p32 labeled U1snRNA in buffer with no protein (lane 1) or with purified U1A (lanes 2–5). Bottom panel: U1snRNA *P*-values = 0.0008, <0.0001, <0.0001, <0.0001 in increasing order of U1A concentration. ***P*-value < 0.005, ****P*-value < 0.001.

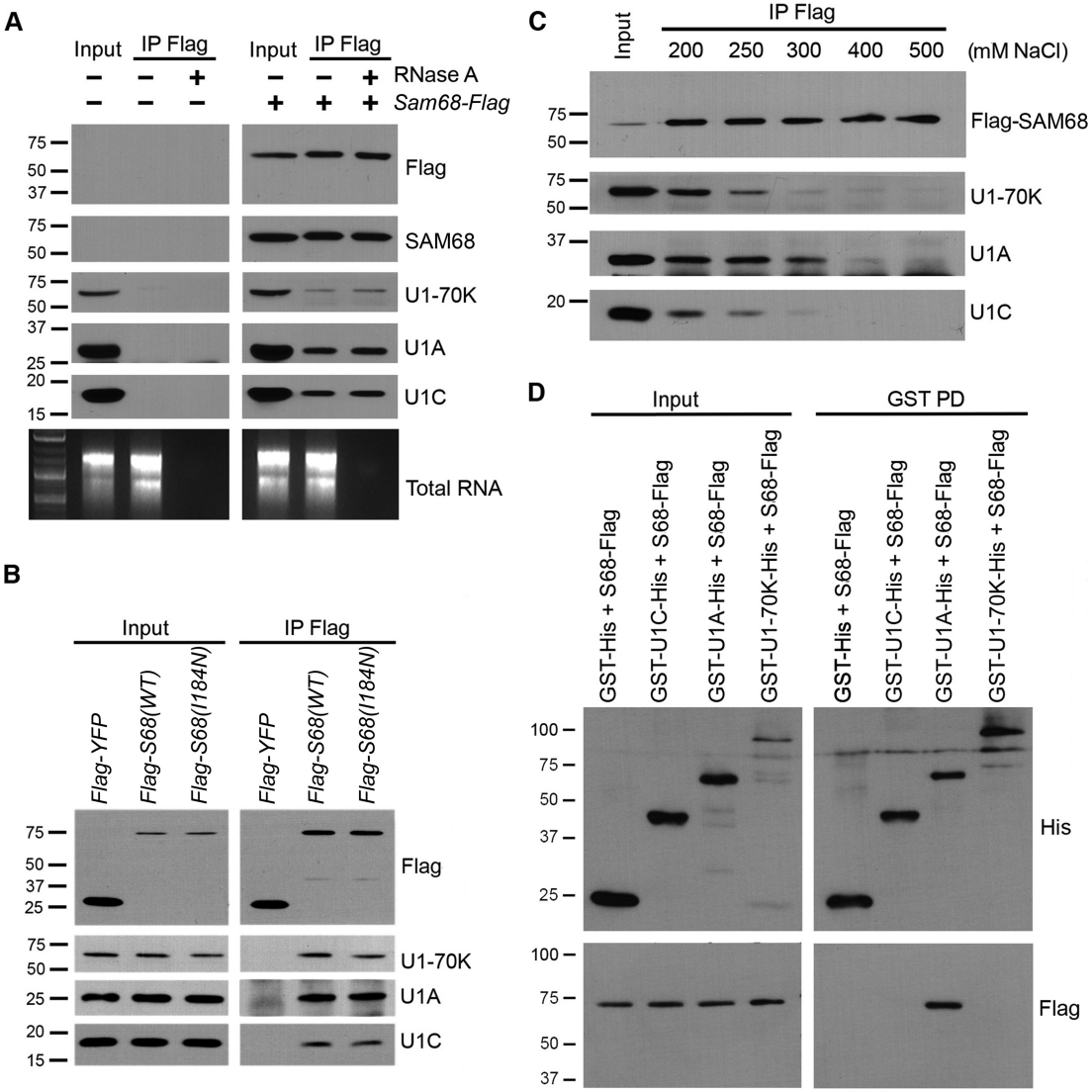

Figure 2. *In vitro* Purified SAM68 associated with U1 snRNP in an RNA-independent manner. (**A**) *In vitro* purified hSAM68-Flag was added to sh*SAM68* HEK-293T cell lysates for 1 h at 4°C, in the presence or absence of 50 μg/ml RNaseA. hSAM68-Flag and associated proteins were immunoprecipitated using Flag-M2 affinity beads and treated further with RNaseA at 37°C for 30 min. Bound proteins were eluted with Laemmli and immunoblotted with antibodies specific to U1–70K, U1A and U1C. To assess RNaseA treatment efficiency, total RNA from sh*SAM68* HEK-293T was treated with either Mock or RNaseA for 30 min at 37°C, and the remaining total RNA was assessed on agarose gel. (**B**) RNA-binding defective mutant hSAM68$_{I184N}$ interacts with U1 snRNP. sh*SAM68* HEK-293T were transiently transfected with Flag-hSAM68, Flag-hSAM68$_{I184N}$ and Flag-YFP (negative control). The Flag-tagged proteins were immunoprecipitated using anti-Flag M2 agarose beads and immunoblotted with antibodies directed against U1–70K, U1A and U1C. (**C**) Association of hSAM68-Flag with U1 snRNP withstands high salt washes. Purified *in vitro* produced hSAM68-Flag was added to cell lysates of sh*SAM68* HEK-293T for 1 h at 4°C. Flag-M2 affinity beads were added to the reaction and left for 1 h at 4°C. The washes were done, by increasing salt concentration, from 150 to 500 mM of NaCl. Bound proteins were eluted with Laemmli and immunoblotted with antibodies directed against U1–70K, U1A and U1C. (**D**) SAM68 interacts with U1A *in vitro*. About 300 ng of purified hSAM68-Flag was incubated with 100 ng of glutathione-agarose bound GST-U170k-His, GST-U1A-His, GST-U1C-His and GST-His. Following washes, the beads were washed five times in binding buffer and the bound proteins eluted with Laemmli and immunoblotted using anti-Flag or anti-His antibodies.

SAM68 interaction with U1A is mediated through its tyrosine-rich (YY) domain

Being an adaptor protein, SAM68 comprises many protein–protein interaction domains such as SH3 binding proline-rich motifs and SH2 binding tyrosine-rich domain (28). In order to determine which domain was responsible for the association with U1A, we first truncated hSAM68 in two fragments (Figure 3A). The first fragment (N-term) contains the RNA-binding domain of hSAM68, spans from amino acids 1–280 and comprises the KH domain and proline-rich motifs (P0-P2) and the hSAM68 nuclear-localization signal (NLS). The second fragment (C-term) spanning from amino acids 281–443 comprises proline-rich motifs (P3-P5), the tyrosine-rich domain (YY), the SH2-binding tyrosine residues and the hSAM68 NLS. Both fragments were Flag-tagged at their N-terminus (Supplementary Figure S3A). Both flag-tagged hSAM68 fragments (N- and C-term), along with positive control Flag-hSAM68 and

