Published online 9 January 2023

Nucleic Acids Research, 2023, Vol. 51, No. 3 ***1393–1408***
https://doi.org/10.1093/nar/gkac1250

The hnRNP C tetramer binds to CBC on mRNA and impedes PHAX recruitment for the classification of RNA polymerase II transcripts

Sayaka Dantsuji[1], Mutsuhito Ohno[1] and Ichiro Taniguchi [1,2,*]

[1]Institute for Life and Medical Sciences, Kyoto University, Kyoto, Kyoto 606-8507, Japan and [2]Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan

Received March 26, 2022; Revised December 11, 2022; Editorial Decision December 12, 2022; Accepted December 15, 2022

ABSTRACT

In eukaryotic cells, various classes of RNAs are exported to the cytoplasm by class-specific factors. Accumulating evidence has shown that export factors affect the fate of RNA, demonstrating the importance of proper RNA classification upon export. We previously reported that RNA polymerase II transcripts were classified after synthesis depending on their length, and identified heterogeneous nuclear ribonucleoprotein (hnRNP) C as the key classification factor. HnRNP C inhibits the recruitment of PHAX, an adapter protein for spliceosomal U snRNA export, to long transcripts, navigating these RNAs to the mRNA export pathway. However, the mechanisms by which hnRNP C inhibits PHAX recruitment to mRNA remain unknown. We showed that the cap-binding complex, a bridging factor between m^7G-capped RNA and PHAX, directly interacted with hnRNP C on mRNA. Additionally, we revealed that the tetramer-forming activity of hnRNP C and its strong RNA-binding activity were crucial for the inhibition of PHAX binding to longer RNAs. These results suggest that mRNA is wrapped around the hnRNP C tetramer without a gap from the cap, thereby impeding the recruitment of PHAX. The results obtained on the mode of length-specific RNA classification by the hnRNP C tetramer will provide mechanistic insights into hnRNP C-mediated RNA biogenesis.

INTRODUCTION

In eukaryotic cells, different classes of RNAs are exported from the nucleus after synthesis by distinct sets of export factors (1). The composition of the export complex influences the fate of the RNA, such as its functions, stability and cytoplasmic localization, demonstrating the importance of proper RNA classification upon export (2–7). Nascent spliceosomal U small nuclear RNA (snRNA) and mRNA, which are synthesized by RNA polymerase II, share common features at their 5′ termini; they receive a 7-methyl guanosine (m^7G) cap structure. The cap structure recruits the cap-binding complex (CBC) and induces their nuclear export (8,9). Notably, different export factors for U snRNA and mRNA are recruited for their export (Figure 1A). In U snRNA export, the adapter protein PHAX is recruited to the cap-proximal region through a direct interaction with both CBC and RNA (10). PHAX subsequently recruits the nuclear export receptor CRM1-RanGTP (10). In bulk mRNA export, the transcription and export (TREX) complex, rather than PHAX, is recruited to the same region through an interaction with both CBC and RNA (11). SR proteins and the exon junction complex (EJC) are also recruited to mRNA (12,13). These complexes/proteins in turn recruit the mRNA export receptor TAP/NXF1 (14,15). U snRNA export factors, such as PHAX, do not bind to mRNA, despite the common cap structure (2,10), suggesting the existence of other features that distinguish these two classes of RNAs.

We previously identified RNA length as one of the features that distinguish classes of RNAs for correct nuclear export (2,16,17). When different lengths of unstructured exonic sequences were artificially inserted into U snRNA, PHAX binding to elongated RNA was inhibited and the elongated RNA was exported by the mRNA export machinery. When intronless mRNA was artificially truncated, the shortened mRNA was bound by PHAX and exported by the U snRNA export machinery. The threshold length of export pathway switching was ∼200–300 nucleotides (nt). We also identified the heterogeneous nuclear ribonucleoprotein (hnRNP) C1/C2 as a key classification factor (18). HnRNP C selectively bound to unstructured RNA regions >200–300 nt and inhibited PHAX binding to these RNAs. We further found that hnRNP C directly interacted with CBC. These findings led us to propose the following model for the inhibition of PHAX recruitment onto long transcripts by hnRNP C (18) (Figure 1A). When a transcript

*To whom correspondence should be addressed. Email: taniguchi.ichiro.fbs@osaka-u.ac.jp

of >200–300 nt is synthesized, hnRNP C stably binds to the RNA, and the RNA-bound hnRNP C interacts with CBC. This cap-proximal RNP formation consequently inhibits PHAX recruitment to RNA, leading to mRNA-type nuclear export. When a transcript is <200–300 nt, hnRNP C cannot stably bind to the transcript, thereby allowing PHAX recruitment and subsequent U snRNA-type export. However, the molecular mechanisms underlying the PHAX-inhibitory and length-specific RNA-binding activities of hnRNP C have not yet been elucidated.

HnRNP C1/C2 is one of the most abundant pre-mRNA-binding proteins in the nucleus (19), which, together with hnRNP A/B proteins, packages nascent RNA into a repeating array of regular particles—40S hnRNP particles (20–23). In addition to RNA packaging, hnRNP C1/C2 has important functions in mRNA maturation, such as in the regulation of splicing (24–26) and RNA sorting (18). C1 and C2 are alternatively spliced gene products that only differ by the presence of 13 amino acids in C2 (27). HnRNP C contains several functional regions: the RRM, BASIC, ZIPPER and ACIDIC regions (Figure 1B). Its RNA-binding activity is mainly attributed to two distinct RNA-binding regions: a consensus RNA recognition motif (RRM) and a region rich in basic residues (BASIC). The well-conserved RRM, also referred to as the RNA-binding domain (RBD), is a ∼90-amino-acid domain located in the N-terminus of hnRNP C. The RRM has two conserved octa- and hexapeptide sequences termed RNP1 and RNP2 (28,29). HnRNP C preferentially binds to uridine stretches via the two RNPs (30–32). An auxiliary region (AUX) is located between the RRM and BASIC regions. The BASIC region provides additional sequence-independent RNA-binding activity (33–35); this region of 40 amino acids consists of 11 basic residues (Arg/Lys) and no acidic residues (36). Together with the adjacent leucine zipper (ZIPPER) domain, the BASIC region forms a bZIP-like motif, as seen in transcription factors (25). HnRNP C1 and C2 form a 3:1 heterotetramer (37,38). The tetramer is highly stable in solution because of the ZIPPER domain (35,39); it retains its tetramer structure in 2 M NaCl and does not disassociate at pH 5.5 or 11.5 (37). This stability is supported by the presence of hydrophobic residues at every seventh position in 30 amino acids (Figure 2B) and its location at the center of the tetramer complex, forming an antiparallel four-helix coiled-coil structure (27). Additionally, the C-terminal acidic region (ACIDIC) stabilizes tetramer formation (35,39). One tetramer binds RNA 230–240 nt in length (23,38), which notably is the threshold length of export pathway switching between U snRNA and mRNA (16). While it is possible that tetramer formation contributes to the binding to long RNAs, there is no evidence to prove this possibility.

To obtain a more detailed understanding of RNA classification mechanisms, we investigated the regions in hnRNP C that are essential for the correct classification of RNAs for nuclear export. We found that the interaction between CBC and hnRNP C occurred on RNA, and an N-terminal region was necessary for the direct interaction. We also showed that the BASIC and ZIPPER regions were both required for hnRNP C to inhibit PHAX binding to longer RNAs and that these regions contributed to strong binding to longer RNAs. Furthermore, our results suggested that the formation of tetramers was crucial for the RNA-binding activity of hnRNP C both *in vitro* and *in vivo*. From the current results, the following model was proposed: RNA longer than a certain threshold is wrapped around the tetramer without a gap from the cap through a direct interaction with CBC via the N-terminal region, thereby inhibiting the access of PHAX. The present results demonstrate how hnRNP C inhibits PHAX recruitment by interacting with CBC in an RNA-length-specific manner.

MATERIALS AND METHODS

DNA constructs

To generate hnRNP C1 plasmids, the full-length hnRNP C1 fragment was inserted into the BamHI–XhoI sites of pGEX-6p-1 (pGEX-6p-1-C1) or pET28a (pET28a-C1) (18). For the hnRNP C1 mutants N-half, C-half, Middle, RRM, AUX, BASIC, ΔRRM, ΔBASIC, ΔRRMΔBASIC, ΔZIPPER and ZIP-MT4, pGEX-6p-1-C1 and pET28a-C1 were used as the templates for mutagenesis. To generate the Raly plasmid, the Raly fragment was inserted into the BamHI–XhoI sites of pET28a (pET28a-Raly). To generate plasmids to express chimeric proteins, pET28a-C1 and pET28a-Raly were used as the template for the polymerase chain reaction (PCR).

Recombinant proteins

Recombinant His-CBC (His-CBP80 and His-CBP20) and glutathione *S*-transferase (GST)–PHAX were prepared as previously described (18). To purify His-T7-C1 wild-type (WT) and mutant proteins, plasmids were first transformed into *Escherichia coli* BL21-Gold (DE3). The expressed protein in the bacterial lysate was bound to Ni Sepharose beads (GE Healthcare) at 4°C for 1–3 h. Bound beads were washed five times with Buffer-1 [20 mM Tris–HCl, pH 8.0, 0.5 M NaCl, 50 mM imidazole, 10% glycerol, 1 mM 2-mercaptoethanol and 0.1% Nonidet P-40 (NP-40)], while beads bound with C1 WT, ΔRRM, ΔBASIC and ΔRRMΔBASIC were washed with Buffer-1 containing 150 mM imidazole. Proteins were eluted in Buffer-1 containing 500 mM imidazole. Eluted proteins were dialyzed against Buffer-2 [20 mM Tris–HCl, pH 8.0, 0.25 M NaCl, 10% glycerol, 1 mM dithiothreitol (DTT) and 0.2 mM EDTA] at 4°C overnight, while ΔACIDIC was dialyzed against Buffer-3 (20 mM Tris–HCl, pH 8.0, 0.5 M NaCl, 10% glycerol, 1 mM DTT and 0.2 mM EDTA). His-T7-Raly and chimeric proteins were purified using the same methods as for His-T7-C1 proteins. To purify GST–C1 WT and mutant proteins, the plasmids were transformed into *E. coli* BL21-Gold (DE3). The expressed protein in the bacterial lysate was bound to Glutathione Sepharose 4B (GE Healthcare) at 4°C for 1–3 h. Bound beads were washed five times with Buffer-3 containing 0.1% NP-40. Proteins were eluted in Buffer-3 containing 10 mM reduced glutathione and then dialyzed against Buffer-3 at 4°C overnight.

In vitro transcription

^{32}P-Labeled RNAs were transcribed in a 10 μl volume containing 20 U of T7 RNA polymerase (Promega), Transcription Buffer (Promega), 1 mM DTT (Promega), 12 U

of RNasin Plus (Promega), NTP mixture (0.5 mM ATP, 0.5 mM CTP, 0.1 mM UTP and 0.1 mM GTP), 1 μg of DNA template, 1 mM $m^7G(5')ppp(5')G$ RNA Cap Structure Analog (New England Biolabs) and 2.96 TBq/mmol [α-^{32}P]UTP (Perkin Elmer). Cap Structure Analog was not added for U6 snRNA or unlabeled 700 nt RNA. After a 60 min incubation at 37°C, RNA was recovered from the supernatants by phenol/chloroform extraction and purified using G-50 microcolumns (GE Healthcare). RNA was then precipitated with ethanol and dissolved in H_2O.

Splint ligation of RNA

Splint ligation of RNA was performed as previously described (40). *In vitro* transcribed, unlabeled and uncapped dihydrofolate reductase (DHFR; 700 nt) RNA was dephosphorylated with calf intestine alkaline phosphatase (TOYOBO) and then phosphorylated with T4 polynucleotide kinase (TOYOBO). The RNA fragment and *in vitro* transcribed, ^{32}P-labeled and m^7G-capped RNA (80 nt) were mixed with bridging DNA oligonucleotide and incubated with T4 DNA ligase (TOYOBO). The ligation product was separated by denaturing polyacrylamide gel electrophoresis (PAGE) and extracted from the gel.

RNA co-immunoprecipitation (co-IP)

RNA co-IP assay was performed as previously described (18). The ^{32}P-labeled RNAs were mixed with purified recombinant His-T7-C1 and incubated at 30°C for 20 min. The mixture was then rotated at 4°C for 1–3 h with Protein A–Sepharose beads (GE Healthcare) that were prebound to an anti-T7 tag antibody (Novagen). Beads were then washed five times with RSB100N buffer (10 mM Tris–HCl, pH 7.5, 100 mM NaCl, 2.5 mM $MgCl_2$ and 0.1% NP-40) and incubated in Homomix [50 mM Tris–HCl, pH 7.5, 5 mM EDTA, 1.5% sodium dodecylsulfate (SDS), 300 mM NaCl and 1.5 mg/ml proteinase K (Nacalai Tesque)] at 50°C for 30 min. RNA was recovered from the supernatant by phenol/chloroform extraction and ethanol precipitation, and then analyzed by denaturing PAGE and autoradiography.