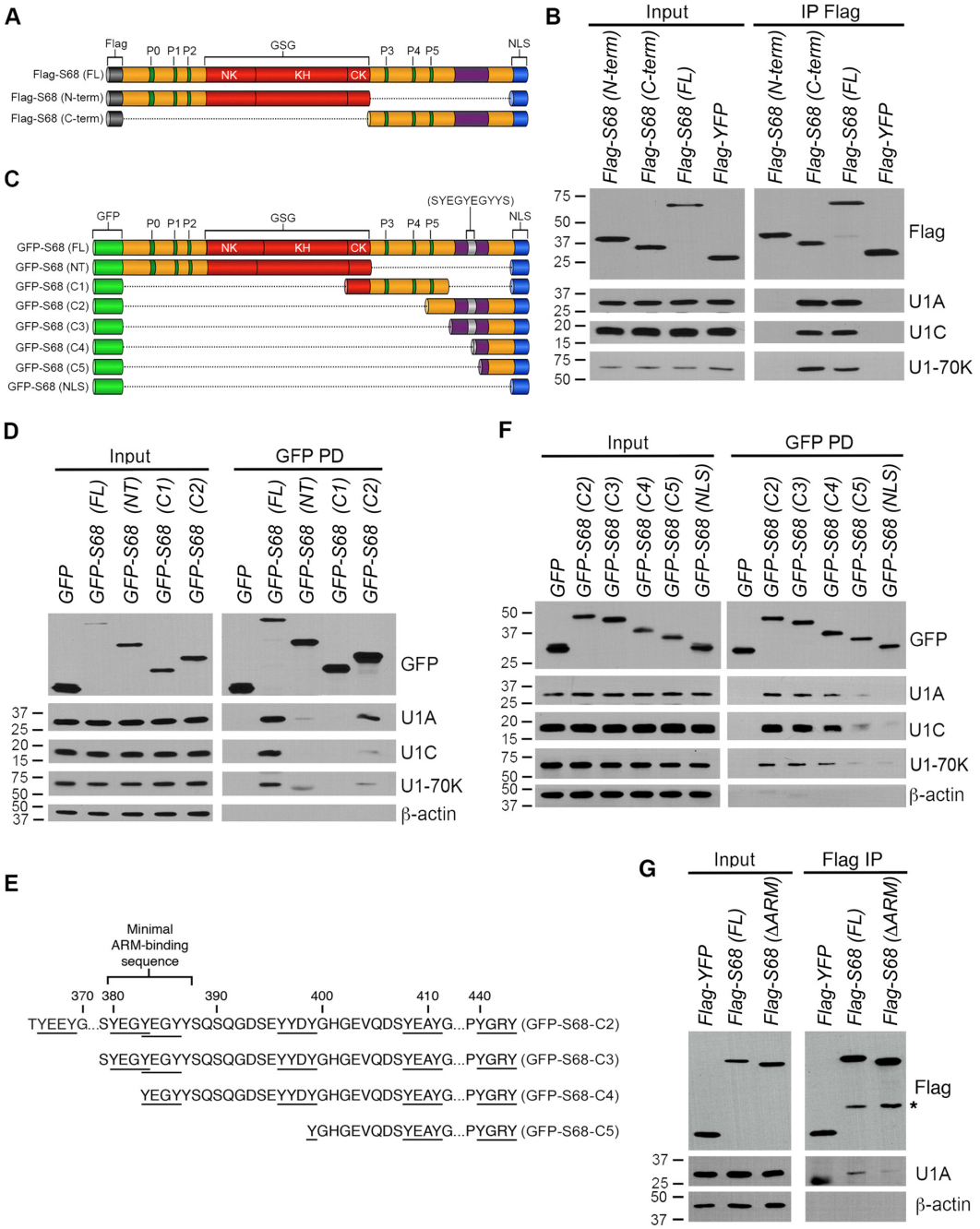

Figure 3. SAM68 interaction with U1A is mediated through its C-terminal portion. (**A**) Schematic representation of C-terminus (aa. 1–280) and N-terminus (aa.281–443) deletion domains of hSAM68 fused to flag. (**B**) sh*SAM68* HEK-293T cells were transiently transfected with Flag-*SAM68*(N-term), Flag-*SAM68*(C-term), Flag-*SAM68*(FL) and flag-YFP (negative control). Forty-eight hours post transfection, the flag-tagged proteins were immunoprecipitated using anti-flag M2 agarose beads and immunoblotted with antibodies specific to U1–70K, U1A and U1C. (**C**) Schematic representation of full-length SAM68, C-terminus deleted SAM68 (NT, aa. 1–280), C-terminus truncated to proline rich C1 (aa. 269–364) and tyrosine rich C2 (aa. 365–443), C3 (aa. 370–443), C4 (aa. 385–443), C5 (aa. 340–443) and NLS (aa. 430–443). Fragments were fused to GFP tag at their N-terminus and all fragments had SAM68 NLS at their C-terminus. (**D**) GFP-Trap-A pulldown of GFP-tagged proteins. sh*SAM68* HEK-293T cells were transiently transfected with *GFP, GFP-SAM68*(FL), *GFP- SAM68*(NT), *GFP- SAM68*(C1) and *GFP-SAM68*(C2). Forty-eight hours post transfection, cells were lysed and GFP-Trap-A beads were used to pull down GFP-tagged proteins, and their association with U1A was validated by western blot using specific antibodies. (**E**) Primary amino acid sequence of the various deletion constructs of SAM68 YY domain (GFP-hSAM68 C2 to C5). Underlined indicates YXXY motifs in the YY domain. Also highlighted is the minimal ARM-binding region. (**F**) GFP-Trap-A pulldown of GFP-tagged proteins. shSAM68 HEK-293T cells were transiently transfected with *GFP*, *GFP-SAM68*(C2), *GFP-SAM68*(C3), *GFP-SAM68*(C4), *GFP-SAM68*(C5) and *GFP-SAM68*(NLS). Forty-eight hours post transfection, cells were lysed and GFP-Trap-A beads were used to pull down GFP-tagged proteins, and their association with U1A was validated by western blot using specific antibodies. (**G**) U1A binds preferentially to the minimal ARM motif (YEGYEGY) within the YY domain of SAM68. Flag-hSAM68(FL) and Flag-hSAM68(ΔARM) were transiently transfected in sh*SAM68* HEK-293T cells. Forty-eight hours post transfection, cells were lysed and Flag-tagged proteins were immunoprecipitated using anti-flag M2 agarose beads, and U1A association was assessed using U1A antibody. ⋆: denotes an unspecific band.

negative control flag-YFP, were transfected in sh*SAM68* HEK-293T cells. Confocal images showed that while the N-terminal fragment remained largely cytoplasmic, it also partially localized to the nucleus, suggesting that the addition of the NLS allows efficient nuclear localization of the N-terminal fragment (Supplementary Figure S3A). Flag-tagged proteins were immunoprecipitated and immunoblotted with U1A antibody, while U1–70K and U1C antibodies were used as positive controls. As expected, full-length hSAM68 showed a strong association with U1 A (Figure 3B) as well as U1–70K and U1C, while they were not detected in Flag-YFP immunoprecipitates (negative control). The U1 snRNP components co-immunoprecipitated with the Flag-hSAM68 C-terminal fragment with as strong of an association as the full-length protein, but not the N-terminal fragment containing the KH RNA-binding domain (Figure 3B). While confocal immunofluorescence detection shows that the N-terminal fragment can also be located in nucleus due to the addition of the NLS (Supplementary Figure S3A), the partial cytoplasmic localization of this fragment could explain the decreased association with U1A. To assess this possibility, equal amounts of *in vitro* purified N-term and C-term fragments were immobilized and incubated with whole cell lysate taken from sh*SAM68* HEK-293T, validating our initial observation where U1A specifically binds the C-terminal fragment and not the N-terminal fragment (Supplementary Figure S3B). This result further corroborates our previous observation that the interaction is RNA-independent and implies that the SAM68 binds U1A through its C-terminal region.

To refine our mapping of the SAM68–U1A interaction, we truncated the C-terminal fragment of SAM68 into two smaller parts: GFP-hSAM68 (C1) and GFP-hSAM68 (C2) (Figure 3C). As the fragments were very short, we opted to clone them in frame with the larger, GFP-tag and performed a GFP-binding protein pulldown and as expected showed a strong nuclear localization by confocal imaging (Supplementary Figure S4A). Using these constructs, we found that like the full-length hSAM68, the C2 fragment of hSAM68 was associated with U1A, while the C1 fragment did not (Figure 3D). This initial interaction mapping suggests that the YXXY motif rich domain (YY domain) of SAM68, located in the C2 fragment (six YXXY motifs) but not C1, could be involved in the interaction with U1 snRNP. This domain consists of tyrosine-rich motifs involved in the association between SAM68 and the armadillo repeat domain (ARM) of the adenomatous polyposis coli (APC) protein (54,55). To assess the importance of this motif in the SAM68–U1A interaction, we further divided the GFP-hSAM68 (C2) domain in four different fragments with decreasing number of YXXY motifs (Figure 3E). GFP-hSAM68 fragment (C3) contains five YXXY motifs, (C4) has four, (C5) has two, while (NLS) has none. We found that the co-immunoprecipitation efficiency of U1A was directly dependent on the number of YXXY motifs, since the C4 fragment showed a slight decrease in association strength, while the C5 and the NLS fragments showed little or no association with U1A (Figure 3F). This change in association is unlikely to be related to mislocalization caused by the GFP moiety, since all the fragments showed a predominant nuclear localization (Supplementary Figure S4B).

Interestingly, there are six YXXY motifs within SAM68 YY domain, of which two successive YEGY motifs (SYEGYEGYYS) are defined as the minimal ARM interaction domain (54). Results obtained in Figure 3F suggested that losing this minimal motif could drastically affect the ability of SAM68 to bind U1A, like it was observed with APC. To validate this possibility, we proceeded with the deletion of this ten amino acid stretch (SYEGYEGYYS) within the Flag-hSAM68 (FL) construct. Similarly to what was observed with APC, deletion of the minimal ARM-binding motif (ΔARM) was enough to abrogate most of SAM68 association with U1A (Figure 3G), while it did not affect SAM68 affinity for its RNA target (Supplementary Figure S4C). Taken together, these results strongly indicate that SAM68 interaction with U1A is mediated through a specific sequence found in the YY domain (55).

SAM68 interact with U1A–RRM1 domain

To further confirm that SAM68 was directly interacting with U1A and to identify which part of both proteins interact, we used solution-state NMR spectroscopy. The hSAM68 (C2) fragment was isotopically labeled and resonances from backbone atoms were assigned using classical approaches. Analysis of the backbone chemical shifts revealed that in solution, the hSAM68 (C2) fragment adopts a random coil conformation without any secondary structure. Upon addition of unlabeled GB1-U1A, several resonances of the ^{15}N-labeled hSAM68 (C2) experienced chemical shift changes (Figure 4A) that were reproduced with U1A RRM1 (1–126) but not by the C-terminal part of U1A containing RRM2 (156–282). Reverse NMR titration performed with ^{15}N-labeled GB1-U1A revealed that upon addition of unlabeled GB1-hSAM68 (C2), the NMR signals from the RRM1 of U1A experienced strong line broadening and almost disappeared from the spectra (Supplementary Figure S5). However, when the ^{15}N-labeled U1A RRM1 (1–126) was titrated with unlabeled hSAM68 (C2), several signals of the U1A RRM1 shifted from their initial positions (Figure 4B). The chemical shift perturbations observed on the N-terminal part of U1-A reveal an interaction surface with the C-terminal part of hSAM68 located between the edge of β-sheet surface (β2), the C-terminal helix α3 and the interdomain linker (Figure 4C). In addition, the NMR titration of hSAM68 (C2) by unlabeled U1A shown that the NMR signals from the tyrosine-rich sequence (370–400) of SAM68 (C2) are the most affected and thus strongly support that U1A RRM1 interacts with this aromatic rich sequence of hSAM68 *in vitro* at G-Y-E/D triplets.

SAM68 recruits U1 snRNP to the 5′ splice site of *mTor* intron 5

We next sought to determine if SAM68, through its association to U1A, could serve as an adaptor protein mediating the interaction between *mTor* pre-mRNA and the U1 snRNP. More specifically, we assessed if this interaction is mediated through the simultaneous association of SAM68 with its binding motifs (SBs) found near the 5′ splice site of

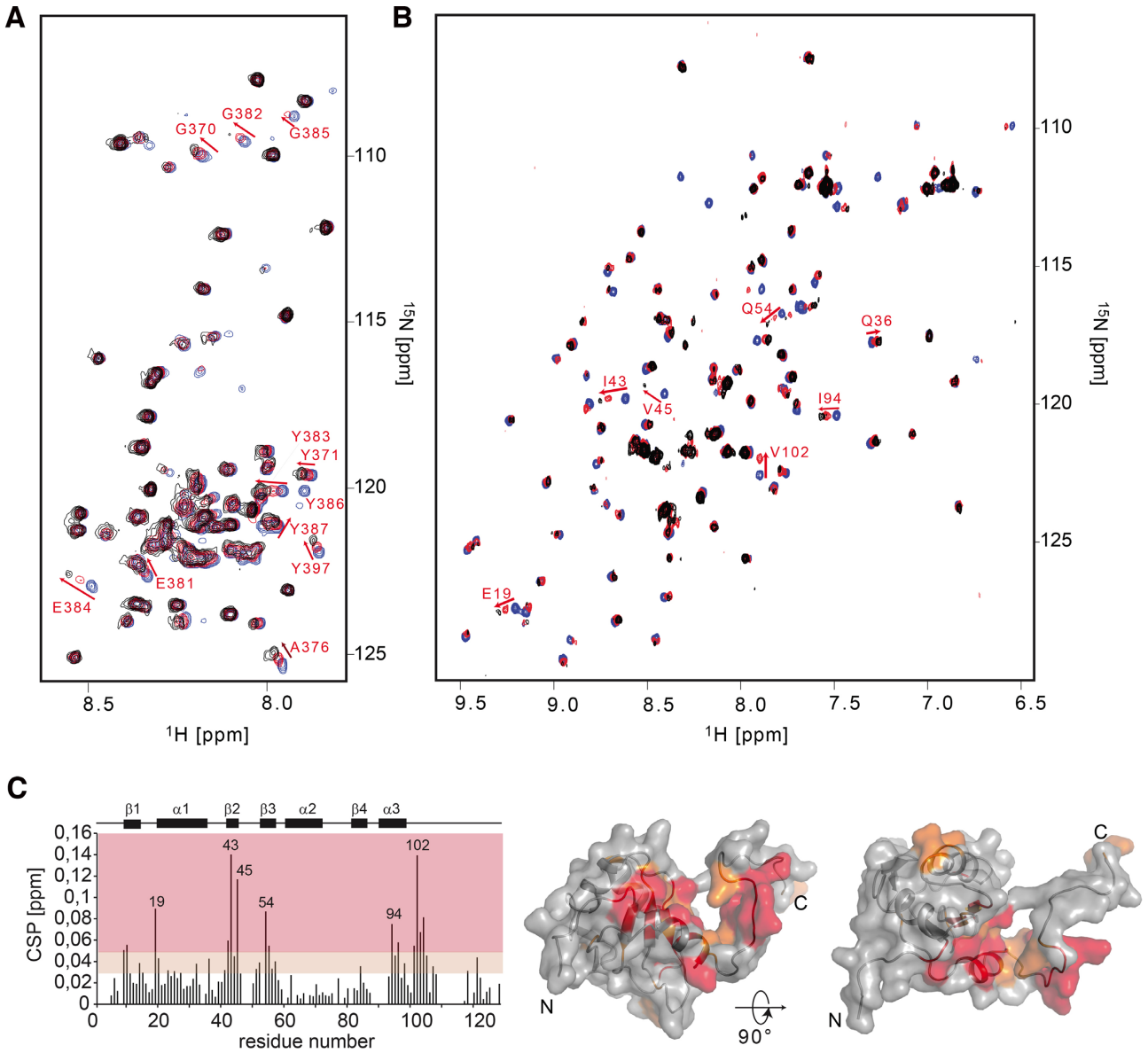

Figure 4. Tyrosine-rich (YY) domain of SAM68 mediates the interaction with U1 snRNP via YXXY repeated motif. (**A**) Overlay of the 2D ^{15}N-^{1}H HSQC spectra of GB1-hSAM68 (C2) recorded before and after the addition of unlabeled GB1-U1A. The spectra are colored according to the molar ratio hSAM68 (C2):U1A (1:0; 1:0.6 and 1:1.4 are colored in blue, red and black, respectively). Strongly perturbed signals are marked by red arrows and their assignment is indicated. (**B**) Overlay of the 2D ^{15}N-^{1}H HSQC spectra of U1A RRM1 recorded before and after the addition of unlabeled GB1-hSAM68 (C2). The spectra are colored according to the molar ratio U1A RRM1:hSAM68 (C2) (1:0; 1:0.6 and 1:1.4 are colored in blue, red and black, respectively). (**C**) Plot of the normalized chemical shift perturbations observed in panel (B) in function of the sequence of U1A RRM1. The chemical shift perturbations are then plotted onto the surface representation of the structure of the free form of the RRM1 of U1A (47). Amino acids that experienced chemical shift perturbation between 0.03 and 0.05 are colored in orange while the CSP higher than 0.05 are colored in red.

mTor intron 5 and with U1A. Indeed, while SAM68 binding sequence 1 (SB-1) does not correspond to the *bona fide* U/AAA consensus sequences identified by SELEX (35), it was previously shown that SAM68 binding to its target sequence closest to the 5′ splice site (SB-1) was essential for *mTor* normal splicing and expression during adipogenesis (21). Moreover, SB-1 shares a high level of homology to a specific SAM68 binding sequence identified in the β-actin mRNA (27). On the other hand, the SAM68 binding motif embedded within the cryptic poly(A) signal (SB-A) has the U/AAAA consensus sequences. Both SB-1 (UUUUAU) and SB-A (AUAAAAAU) were shown to be bound by mSAM68 *in vivo* (21). Interestingly, these two sequences are separated by only 11 nucleotides, which correspond to a bipartite pattern (UUUUAU-(n_{11})-AUAAAAAU), found to favor binding and homodimerization of GSG protein family, including SAM68 (56,57). As such, we reasoned that disrupting SAM68 binding to intron 5 by mutating both these sequences should drastically hinder U1 snRNP recruitment at the 5′ splice site of *mTor* intron 5 (Figure 5A).

To assess this, we *in vitro* transcribed different RNA baits using a minimal portion of the *mTor* minigene that