GST–PHAX pull-down assay

GST–PHAX pull-down assay was performed as previously described (18). The ^{32}P-labeled RNAs were mixed with purified recombinant His-T7-C1, His-CBC and GST–PHAX, and the mixture was incubated at 30°C for 20 min. The reaction mixture was added to Glutathione Sepharose 4B equilibrated with RSB100N buffer and incubated on a rotating platform at 4°C for 1 h. Beads were then washed five times with RSB100N buffer and incubated in Homomix at 50°C for 30 min. RNA was recovered from the supernatant by phenol/chloroform extraction and ethanol precipitation. The samples were then analyzed by denaturing PAGE and autoradiography.

In vitro protein–protein interaction

Purified recombinant His-CBC and GST–C1 WT or mutants were incubated in the presence of RNase A. The mixture was rotated at 4°C for 1 h with Glutathione Sepharose 4B beads. After the beads were washed five times with RSB100N buffer, the bound material was recovered and analyzed by SDS–PAGE and western blotting (WB).

Gel filtration chromatography

Size exclusion chromatography of recombinant C1 proteins was performed with Superdex 200 (SMART System; GE Healthcare) equilibrated with Buffer-2 or Buffer-3.

Electrophoretic mobility shift assay (EMSA)

The ^{32}P-labeled RNAs were incubated with recombinant His-T7-C1 proteins in Buffer-2 at 30°C for 20 min. Samples were subsequently fractioned by native 5% PAGE followed by autoradiography. Data were analyzed using GraphPad Prism.

Ultraviolet (UV) cross-linking assay

The ^{32}P-labeled RNAs (3000 cpm/μl) were mixed with purified recombinant His-T7-C1 and/or His-CBC in a 20 μl volume containing 8 mM Tris–HCl (pH 8.0), 125 mM NaCl, 1.6 mM $MgCl_2$, 4% glycerol and RNasin. After an incubation at 30°C for 15 min, the mixture was irradiated by 254 nm UV light (FUNA-UV-LINKER FS-800; Funakoshi, Tokyo, Japan) at 100 mJ/cm^2 on ice. The irradiated sample was treated with RNase A (1 mg/ml; Nacalai Tesque) at 37°C for 15 min and cross-linked proteins were analyzed by SDS–PAGE and autoradiography.

Chemical cross-linking assay

Reaction mixtures containing 2 μg of purified recombinant His-T7-C1 in 30 μl of cross-linking buffer (20 mM HEPES-KOH, pH 7.9, 250 mM NaCl, 10% glycerol, 1 mM DTT and 0.2 mM EDTA) were incubated with glutaraldehyde to a final concentration of 0.005% at room temperature for 1, 2 and 5 min. Cross-linking was terminated by adding Tris–HCl (pH 8.0) to a final concentration of 150 mM. Reaction mixtures were mixed with SDS–PAGE sample buffer and heated at 85°C for 3 min. Samples were analyzed by 7.5% SDS–PAGE.

Protein–protein interactions in HEK293T cells

Huan embryonic kidney (HEK)293T cells were transfected with pCI-FLAG-C1 WT or ΔZIPPER plasmids. Cells were lysed in RSB100N buffer and lysates were rotated at 4°C for 1 h with anti-FLAG M2 affinity agarose gel (Sigma-Aldrich). After washing five times with RSB100N buffer, the bound material was recovered and analyzed by SDS–PAGE and WB.

Protein–mRNA interactions in HEK293T cells

In vivo protein–mRNA interaction assay was performed as previously described (41,42). HEK293T cells were transfected with pCI-FLAG-C1 WT or ΔZIPPER plasmids. Cells were irradiated by 254 nm UV light (FUNA-UV-LINKER FS-800) at 200 mJ/cm^2 on ice. Cells were lysed

in lysis/binding buffer (100 mM Tris–HCl, pH 7.5, 500 mM LiCl, 10 mM EDTA, 1% lithium-dodecylsulfate and 5 mM DTT) and homogenized using a narrow-gauge needle. The supernatants were added to oligo(dT)$_{25}$ Dynabeads (Thermo Fisher) and rotated at room temperature for 1 h. After washing three times with lysis/binding buffer and three times with NP-40 washing buffer (50 mM Tris–HCl, pH 7.5, 140 mM LiCl, 2 mM EDTA, 0.5% NP-40 and 0.5 mM DTT), protein–mRNA complexes were heat eluted from beads in 10 mM Tris–HCl (pH 7.5) for 2 min at 80°C. The samples were treated with RNase T1 and RNase A (Thermo Fisher). The released proteins were analyzed by SDS–PAGE and WB.

Xenopus oocyte microinjection

Xenopus oocyte microinjection was performed as previously described (1,16). Briefly, a ^{32}P-labeled RNA mixture was injected alone or with either His-T7-hnRNP C1 WT or ΔZIPPER into the nucleus. The nuclear fraction was prepared after 45 min, and IP was performed. The precipitated RNA was recovered and analyzed.

Antibodies

The antibodies were as follows: anti-CBP80 (43) and anti-PHAX (16) antibodies (generated in our laboratory), T7 for the T7 tag, M2 for the FLAG tag, 4F4 for hnRNP C, 9H10 for hnRNP A1 and 11G5 for ALY (Sigma-Aldrich), and anti-Lamin B1 and anti-TATA-binding protein (TBP) polyclonal antibodies (ProteinTech).