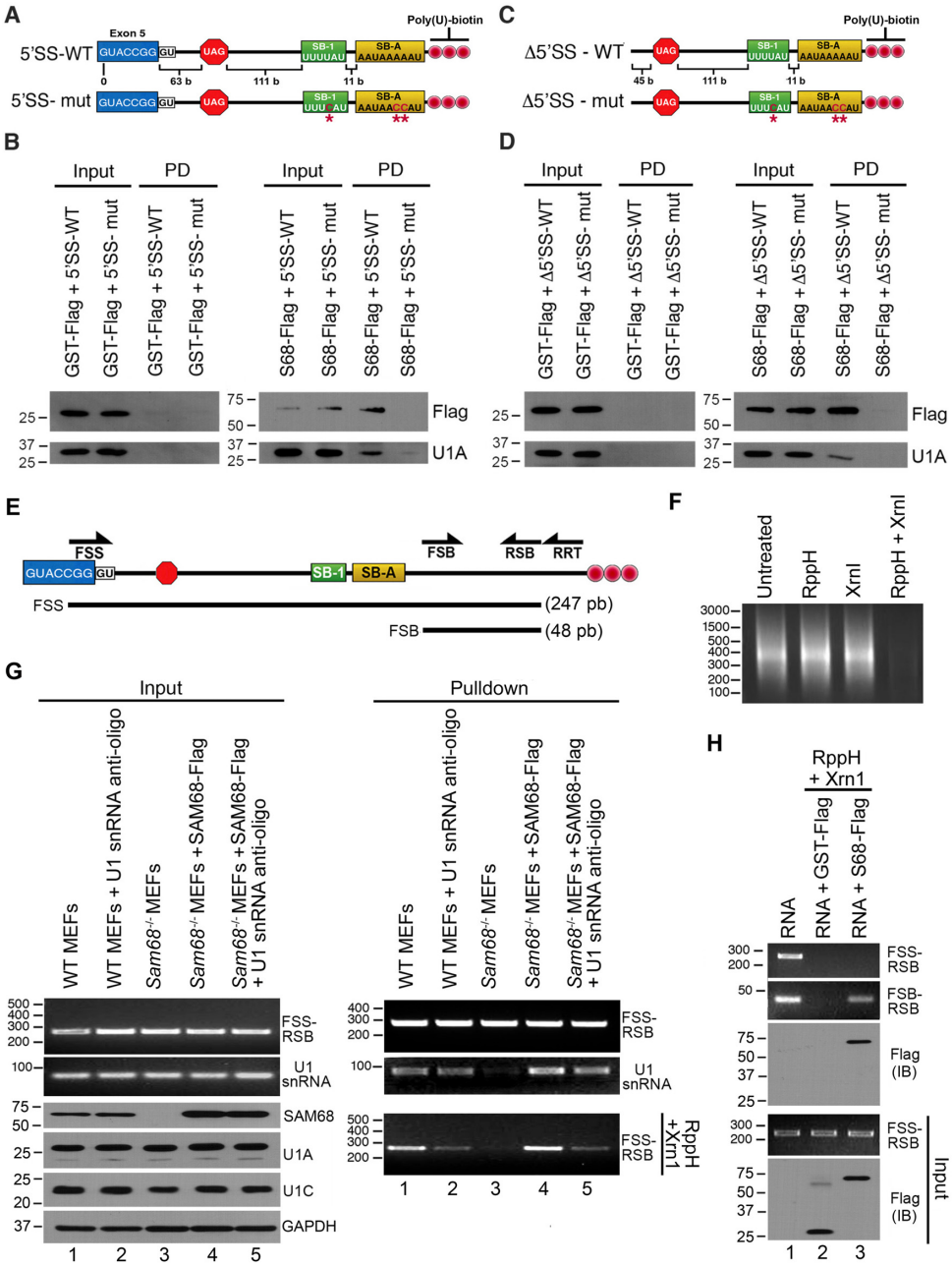

Figure 5. Both SAM68 and intronic enhancer sequences in *mTor* intron 5 are required for U1A recruitment to 5′SS *in vitro*. (**A**) Schematic representation of the various *in vitro* transcribed *mTor* minigene baits with the 5′ splice site. As shown, the baits span from last 7 nucleotides of exon5 to the poly-adenylation signal in intron 5. WT refers to the wild-type intronic SAM68-binding sequences of SB-1 (UUUUAU) and SB-A (UAAAA), the latter is embedded in the cryptic poly-adenylation signal (AAUAAA). The 'mut' denotes the combined mutations of SB-1 (UUUUAU to UUUCAU) and SB-A (AAUAAAA to AAUAACC). (**B**) SAM68 recruits U1A to 5′ splice site *in vitro*. Recombinant *in vitro* purified hSAM68-Flag was tested for its ability to recruit U1A to *mTor* intron 5 baits with either WT or mutated SAM68-binding sites. GST-Flag was used as negative control. (**C**) Schematic representation of the various *in vitro* transcribed *mTor* minigene baits that are deleted for the 5′ splice site. As shown, the baits span 18 nucleotides downstream of the 5′ splice site to the poly-adenylation signal of intron 5. WT refers to the wild-type intronic SAM68-binding sequences, SB-1 (UUUUAU) and SB-A (UAAAA). The 'mut' denotes the combined mutations of SB-1 (UUUUAU to UUUCAU) and SB-A (AAUAAAA to AAUAACC). (**D**) SAM68 recruits U1A in the absence of 5′ splice site *in vitro*. Recombinant *in vitro* purified hSAM68-Flag was tested for its ability to recruit U1A to *mTor* intron 5 baits lacking 5′SSs with either WT or mutated SAM68-binding sites. GST-Flag was used as negative control. (**E**) Schematic representation of the *in vitro* transcribed mTor minigene bait and the primers used for the RppH/Xrn1 protection assays. (**F**) Assessment of the processivity of RppH and Xrn1 enzyme on the naked mRNA bait, showing that RppH treatment is necessary for Xrn1-mediated degradation of the mRNA bait. (**G**) RppH and Xrn1 protection assays *in vitro* produced mRNA bait incubated with either WT MEFs cell lysate (lane 1), *Sam68*$^{-/-}$ MEFs cell lysate (lane 2), *in vitro* produced mSAM68(WT) + *Sam68*$^{-/-}$ MEFs cell lysate (lane 3) or *in vitro* produced mSAM68(WT) + *Sam68*$^{-/-}$ MEFs cell lysate + U1 nRNAs antisense oligo (lane 4). U1snRNP components (U1A, U1C) and mSAM68 levels were assessed by western blot, while U1 snRNA levels was assessed by RT-PCR. GAPDH served as loading control for the western blot. (**H**) SAM68 protects the mTor RNA bait from Xrn1 degradation. Biotinylated RNA baits were incubated with buffer (lane 1), 100 ng of GST-Flag (lane 2) or 100 ng of mSAM68-Flag (lane 3) for 30 min on ice. Sam68 levels were assessed by western blotting using anti-Flag, while baits levels were measured by semi-quantitative RT-PCR using FSS-RSB primers for the full-length RNA and FSB-RSB for the SAM68 protected fragment.

span from the last 7 nucleic acid of exon5 to the cryptic polyadenylation signal at the intron 5 (Figure 5A). To determine if U1 snRNP recognition of the 5′ splice site was driven by SAM68, we mutated SAM68-binding sequence closest the 5′ splice site (SB-1), as well as the binding site embedded in the polyadenylation (SB-A) (Figure 5A). The 3′-end of the RNA baits were labeled with UTP-Biotin tails and immobilized on streptavidin-agarose beads. The baits were first incubated with *in vitro* purified hSAM68-Flag recombinant protein and then, sh*SAM68* HEK-293T cell extract was added to the mix. As observed in figure 5B, U1A was mostly detected on the baits harboring WT SAM68-binding sites, while its presence was greatly decreased on baits lacking the SAM68-binding sites. This suggests that not only SAM68 association with *mTor* intron 5 strengthen the 5′ splice site recognition by U1 snRNP, but also that SAM68 association could recruit U1 snRNP to the 5′ splice site of *mTor* intron 5 (Figure 5B). To assess this, we used RNA baits lacking the 5′ splice site (Δ5′SS) harboring either WT or mutated SB-1 and SB-A (Figure 5C). Surprisingly, we found that SAM68 was able to recruit the U1A even when the 5′ splice site was absent (Figure 5D).

While these results suggest a Sam68-dependent recruitment of U1A, it remains unclear whether U1 snRNP is only tethered to the RNA bait by a direct protein interaction with SAM68 or if there is an improved recognition by U1 snRNP to the 5′ splice site. To determine that, we incubated a mRNA bait produced *in vitro*, similar to the one used in the previous pulldown experiments (Figure 5E), which was incubated with different nuclear extracts and subjected to Xrn1, a 5′→3′ exoribonuclease (58). As shown in figure 5F, Xrn1 can efficiently degrade the RNA bait following treatment with RppH, an RNA 5′ pyrophosphohydrolase that removes pyrophosphate from the 5′ end of triphosphorylated RNA to leave a 5′ monophosphate RNA, the substrate of Xrn1 (59–61). To determine if SAM68 could recruit and enhance U1 snRNP 5′ splice site recognition, we incubated our bait with either WT MEFs cell lysate, Sam68$^{-/-}$ MEFs cell lysate or Sam68$^{-/-}$ MEFs cell lysate supplemented with an *in vitro* produced flag-tagged mSAM68 (Figure 5G). RNP complexes were pulled down using streptavidin agarose beads and treated with RppH prior to their digestion with Xrn1. Using specific primers (Figure 5E), we found that, in the presence of mSAM68 (from WT MEFs cell lysate or *in vitro* production), 5′ splice sites were highly protected from RppH/Xrn1, while the baits were completely degraded in the absence of mSAM68 (Figure 5G). Using *in vitro* produced mSAM68 and our RNA bait, we found that SAM68, by itself, could only partially protect the RNA downstream of the SAM68-binding site, while the 5′ end of the bait was completely degraded by Xrn1. This result indicates that Xrn1 accessibility at the 5′ end of our bait (5′ splice site of *mTor* intron 5) is efficiently impeded only by steric hindrance caused by the mSAM68–U1 snRNP complex. Hence, change in Xrn1 accessibility observed in Figure 5G suggests that recruitment and association of U1 snRNP is dependent on SAM68. This was further confirmed with a competing U1 snRNA antisense oligo directed against the RNA-binding site of U1 snRNP (Figure 5G, lanes 2 and 5). In both cases, the SAM68-induced U1 snRNP protection of the 5′ splice site was significantly reduced when the cell extracts were supplemented with the U1 snRNA antisense oligo in order to impair U1 snRNA hybridization with the 5′-splice site of the *mTor* RNA bait.

SAM68 recruitment of U1 snRNP is specific to endogenous *mTor* intron 5

To confirm that SAM68 binds to the endogenous intron 5 sequences of mTOR, we performed cross-linking immunoprecipitation (CLIP) on WT and *Sam68*$^{-/-}$ MEFs and assessed mSAM68 binding to different intron of mTOR pre-mRNA. mSAM68 specifically associated in intron 5 near the exon-intron junction in WT MEFs, while no binding was detected in two other introns (4 and 37), which lack mSAM68 binding sites (Supplementary Figure S6A). Furthermore, no signal was observed in the *Sam68*$^{-/-}$ background. To rescue mSAM68 binding, we performed SAM68 CLIP on *Sam68*$^{-/-}$ MEFs transfected with *mSam68(WT)*, *mSam68(ΔARM)* or *GFP* as a negative control. Enrichment of the intron 5 sequence was observed in MEFs expressing either the WT or ΔARM versions of mSAM68 but was undetectable at the other tested intron (Supplementary Figure S6B).

We then sought to determine if mSAM68 could recruit U1 snRNP on endogenous mTOR, as observed with *in vitro* synthesized baits. To address this, we performed RNA immunoprecipitation (RIP) of U1A and assessed U1 snRNP coverage of the different exon–intron junctions in both WT and *Sam68*$^{-/-}$ MEFs. No U1A signal was found at mTOR exon5–intron 5 junction (ei5) of *Sam68*$^{-/-}$ MEFs (Figure 6A), while U1A signal was easily detectable in WT MEFs, and in the other exon–intron junctions (ei4 and ei37) of both WT and *Sam68*$^{-/-}$ MEFs (Figure 6A). These results confirm that mSAM68 directly facilitates the recruitment of U1 snRNP to ei5, while not affecting other exon–intron junctions where there is no SAM68-binding site. Moreover, mSAM68 ability to recruit U1 snRNP was not shared with the mSAM68(ΔARM) mutant in a *Sam68*$^{-/-}$ MEF background, further suggesting that the recruitment of U1 snRNP to ei5 is indeed through SAM68 (Figure 6B). Accordingly, U1A immunoprecipitation in *Sam68*$^{-/-}$ MEFs expressing mSAM68(WT) or mSAM68(ΔARM) revealed that only the WT protein allowed the detection of ei5, and no change of U1 snRNP association was observed in other exon–intron junctions.

Taken together, these results confirm that SAM68 binding to its target intronic sequences is sufficient to recruit the U1 snRNP through U1A to the 5′ splice site of *mTor* intron 5 and thus facilitates its recognition and the stabilization of the U1 complex at the 5′ splice site.

SAM68 regulates *mTor* splicing through the recruitment of U1 snRNP

To determine the role of SAM68–U1snRNP interaction in the regulation of *mTor* intron 5 splicing endogenously, we performed *in vivo* splicing assay on *mTor* in WT or Sam68$^{-/-}$ MEFs, using a common forward primer and a splicing specific reverse primer (i5 versus e6) (Figure 7A). As expected, we observed a drastic decrease of $mTor_{4-6}$ amplicon upon *Sam68* inactivation, which was replaced

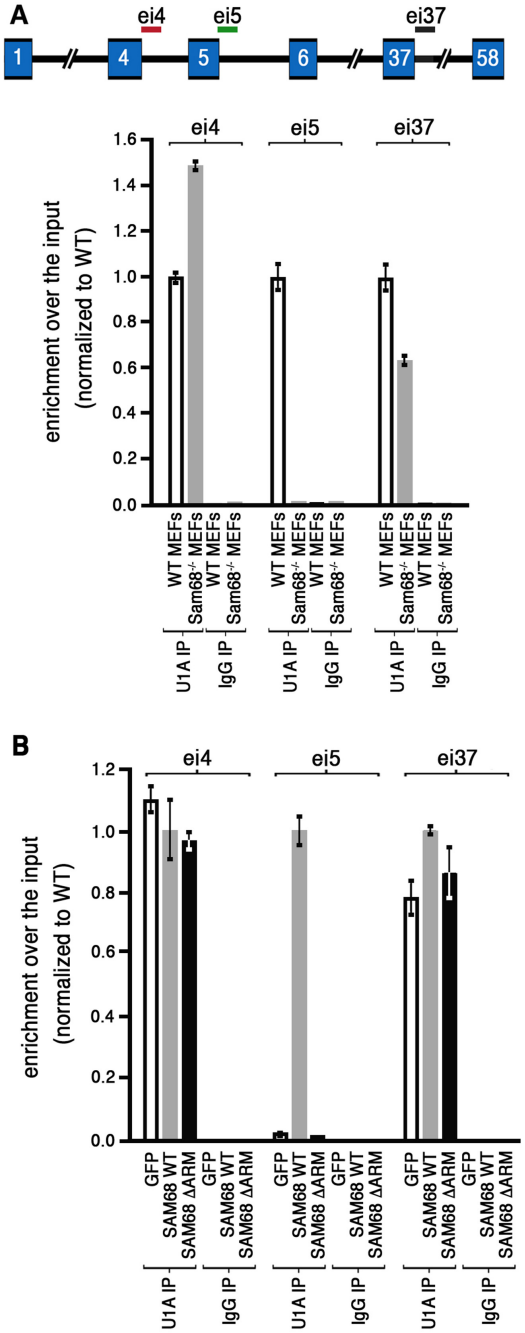

Figure 6. U1snRNP is recruited in a SAM68-dependent manner at the exon5–intron 5 junction (ei5) in *mTor* pre-mRNA. (**A**) RNA immunoprecipitation (RIP) assay of mSAM68 on *mTor* pre-mRNA. (Top) Schematic representation of *mTor* pre-mRNA showing location of amplicon used to detect U1 snRNP binding by RIP (ei4 in red, ei5 in green and ei37 in black). (Below) U1A-RIP was done from WT MEFs or *Sam68*$^{-/-}$ MEFs using anti-U1A or control IgG antibodies. Bound RNA was analyzed by RT-qPCR using the highlighted primers. Mean values are expressed as fold enrichment over input and normalized to WT signal. Error bars represent ± standard deviations of the means. (**B**) U1snRNP recruitment is restored at *mTor* EI5, in *Sam68*$^{-/-}$ MEFs expressing *mSam68(WT)* but not with *mSam68(ΔArm)*. U1A-RIP was done using anti-U1A or control IgG antibodies in *Sam68*$^{-/-}$ MEFs, *Sam68*$^{-/-}$ MEFs rescued with *mSam68(WT)* or *mSam68(ΔArm)*. Bound RNA was analyzed in triplicates by RT-qPCR using the highlighted primers. Mean values are expressed as fold enrichment over input and normalized to WT signals. Error bars represent ± standard deviations of the means.

by a robust increase of the $mTor_{i5}$ transcript (Figure 7B), confirming our previous observation that *Sam68* depletion leads to increased intron 5-induced termination (21). To rescue Sam68 depletion splicing effects, *Sam68*$^{-/-}$ MEFs were transduced with either *Sam68(WT)* or *Sam68(ΔARM)*. Total RNA was isolated 48 h post transfection and analyzed by semi-quantitative RT-PCR. As expected, upon expression of mSAM68(WT), there was a significant decrease in intron 5-induced premature termination and polyadenylation (Figure 7B). On the other hand, cells that expressed mSAM68(ΔARM) were unable to rescue the splicing defect. This result confirms that reduced association with U1A leads to U1 snRNP recruitment impairments and thus, *mTor* intron 5 proper splicing. While robust level of intron 5 detection was still observed in cells expressing mSAM68(ΔARM), levels were lower than vector control (Figure 7B). This partial decrease in intron 5-induced termination is assumingly provoked by the remaining hSAM68(ΔARM)–U1A interaction, as observed in Figure 4C. As proposed in our previous work using the 3T3-L1 cell lines, this SAM68 splicing defect was also associated with decreased *mTor* expression level in MEFs. Indeed, mTOR protein level was highly decreased in *Sam68*$^{-/-}$ MEFs and *Sam68*$^{-/-}$ MEFs expressing mSAM68(ΔARM), when compared to WT MEFs or *Sam68*$^{-/-}$ MEFs expressing mSAM68(WT) (Supplementary Figure S7).

These results were also observed in an *in vivo* splicing assays using the *mTor* minigene, a plasmid that drives the expression of the 2.3 kb genomic fragment spanning from exon4 to exon6 of mouse *mTor* (Supplementary Figure S7A). Indeed, hSAM68-depleted HEK-293T cells cotransfected with the minigene and either Flag-YFP, Flag-hSAM68(WT) or Flag-hSAM68(ΔARM) showed similar results that Flag-hSAM68(WT) was able to revert the increased $mTor_{i5}$ / $mTor_{tot}$ ratio, while the Flag-hSAM68(ΔARM) behaved like Flag-YFP (Supplementary Figure S6B).

DISCUSSION

The KH domain RNA-binding protein, SAM68, regulates splicing of *mTor* as well as the ribosomal S6 kinase *(Rps6kb1)* transcripts in pre-adipocytes (21,39). In turn, pre-adipocytes of *Sam68*$^{-/-}$ mice do not differentiate to adipocytes due to defective mTOR signaling. Our data show that SAM68 modulates *mTor* splicing by binding to specific regulatory elements found in intron 5 (SB-1 and SB-A), of which SB-A overlaps with the poly-adenylation signal (AAUAAA). This led us to postulate that these AU-rich *cis*-acting elements were intronic splicing enhancers to which SAM68 bound with great affinity to modulate the recruitment of U1 snRNP at the upstream 5′ splice site.