RESULTS

CBC stimulates the RNA binding of hnRNP C

We previously showed that hnRNP C selectively binds unstructured RNA regions longer than 200–300 nt and that it directly interacts with CBC (18). On the basis of these findings, we hypothesized that hnRNP C interrupts CBC–PHAX interaction on m^{7}G-capped long RNA (Figure 1A). However, whether CBC–hnRNP C interaction occurs on m^{7}G-capped long RNA is unclear. Therefore, we first performed an *in vitro* RNA–protein binding assay. A mixture of ^{32}P-labeled *in vitro* transcribed RNAs containing intronless DHFR mRNA (700 nt), three shortened DHFR RNAs (300–120 nt) and U1 and U6 snRNAs was incubated with purified recombinant CBC and T7-tagged hnRNP C proteins (Supplementary Figure S1A, B). All RNAs were m^{7}G capped, except for U6 snRNA. After the incubation, the hnRNP C protein was precipitated with the anti-T7 tag antibody, and co-precipitated RNA was analyzed by denaturing PAGE (RNA co-IP). When hnRNP C alone was used, DHFR mRNA (700 nt) was co-precipitated, whereas shorter RNAs were not (Figure 1C, lane 2), supporting the previous finding showing that hnRNP C *per se* has an intrinsic property of selective binding to long RNAs (17). When an increasing amount of recombinant CBC was added, hnRNP C binding to DHFR mRNA (700 nt) was increased and that to DHFR (300 nt) and DHFR (200 nt) became detectable (Figure 1C, lanes 3 and 4; Supplementary Figure S2), suggesting that hnRNP C directly interacted with CBC on RNA. Notably, CBC did not affect the selectivity of hnRNP C binding to longer RNAs.

To analyze whether CBC and hnRNP C interact on the cap-proximal region, we performed a UV cross-linking assay (Figure 1D). First, m^{7}G-capped RNA, ^{32}P-labeled only near the cap-proximal region, was generated by the splint ligation method (40). The RNA was incubated with purified recombinant hnRNP C and then irradiated by UV light. If hnRNP C binds to the cap-proximal region of labeled RNA, it is cross-linked to the RNA and labeled by ^{32}P. After incubation with RNase A to digest the RNA moiety, the sample was separated by SDS–PAGE and analyzed by autoradiography. A band corresponding to 43 kDa was increased by the addition of CBC (Figure 1D), indicating that hnRNP C interacted with CBC on the cap-proximal region.

To identify the regions responsible for the interaction of hnRNP C with CBC, we constructed a series of mutants in which the C-terminal, N-terminal and both the RRM and ACIDIC regions of C1 were deleted, named N-half, C-half and Middle, respectively (Figure 1B). Mutants and WT C1 with an N-terminal GST tag were expressed in *E. coli* and purified. A pull-down assay using purified recombinant CBC and GST-tagged proteins was performed in the presence of RNase A. CBP80, a component of CBC, was pulled down with WT, N-half and Middle, but not with GST or C-half (Figure 1E, lanes 2–6). To narrow down the interaction regions, we constructed the deletion mutants RRM, AUX and BASIC regions (Figure 1B). CBP80 was slightly pulled down by RRM, but not by AUX or BASIC (Figure 1E, lanes 7–9). The specific region in Middle for the interaction with CBC could not be determined, possibly because fragmentation changed the tertiary structure, including the interaction sites. These results suggest that multiple interaction sites in the N-terminal region of hnRNP C contribute to the direct interaction with CBC.

Mutations in the ZIPPER domain, but not the RNA-binding domains, disrupt tetramer formation of hnRNP C

To systematically analyze the regions of hnRNP C involved in PHAX-inhibitory activity, we constructed a series of mutants deleted for the RRM, BASIC, ZIPPER and ACIDIC regions of hnRNP C1, named ΔRRM, ΔBASIC, ΔZIPPER and ΔACIDIC, respectively (Figure 2A). We also constructed ΔRRMΔBASIC, a mutant deleted for both RRM and BASIC, to evaluate the function of RNA-binding domains in the inhibition of PHAX (Figure 2A). Additionally, the site-directed mutant ZIP-MT4, which has four amino acid substitutions in ZIPPER, was designed to analyze the importance of oligomer formation of hnRNP C in PHAX-inhibitory activity (Figure 2A, B) (2). While hnRNP C exists as a heterotetramer of C1 and C2 in cells, C2 was not used in the present study because we previously demonstrated that the C1 homotetramer also inhibited PHAX binding to longer RNAs as a heterotetramer (18). The five mutants and WT C1 with N-terminal His and T7 tags were expressed in *E. coli* and purified (Supplementary Figure S1A).

To analyze the oligomer formation of the purified recombinant proteins, gel filtration chromatography was performed (Figure 2C; uncropped images are shown in Sup-