In this study, we report that SAM68 functionally interacts with U1 snRNP, the spliceosomal component that recognizes 5′ splice sites (10). While SAM68 can bind with low affinity to the U1 snRNA *in vitro*, it is highly unlikely that it will happen *in vivo*. Furthermore, the potential AU-rich SAM68 binding motif found in the U1 snRNA is located in the Sm site, which is rapidly masked by the Sm protein ring during U1 snRNP assembly (49,62). Rather, we found that SAM68 interacts directly with U1A, the stem–

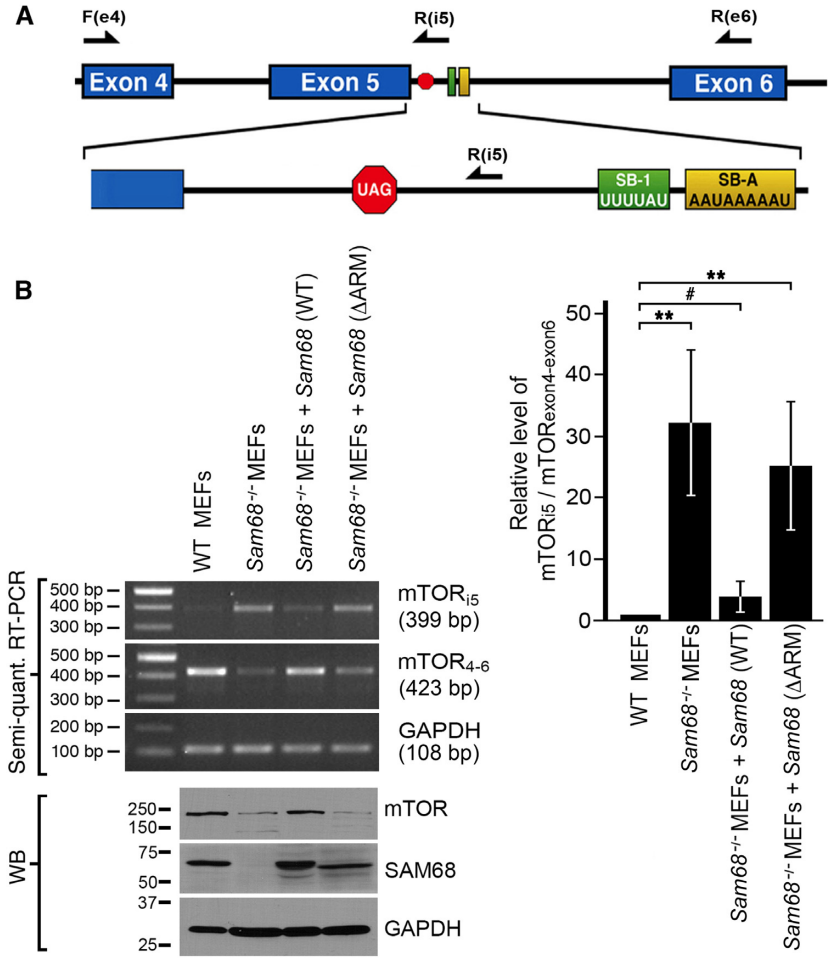

Figure 7. SAM68 deletion of 'ARM binding region' shows decrease in U1A binding. (**A**) Schematic of the pcDNA $mTor_{4–6}$ minigene, comprising the *mTor* genomic fragment from exon4 to exon6. (**B**) (Left panel) $Sam68^{-/-}$ MEFs cells were infected with *Sam68*(WT) or *Sam68*(ΔARM) and compared to uninfected $Sam68^{-/-}$ MEFs or WT MEFs. Total RNA was extracted in each cell lines and semi-quantitative RT-PCRs were performed using endogenous *mTor* specific primers. Forward (Fe4) and Reverse (Re6) were used to quantify *mTor* transcripts that were spliced normally ($mTor_{4–6}$), while Forward (Fe4) and Reverse (Ri5) were used to quantify intron 5 including *mTor* transcripts ($mTor_{i5}$). *Gapdh* was used to normalize the values obtained. Total protein was also extracted and run on 10% SDS-PAGE and blotted with SAM68, U1A and GAPDH antibodies. (Right panel) Quantification of intron 5-induced termination over normally spliced mRNA based on three independent experiments using endogenous *mTor* specific primers. $^{**}P$-value ≤ 0.05 and # = non-significant (two-tailed *t*-test).

loop II binding protein of U1 snRNA. This SAM68–U1A interaction was shown to be resistant to RNaseA treatment and RNA-binding defective mutant versions of SAM68 bearing point mutations in the KH domain; $SAM68_{V229F}$ (23) and $SAM68_{I184N}$ (36,52) could still bind U1A and the U1snRNP. Domain mapping studies showed that U1A binds to the 'minimal ARM binding region of SAM68 and that deleting this region (379–389 aa) in SAM68(WT) was sufficient to impair U1A association. This region located within SAM68 YY domain was initially identified to regulate T-cell factor 1 splicing by binding the armadillo repeat (ARM) domain of Adenomatous polyposis coli (APC) (54,55). Hence, our results confirmed that SAM68 can directly interact with U1A, making it the only identified protein interactor of this U1 snRNP core protein to date. NMR spectroscopy also confirmed that this protein–protein interaction is mediated through the tyrosine-rich 'ARM-binding domain' of SAM68 and the RRM1 domain of U1A. Conversely, we observed that deleting the ARM domain of SAM68 greatly impaired its association with U1A, resulting in increased intron 5 inclusion. Residual splicing activity could still be observed, which might be due to the remaining YXXY motifs in SAM68(ΔARM) that may act as weak surrogate binding sites in the absence of the ARM motif.

Results obtained with synthetized RNA baits suggest that mSAM68 initiate U1 snRNP recruitment, which then allows the recognition of the sub-optimal splice site. Hence, the presence of mSAM68 increases the recognition rate of the 5′ splice site of *mTor* intron 5, by increasing U1 snRNP stoichiometry close to the splice site. This was further confirmed by the endogenous mSAM68 CLIP, the U1A RNA immunoprecipitation assays and the *mTor* splicing disparity observed using either the *mTor* $minigene_{exon4–6}$ or the endogenous *mTor*. These findings highlight the importance of intronic enhancer mediated binding of SAM68, subsequent recruitment of U1snRNP via U1A and the resultant

splicing of *mTor* intron 5. This is strikingly similar to the mechanism by which other RNA-binding proteins, such as TIA1 and RBM24, modulate 5′ splice site recognition of Drosophila male specific lethal (msl) gene and inhibitor of κ light polypeptide gene enhancer in B cells, kinase complex-associated protein (IKBKAP) gene in familial dysautonomia (FD) by interacting with U1snRNP (17,63). Indeed, both of these RNA-binding proteins were shown to promote splicing of pre-mRNA's containing sub-optimal splicing signals by interacting with components of U1 snRNP. While TIA1 interacts with U1C and SAM68 with U1A, both act as intronic splicing enhancers by binding sequences close to the 5′ splice sites, thus facilitating U1 snRNP recruitment.

As shown here, SAM68 seems to initiate spliceosome assembly at the 5′ splice site of *mTor* intron 5, through the recruitment the U1 snRNP. This is mainly achieved when SAM68 binds specific sequence within the 5′ portion of mTor intron 5, mainly the bipartite sequences designated as SB-1 and SB-A. While these two sequences are clearly involved in U1 snRNP recruitment, the role of the downstream SAM68-binding sites found throughout intron 5 remains to be defined. One main hypothesis could be that SAM68 regulates early steps of spliceosome assembly (e.g. complex E formation) in specific introns. This seems to be in line with the observation that SAM68 binds U2AF65, a subunit of the U2AF complex known to initiate 3′ splice site recognition (64). Moreover SAM68 binding to U2AF65 increased its efficiency. As SAM68 homodimerizes through its association with RNA, it is also possible that multiple binding sites favor its dimerization, thus decreasing the gap between the 5′ and the 3′ splice site, which would again facilitate early spliceosome assembly.

It has been recently reported that introns with weak 5′ splice sites are susceptible to premature cleavage and polyadenylation events in the absence of U1 snRNP binding and RNA-binding proteins are required for U1 snRNP occupancy at such sites (65). Indeed, morpholino oligonucleotides interfering with U1 snRNA binding causes premature cleavage and polyadenylation in numerous pre-mRNAs at cryptic polyA sites in introns and this occurs near the start of the transcripts (66). This mechanism seems to involve core protein components of the U1 snRNP, since both U1–70K and U1A were reported to inhibit polyadenylation by mediating a direct interaction with poly-A-polymerase (67,68). This U1-mediated inhibition of polyadenylation even led to the characterization of a Polyadenylation inhibitory Elements (69). Concomitantly, SAM68 has also recently been reported to regulate 3′ end processing of *Aldh1a3* pre-mRNA, a mechanism that is required to maintain glycolytic metabolism and self-renewal of mouse neural progenitor cells (37). While our results show that mSAM68 recruits U1 snRNP to the 5′ splice site of *mTor* intron 5 and increases the rate of intron 5 excision, it also seems to cooperate in masking cryptic polyadenylation sites, a process that is known to be mediated by U1 snRNP (66,70). Indeed, SAM68 alone (without U1 snRNP) is unable to modulate alternative intronic polyadenylation usage, since SAM68(ΔARM) can bind the poly(A) signal as efficiently as the WT version, but still allows polyadenylation of the $mTor_{i5}$ transcript. Hence, these results suggest that 5′ splice site recognition by the SAM68–U1 snRNP complex precludes intronic cryptic polyadenylation usage, while allowing proper excision of *mTor* intron 5. This also suggest that SAM68 could also be involved in the modulation of the previously described U1-dependent modulation of alternative poly(A) usage (37,66,70,71).