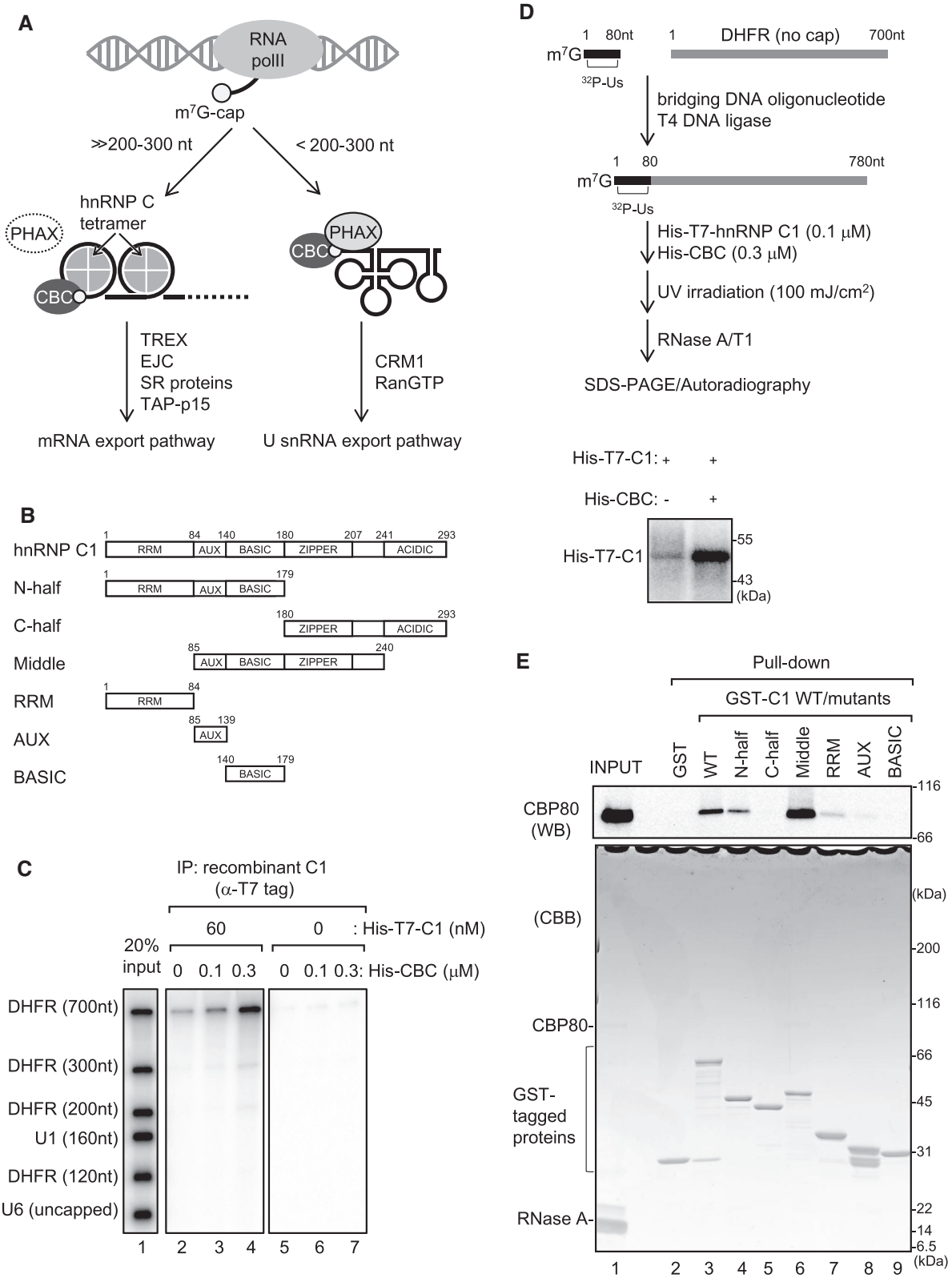


Figure 1. CBC stimulates RNA binding of hnRNP C. (**A**) A model of the classification of RNA polymerase II transcripts by hnRNP C. Refer to the Introduction. (**B**) Schematic representation of the deletion constructs. (**C**) The ^{32}P-labeled RNAs, recombinant His-T7-C1 and His-CBC were incubated, and IP assays were performed using an anti-T7 monoclonal antibody. Precipitated RNA was analyzed. (**D**) m^{7}G-capped RNA, ^{32}P-labeled only near the cap-proximal region, was generated through the splint ligation method. The RNA was incubated with purified recombinant hnRNP C and then irradiated by UV light. The cross-linked hnRNP C was separated by SDS–PAGE and analyzed by autoradiography. (**E**) Top: recombinant His-CBP80 and GST-tagged proteins were incubated in the presence of RNase A and pull-down assays were performed. Proteins were detected by WB using an anti-CBP80 antibody (top) or Coomassie Brilliant Blue staining (bottom).

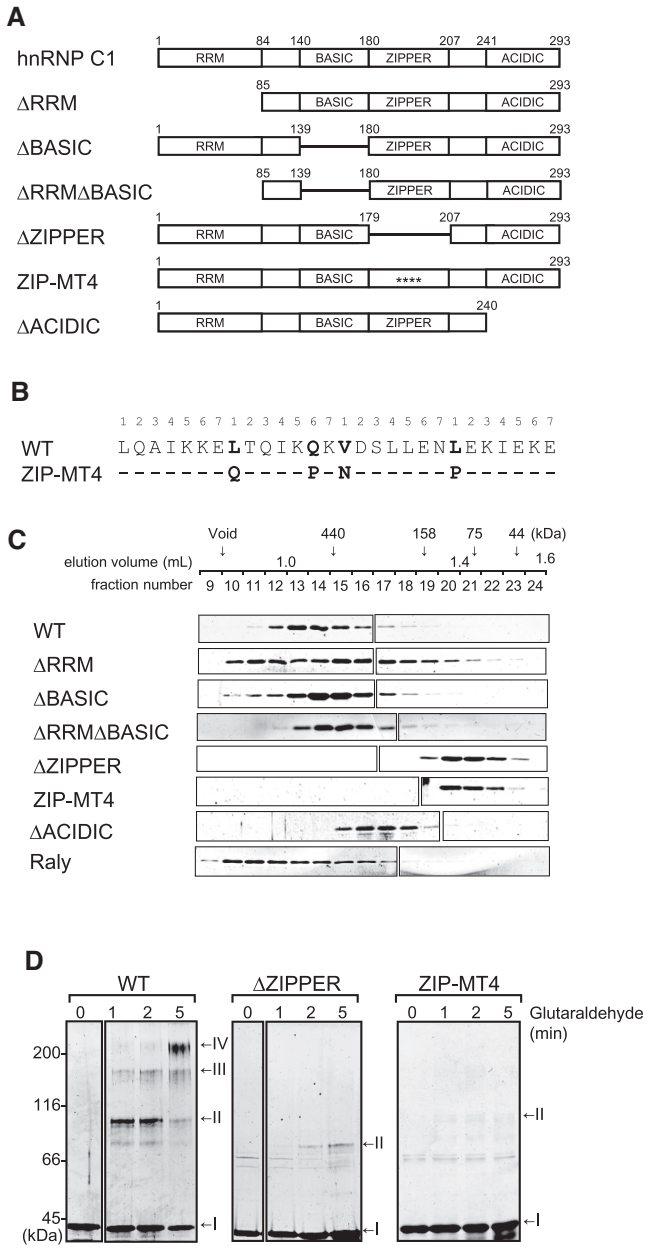


Figure 2. Effects of the ZIPPER domain of hnRNP C on oligomer formation. (**A**) Schematic representation of the deletion and site-directed constructs. '*' indicates an amino acid substitution. (**B**) The ZIPPER sequence (amino acids 180–208) and substituted amino acids in the site-directed mutant ZIP-MT4 are shown. Numbers 1–7 above the sequence denote the position of each amino acid in the heptad repeat. (**C**) Purified recombinant WT and mutant hnRNP C1 proteins were subjected to gel filtration chromatography. The molecular weight of each marker protein is denoted above the ruler. Uncropped images are shown in Supplementary Figure S3. (**D**) Coomassie-stained SDS–PAGE of chemically cross-linked recombinant hnRNP C1 WT and ZIPPER mutant proteins. Proteins were treated with glutaraldehyde for 0, 1, 2 and 5 min. 'I' indicates the unshifted band (monomers), while 'II' 'III' and 'IV' indicate the expected sizes of the dimers, tetramers and larger oligomers of the recombinant proteins, respectively.

plementary Figure S3). WT C1 was eluted with an elution volume corresponding to an apparent molecular mass of 400–440 kDa, which is consistent with previous studies that analyzed hnRNP C tetramers (18,37,44). Deletion mutants of RNA-binding domains (ΔRRM, ΔBASIC and ΔRRMΔBASIC) were eluted with similar volumes to WT C1, indicating that these mutants formed tetramers as did WT C1. ZIPPER domain mutants (ΔZIPPER and ZIP-MT4) were eluted in fractions for apparent molecular masses of 100–120 kDa, strongly indicating deficient oligomerization of these two ZIPPER mutants, which was also consistent with previous studies (35,39). Additionally, the ACIDIC region mutant (ΔACIDIC) was eluted in fractions with smaller sizes than the WT and the deletion mutants of RNA-binding domains (ΔRRM, ΔBASIC and ΔRRMΔBASIC), but it was still larger than the ZIPPER mutants (ΔZIPPER and ZIP-MT4), indicating that ΔACIDIC was eluted as dimers, as previously reported (35). ΔACIDIC was not used in subsequent RNA-protein binding assays, because it was unstable in the assays.

Oligomer formation was also analyzed using a chemical cross-linking assay (Figure 2D). When recombinant WT C1 was incubated in the presence of the effective protein cross-linking reagent glutaraldehyde, proteins were detected at the sizes of the dimers and tetramers on SDS–PAGE. In the case of ZIPPER mutants, only weak dimer signals and no tetramer signals were observed. These results further confirmed that the ZIPPER domain mutants were deficient in oligomerization.

The BASIC domain of hnRNP C is required for PHAX-inhibitory activity

To evaluate the inhibitory effects of hnRNP C on PHAX binding to longer RNAs, we utilized an *in vitro* system that we previously developed for the identification of RNA classification factors (18). This system recapitulates the remodeling of RNA–protein export complex formation in accordance with RNA length. The same RNA mixture as that in Figure 1C was incubated with purified recombinant CBC and GST-tagged PHAX proteins (Supplementary Figure S1B, C), and a GST pull-down assay was performed to examine U snRNA export complex formation. WT C1 and mutants were added to the GST–PHAX pull-down assay, and the RNA binding of PHAX was examined (Figure 3). An RNA co-IP assay using the anti-T7 tag antibody, similar to that in Figure 1C, was also performed to observe the RNA binding of hnRNP C1 WT and mutants (Figure 3).

We confirmed that all m^7G-capped RNAs were pulled down in the absence of C1 (Figure 3A, lane 2) (18). When increasing amounts of recombinant C1 were added to the system, PHAX binding to longer RNAs was inhibited, while binding to shorter RNAs was not (Figure 3A, lanes 3–5, and B) (18). We also confirmed that recombinant C1 preferentially bound to longer RNAs, suggesting that PHAX-inhibitory activity by hnRNP C depended on its RNA-binding activity (Figure 3F, lanes 3–5, and G) (18).

We then examined the contribution of RNA-binding domains to PHAX-inhibitory activity. The ΔRRM mutant inhibited PHAX binding to longer RNAs almost as efficiently as WT C1 (Figure 3A, lanes 6–8, and C). The strength of its