Our findings illustrate the first mechanistic evidence showing that SAM68 recruits U1 snRNP via direct interaction with U1A to upstream 5′ splice site, while it binds AU-rich regulatory sequences in *mTor* intron 5, through its RNA-binding KH domain. This is subsequently mediated through its tyrosine-rich 'ARM-binding domain', and in turn enhances splicing of intron 5, while SAM68 depletion inhibits U1 snRNP recruitment and promotes premature intron 5-induced termination and polyadenylation. Although the regulatory mechanism of this interaction remains elusive, the presence of numerous tyrosine residues located within the 'YY region' of SAM68 suggest that phosphorylation might be involved. Indeed, SAM68 tyrosine phosphorylation was shown to negatively affect its ability to bind RNA (31,33,34), and its splicing functions (i.e. *Bcl-X*, *CD-44* and *Cyclin D1*) (22,23,33).

In the future, the global identification of mRNA targets regulated by this SAM68-dependent alternative splicing should allow us to better understand these possible regulatory mechanisms and determine how a binding site situated 174-nt downstream of the 5′ splice site can modulate its recognition by the U1 snRNP. Indeed, it would be interesting to assess if other alternatively spliced mRNAs targeted by SAM68 could be differentially regulated through distance changes between SAM68-binding sites and the 5′ splice site during cellular processes such as development or cancer. Moreover, determining how the SAM68–U1A interaction is modulated by distance constraints, secondary structures or associated proteins to regulate not only *mTor*, but also a specific sub-class of mRNA, remains to be investigated.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Stéphane Richard, Dr Roberto Bonasio and Dr Philip Leder for sharing reagents. We thank Dr Rachid Mazroui and Dr Martin Simard for critical reading of the manuscript. We are grateful to Yan Coulombe and Maripier Hainse for technical assistance. We thank Jonathan Bergeman and Carl St-Pierre from the Cell Imaging Unit of the Centre de recherche du CHU de Québec-Université Laval for their technical assistance.

FUNDING

Fonds de Recherche du Québec - Santé (FRQ-S) (to M.-E., S.M.H., J.-Y.M.); Bourse de la Fondation du CHU de Québec (to J.O's. and V.F.); Natural Sciences and Engineering Research Council of Canada (NSERC) [402880-2012 to M.-E.H.; 2016-05847 to S.M.H.], NCCR RNA and Disease of the Swiss National Foundation (to F.H.T.A.); Canadian

Institutes of Health Research (to J.-Y.M.). Funding for open access charge: Natural Sciences and Engineering Research Council of Canada (NSERC) [402880-2012 to M.-E.H.].
Conflict of interest statement. None declared.

REFERENCES

1. Wahl,M.C., Will,C.L. and Luhrmann,R. (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell*, **136**, 701–718.
2. Faustino,N.A. and Cooper,T.A. (2003) Pre-mRNA splicing and human disease. *Genes Dev.*, **17**, 419–437.
3. Matera,A.G. and Wang,Z. (2014) A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.*, **15**, 108–121.
4. Papasaikas,P. and Valcarcel,J. (2016) The Spliceosome: the ultimate RNA chaperone and sculptor. *Trends Biochem. Sci.*, **41**, 33–45.
5. Baralle,F.E. and Giudice,J. (2017) Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.*, **18**, 437–451.
6. Chabot,B. and Steitz,J.A. (1987) Recognition of mutant and cryptic 5′ splice sites by the U1 small nuclear ribonucleoprotein in vitro. *Mol. Cell. Biol.*, **7**, 698–707.
7. Mount,S.M., Pettersson,I., Hinterberger,M., Karmas,A. and Steitz,J.A. (1983) The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site in vitro. *Cell*, **33**, 509–518.
8. Alibert,C., Tazi,J., Temsamani,J., Jeanteur,P., Brunel,C. and Cathala,G. (1990) Interplay between U2 snRNP and 3′ splice factor(s) for branch point selection on human beta-globin pre-mRNA. *Nucleic Acids Res.*, **18**, 235–245.
9. Ruskin,B., Zamore,P.D. and Green,M.R. (1988) A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell*, **52**, 207–219.
10. Shao,W., Kim,H.S., Cao,Y., Xu,Y.Z. and Query,C.C. (2012) A U1-U2 snRNP interaction network during intron definition. *Mol. Cell. Biol.*, **32**, 470–478.
11. Jurica,M.S. and Moore,M.J. (2003) Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell*, **12**, 5–14.
12. Roca,X., Krainer,A.R. and Eperon,I.C. (2013) Pick one, but be quick: 5′ splice sites and the problems of too many choices. *Genes Dev.*, **27**, 129–144.
13. Sun,S., Ling,S.C., Qiu,J., Albuquerque,C.P., Zhou,Y., Tokunaga,S., Li,H., Qiu,H., Bui,A., Yeo,G.W. *et al.* (2015) ALS-causative mutations in FUS/TLS confer gain and loss of function by altered association with SMN and U1-snRNP. *Nat. Commun.*, **6**, 6171.
14. Rogelj,B., Easton,L.E., Bogu,G.K., Stanton,L.W., Rot,G., Curk,T., Zupan,B., Sugimoto,Y., Modic,M., Haberman,N. *et al.* (2012) Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.*, **2**, 603.
15. Mayeda,A. and Krainer,A.R. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*, **68**, 365–375.
16. Krainer,A.R., Conway,G.C. and Kozak,D. (1990) The essential pre-mRNA splicing factor SF2 influences 5′ splice site selection by activating proximal sites. *Cell*, **62**, 35–42.
17. Forch,P., Puig,O., Martinez,C., Seraphin,B. and Valcarcel,J. (2002) The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5′ splice sites. *EMBO J.*, **21**, 6882–6892.
18. Ohe,K., Yoshida,M., Nakano-Kobayashi,A., Hosokawa,M., Sako,Y., Sakuma,M., Okuno,Y., Usui,T., Ninomiya,K., Nojima,T. *et al.* (2017) RBM24 promotes U1 snRNP recognition of the mutated 5′ splice site in the IKBKAP gene of familial dysautonomia. *RNA*, **23**, 1393–1403.
19. Geuens,T., Bouhy,D. and Timmerman,V. (2016) The hnRNP family: insights into their role in health and disease. *Hum. Genet.*, **135**, 851–867.
20. Mayeda,A., Munroe,S.H., Caceres,J.F. and Krainer,A.R. (1994) Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J.*, **13**, 5483–5495.
21. Huot,M.E., Vogel,G., Zabarauskas,A., Ngo,C.T., Coulombe-Huntington,J., Majewski,J. and Richard,S. (2012) The Sam68 STAR RNA-binding protein regulates mTOR alternative splicing during adipogenesis. *Mol. Cell*, **46**, 187–199.
22. Paronetto,M.P., Cappellari,M., Busa,R., Pedrotti,S., Vitali,R., Comstock,C., Hyslop,T., Knudsen,K.E. and Sette,C. (2010) Alternative splicing of the cyclin D1 proto-oncogene is regulated by the RNA-binding protein Sam68. *Cancer Res.*, **70**, 229–239.
23. Paronetto,M.P., Achsel,T., Massiello,A., Chalfant,C.E. and Sette,C. (2007) The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J. Cell Biol.*, **176**, 929–939.
24. Matter,N., Herrlich,P. and Konig,H. (2002) Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature*, **420**, 691–695.
25. Bielli,P., Busa,R., Paronetto,M.P. and Sette,C. (2011) The RNA-binding protein Sam68 is a multifunctional player in human cancer. *Endocr. Relat. Cancer*, **18**, R91–R102.
26. Taylor,S.J., Anafi,M., Pawson,T. and Shalloway,D. (1995) Functional interaction between c-Src and its mitotic target, Sam 68. *J. Biol. Chem.*, **270**, 10120–10124.
27. Itoh,M., Haga,I., Li,Q.H. and Fujisawa,J. (2002) Identification of cellular mRNA targets for RNA-binding protein Sam68. *Nucleic Acids Res.*, **30**, 5452–5464.
28. Chen,T., Damaj,B.B., Herrera,C., Lasko,P. and Richard,S. (1997) Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain. *Mol. Cell. Biol.*, **17**, 5707–5718.
29. Frisone,P., Pradella,D., Di Matteo,A., Belloni,E., Ghigna,C. and Paronetto,M.P. (2015) SAM68: Signal transduction and RNA metabolism in human cancer. *Biomed. Res. Int.*, **2015**, 528954.
30. Lukong,K.E. and Richard,S. (2003) Sam68, the KH domain-containing superSTAR. *Biochim. Biophys. Acta*, **1653**, 73–86.
31. Meyer,N.H., Tripsianes,K., Vincendeau,M., Madl,T., Kateb,F., Brack-Werner,R. and Sattler,M. (2010) Structural basis for homodimerization of the Src-associated during mitosis, 68-kDa protein (Sam68) Qua1 domain. *J. Biol. Chem.*, **285**, 28893–28901.
32. Perez-Perez,A., Sanchez-Jimenez,F., Vilarino-Garcia,T., de la Cruz,L., Virizuela,J.A. and Sanchez-Margalet,V. (2016) Sam68 mediates the activation of insulin and leptin signalling in breast cancer cells. *PLoS One*, **11**, e0158218.
33. Huot,M.E., Brown,C.M., Lamarche-Vane,N. and Richard,S. (2009) An adaptor role for cytoplasmic Sam68 in modulating Src activity during cell polarization. *Mol. Cell. Biol.*, **29**, 1933–1943.
34. Lukong,K.E., Larocque,D., Tyner,A.L. and Richard,S. (2005) Tyrosine phosphorylation of sam68 by breast tumor kinase regulates intranuclear localization and cell cycle progression. *J. Biol. Chem.*, **280**, 38639–38647.
35. Lin,Q., Taylor,S.J. and Shalloway,D. (1997) Specificity and determinants of Sam68 RNA binding. Implications for the biological function of K homology domains. *J. Biol. Chem.*, **272**, 27274–27280.
36. Chawla,G., Lin,C.H., Han,A., Shiue,L., Ares,M. Jr and Black,D.L. (2009) Sam68 regulates a set of alternatively spliced exons during neurogenesis. *Mol. Cell. Biol.*, **29**, 201–213.
37. La Rosa,P., Bielli,P., Compagnucci,C., Cesari,E., Volpe,E., Farioli Vecchioli,S. and Sette,C. (2016) Sam68 promotes self-renewal and glycolytic metabolism in mouse neural progenitor cells by modulating Aldh1a3 pre-mRNA 3′-end processing. *Elife*, **5**, e20750.
38. Li,N., Hebert,S., Song,J., Kleinman,C.L. and Richard,S. (2017) Transcriptome profiling in preadipocytes identifies long noncoding RNAs as Sam68 targets. *Oncotarget*, **8**, 81994–82005.
39. Song,J. and Richard,S. (2015) Sam68 regulates S6K1 alternative splicing during adipogenesis. *Mol. Cell. Biol.*, **35**, 1926–1939.
40. Huot,M.E. and Richard,S. (2012) Stay lean without dieting: Lose Sam68. *Adipocyte*, **1**, 246–249.
41. Paronetto,M.P., Messina,V., Bianchi,E., Barchi,M., Vogel,G., Moretti,C., Palombi,F., Stefanini,M., Geremia,R., Richard,S. *et al.* (2009) Sam68 regulates translation of target mRNAs in male germ cells, necessary for mouse spermatogenesis. *J. Cell. Biol.*, **185**, 235–249.
42. Venables,J.P., Vernet,C., Chew,S.L., Elliott,D.J., Cowmeadow,R.B., Wu,J., Cooke,H.J., Artzt,K. and Eperon,I.C. (1999) T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis. *Hum. Mol. Genet.*, **8**, 959–969.
43. Valacca,C., Bonomi,S., Buratti,E., Pedrotti,S., Baralle,F.E., Sette,C., Ghigna,C. and Biamonti,G. (2010) Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene. *J. Cell. Biol.*, **191**, 87–99.
44. Pedrotti,S., Bielli,P., Paronetto,M.P., Ciccosanti,F., Fimia,G.M., Stamm,S., Manley,J.L. and Sette,C. (2010) The splicing regulator

Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy. *EMBO J.*, **29**, 1235–1247.
45. Laplante,M. and Sabatini,D.M. (2012) mTOR signaling in growth control and disease. *Cell*, **149**, 274–293.
46. Johnson,S.C., Rabinovitch,P.S. and Kaeberlein,M. (2013) mTOR is a key modulator of ageing and age-related disease. *Nature*, **493**, 338–345.
47. Avis,J.M., Allain,F.H., Howe,P.W., Varani,G., Nagai,K. and Neuhaus,D. (1996) Solution structure of the N-terminal RNP domain of U1A protein: the role of C-terminal residues in structure stability and RNA binding. *J. Mol. Biol.*, **257**, 398–411.
48. Pagliarini,V., Pelosi,L., Bustamante,M.B., Nobili,A., Berardinelli,M.G., D'Amelio,M., Musaro,A. and Sette,C. (2015) SAM68 is a physiological regulator of SMN2 splicing in spinal muscular atrophy. *J. Cell. Biol.*, **211**, 77–90.
49. Stark,H., Dube,P., Luhrmann,R. and Kastner,B. (2001) Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature*, **409**, 539–542.
50. Consortium,E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
51. Savarese,E., Chae,O.W., Trowitzsch,S., Weber,G., Kastner,B., Akira,S., Wagner,H., Schmid,R.M., Bauer,S. and Krug,A. (2006) U1 small nuclear ribonucleoprotein immune complexes induce type I interferon in plasmacytoid dendritic cells through TLR7. *Blood*, **107**, 3229–3234.
52. De Boulle,K., Verkerk,A.J., Reyniers,E., Vits,L., Hendrickx,J., Van Roy,B., Van den Bos,F., de Graaff,E., Oostra,B.A. and Willems,P.J. (1993) A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat. Genet.*, **3**, 31–35.
53. Satoh,M., Richards,H.B., Hamilton,K.J. and Reeves,W.H. (1997) Human anti-nuclear ribonucleoprotein antigen autoimmune sera contain a novel subset of autoantibodies that stabilizes the molecular interaction of U1RNP-C protein with the Sm core proteins. *J. Immunol.*, **158**, 5017–5025.
54. Zhang,Z., Akyildiz,S., Xiao,Y., Gai,Z., An,Y., Behrens,J. and Wu,G. (2015) Structures of the APC-ARM domain in complexes with discrete Amer1/WTX fragments reveal that it uses a consensus mode to recognize its binding partners. *Cell. Discov.*, **1**, 15016.
55. Morishita,E.C., Murayama,K., Kato-Murayama,M., Ishizuka-Katsura,Y., Tomabechi,Y., Hayashi,T., Terada,T., Handa,N., Shirouzu,M., Akiyama,T. *et al.* (2011) Crystal structures of the armadillo repeat domain of adenomatous polyposis coli and its complex with the tyrosine-rich domain of Sam68. *Structure*, **19**, 1496–1508.
56. Feracci,M., Foot,J.N., Grellscheid,S.N., Danilenko,M., Stehle,R., Gonchar,O., Kang,H.S., Dalgliesh,C., Meyer,N.H., Liu,Y. *et al.* (2016) Structural basis of RNA recognition and dimerization by the STAR proteins T-STAR and Sam68. *Nat. Commun.*, **7**, 10355.
57. Galarneau,A. and Richard,S. (2009) The STAR RNA binding proteins GLD-1, QKI, SAM68 and SLM-2 bind bipartite RNA motifs. *BMC Mol. Biol.*, **10**, 47.
58. Nagarajan,V.K., Jones,C.I., Newbury,S.F. and Green,P.J. (2013) XRN 5′→3′ exoribonucleases: structure, mechanisms and functions. *Biochim. Biophys. Acta*, **1829**, 590–603.
59. Chapman,E.G., Moon,S.L., Wilusz,J. and Kieft,J.S. (2014) RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA. *Elife*, **3**, e01892.
60. Braun,J.E., Truffault,V., Boland,A., Huntzinger,E., Chang,C.T., Haas,G., Weichenrieder,O., Coles,M. and Izaurralde,E. (2012) A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5′ exonucleolytic degradation. *Nat. Struct. Mol. Biol.*, **19**, 1324–1331.
61. Jinek,M., Coyle,S.M. and Doudna,J.A. (2011) Coupled 5′ nucleotide recognition and processivity in Xrn1-mediated mRNA decay. *Mol. Cell*, **41**, 600–608.
62. Raker,V.A., Hartmuth,K., Kastner,B. and Luhrmann,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.
63. Ohe,K., Yoshida,M., Nakano-Kobayashi,A., Hosokawa,M., Sako,Y., Sakuma,M., Okuno,Y., Usui,T., Ninomiya,K., Nojima,T. *et al.* (2017) RBM24 promotes U1 snRNP recognition of the mutated 5′ splice site in the IKBKAP gene of familial dysautonomia. *RNA*, **23**, 1393–1403.
64. Tisserant,A. and Konig,H. (2008) Signal-regulated Pre-mRNA occupancy by the general splicing factor U2AF. *PLoS One*, **3**, e1418.
65. Tian,B., Pan,Z. and Lee,J.Y. (2007) Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. *Genome Res.*, **17**, 156–165.
66. Kaida,D., Berg,M.G., Younis,I., Kasim,M., Singh,L.N., Wan,L. and Dreyfuss,G. (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*, **468**, 664–668.
67. Gunderson,S.I., Polycarpou-Schwarz,M. and Mattaj,I.W. (1998) U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell*, **1**, 255–264.
68. Gunderson,S.I., Beyer,K., Martin,G., Keller,W., Boelens,W.C. and Mattaj,L.W. (1994) The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell*, **76**, 531–541.
69. Klein Gunnewiek,J.M., Hussein,R.I., van Aarssen,Y., Palacios,D., de Jong,R., van Venrooij,W.J. and Gunderson,S.I. (2000) Fourteen residues of the U1 snRNP-specific U1A protein are required for homodimerization, cooperative RNA binding, and inhibition of polyadenylation. *Mol. Cell. Biol.*, **20**, 2209–2217.
70. Berg,M.G., Singh,L.N., Younis,I., Liu,Q., Pinto,A.M., Kaida,D., Zhang,Z., Cho,S., Sherrill-Mix,S., Wan,L. *et al.* (2012) U1 snRNP determines mRNA length and regulates isoform expression. *Cell*, **150**, 53–64.
71. Workman,E., Veith,A. and Battle,D.J. (2014) U1A regulates 3′ processing of the survival motor neuron mRNA. *J. Biol. Chem.*, **289**, 3703–3712.