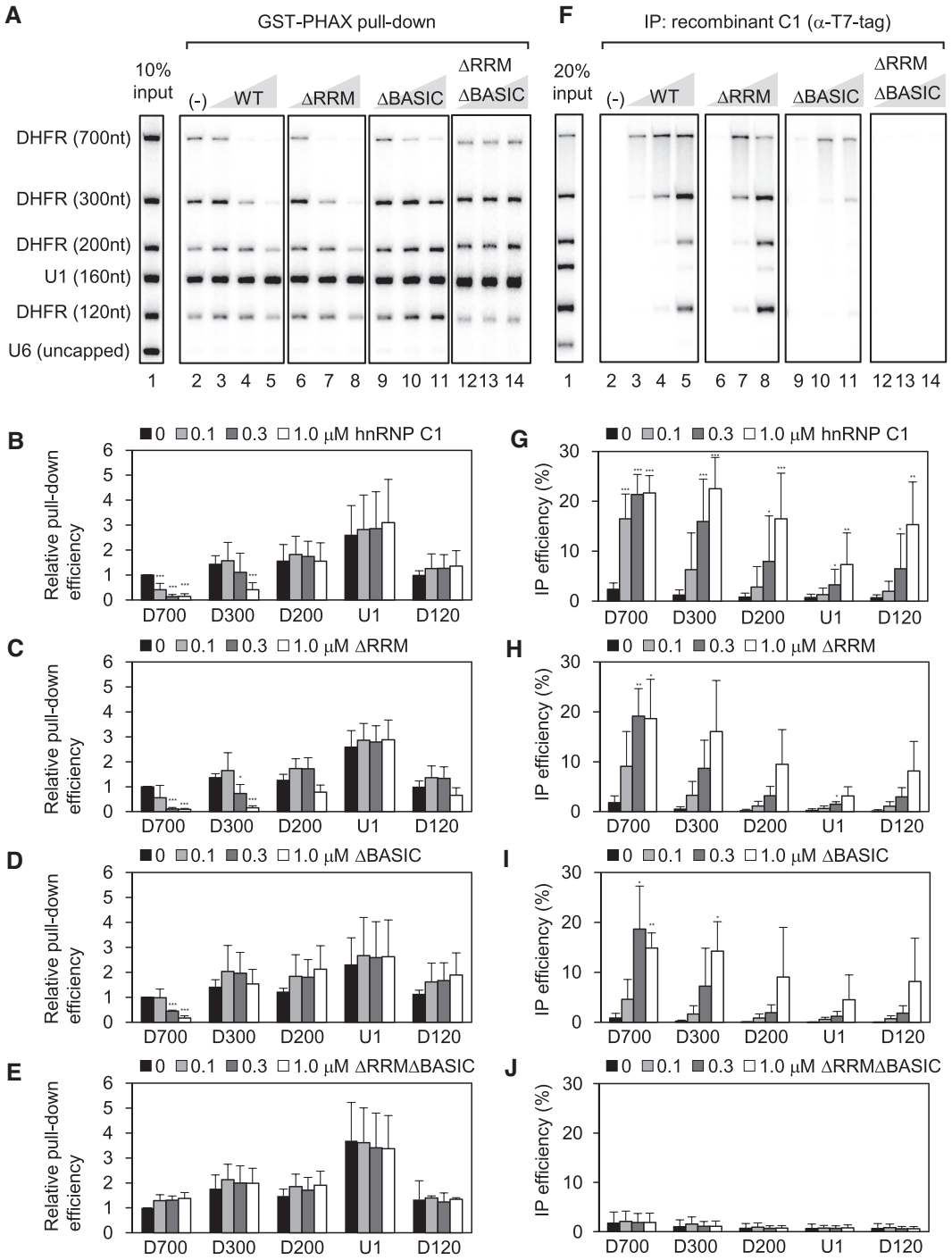


Figure 3. Effects of the RNA-binding domain of hnRNP C on PHAX-inhibitory activity. (**A**) ^{32}P-labeled RNAs were incubated with purified recombinant His-T7-C1 (0.1, 0.3 and 1 μM as the monomer unit), GST–PHAX (0.2 μM) and His-CBC (0.2 μM). GST pull-down assay was performed using glutathione beads and pulled-down RNA was analyzed. (**B**) Quantitation of relative pull-down efficiency from nine independent experiments performed as in (A). (**C**–**E**) Quantitation of relative pull-down efficiency from three independent experiments performed as in (A). The efficiency of the DHFR (700 nt) RNA of the buffer control sample (lane 2) was set to 1. Averages and standard deviations (SDs) are denoted. (**F**) The same ^{32}P-labeled RNAs, recombinant His-T7-C1 and His-CBC were incubated, and IP assay was performed using an anti-T7 monoclonal antibody. Precipitated RNA was analyzed. (**G**) Quantitation of IP efficiency from nine independent experiments performed as in (F). (**H**–**J**) Quantitation of IP efficiency from three independent experiments performed as in (F). Averages and SDs are denoted. Statistically significant differences compared with each buffer control are shown. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

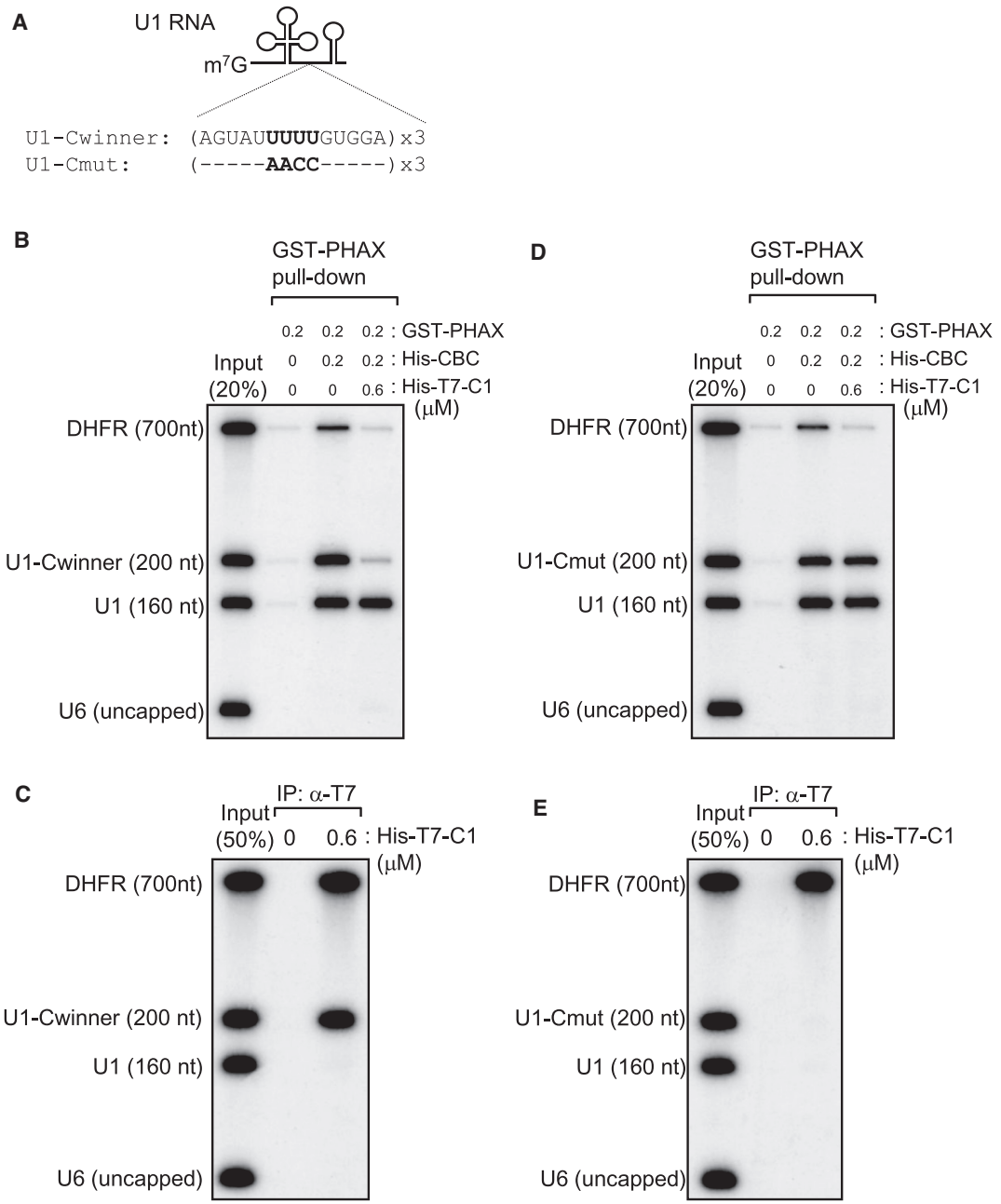


Figure 4. The uridine stretch contributes to the PHAX-inhibitory activity of hnRNP C. (**A**) Schematic representation of U1-Cwinner or U1-Cmut, in which the high-affinity or control sequence of hnRNP C was inserted into U1 RNA. (**B** and **D**) GST–PHAX pull-down assay was performed using ^{32}P-labeled RNAs containing DHFR (700 nt), U1, U6 and U1-Cwinner or U1-Cmut as described in Figure 3. (**C** and **E**) IP was performed as described in Figure 3.

binding to RNAs was equivalent to that of the WT (Figure 3F, lanes 6–8, and H), which was consistent with previous findings on hnRNP C mutants lacking an RRM (35). In contrast, ΔBASIC did not inhibit PHAX binding to DHFR (300 nt) and weakly inhibited binding to DHFR (700 nt) (Figure 3A, lanes 9–11, and D). Additionally, the RNA binding of ΔBASIC was markedly weaker than that of the WT or ΔRRM (Figure 3F, lanes 9–11, and I). The mutant deleted for both RRM and BASIC (ΔRRMΔBASIC) did not exhibit PHAX-inhibitory or RNA-binding activity (Figure 3A, lanes 12–14, E, F, lanes 12–14, and J). RRM or BASIC alone bound to RNAs, and the RNA binding of BASIC was stronger than that of the RRM (Supplementary Figure S4A, B). These results suggested that the BASIC domain was more critical than the RRM for hnRNP C binding to longer RNAs and PHAX inhibition from RNAs. Neither the RRM nor the BASIC domain alone was sufficient for PHAX-inhibitory activity (Supplementary Figure S4C).

To analyze the contribution of RRM to PHAX-inhibitory activity, we constructed two RNAs: U1-Cwinner and U1-Cmut RNAs (Figure 4A). U1-Cwinner was generated by the insertion of three copies of a high-affinity SELEX winner sequence for hnRNP C into U1 RNA (33). To generate the control U1-Cmut RNA, the uridine stretch

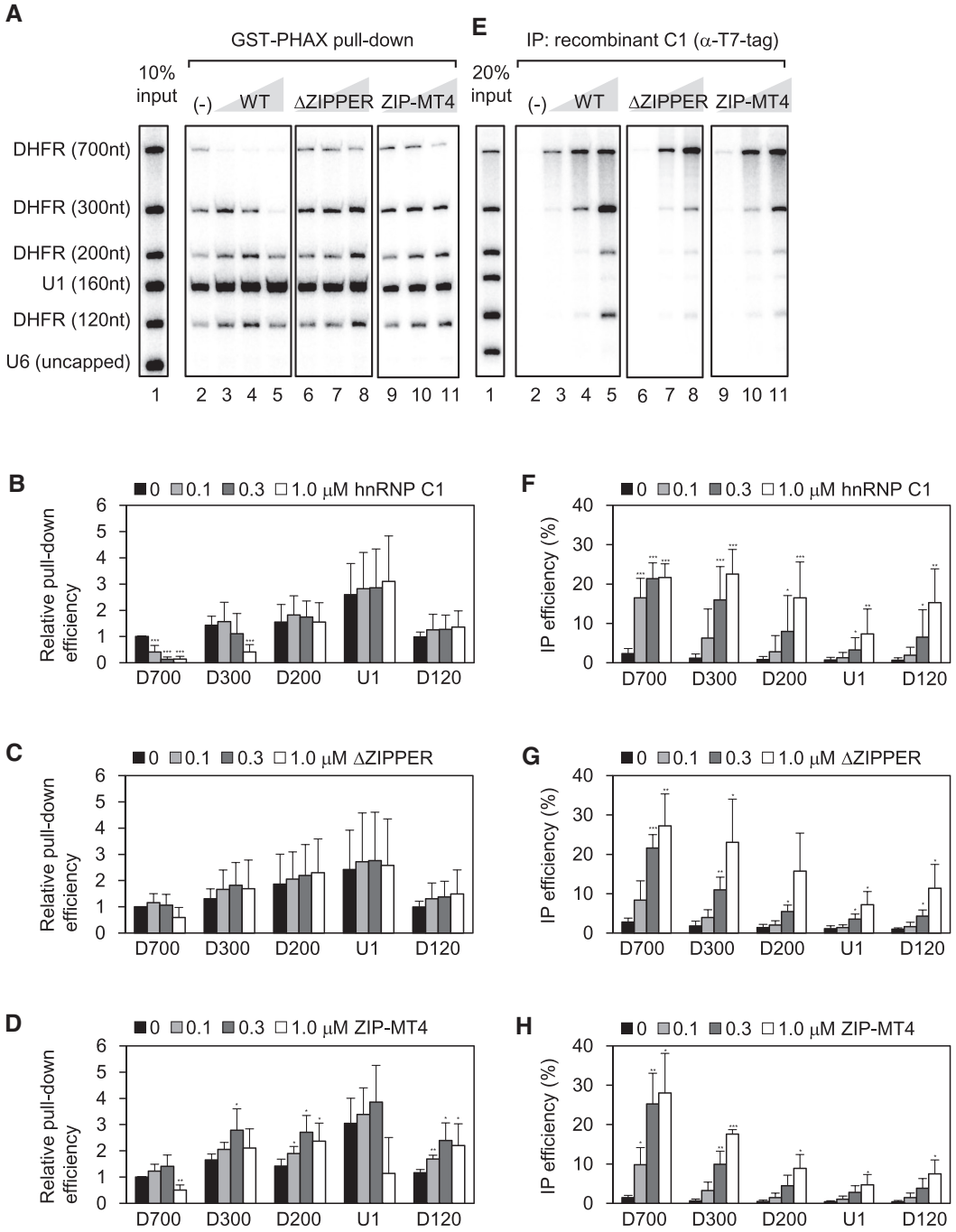


Figure 5. Effects of the ZIPPER domain of hnRNP C on PHAX-inhibitory activity. (**A**) GST–PHAX pull-down assay was performed as described in Figure 3. (**B**) and (**F**) are identical to Figure 3B and G, respectively. (**C** and **D**) Quantitation of relative pull-down efficiency from three independent experiments performed as in (A). (**E**) IP was performed as described in Figure 1. (**G** and **H**) Quantitation of IP efficiency from three independent experiments performed as in (E). Statistically significant differences compared with each buffer control are shown. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

in U1-Cwinner, which is the preferred binding sequence for hnRNP C (30–32), was replaced with adenine and cytidine. HnRNP C was expected to bind to U1-Cwinner through its RRM owing to the uridine stretch. As expected, even though U1-Cwinner was short, hnRNP C efficiently bound to the RNA, and PHAX was excluded (Figure 4B, C). In contrast, hnRNP C did not bind to U1-Cmut, and therefore PHAX was not excluded (Figure 4D, E). These results suggested that the RRM also contributes to PHAX-inhibitory activity when RNA contains a uridine stretch.

The ZIPPER domain is required for PHAX-inhibitory activity

We investigated the contribution of the oligomerization domain (ZIPPER) to the PHAX-inhibitory activity of hnRNP C. The ΔZIPPER and ZIP-MT4 mutants showed

Table 1. A summary of PHAX-inhibitory and RNA-binding activities of hnRNP C1 WT and mutants

	PHAX-inhibitory activity	RNA-binding activity
WT	+++	+++
ΔRRM	++	++
ΔBASIC	+/–	+
ΔRRMΔBASIC	–	–
ΔZIPPER	–	++
ZIP-MT4	–	++

Quantified PHAX-inhibitory and RNA-binding activities of hnRNP C1 WT and mutants at 0.1 μM on DHFR mRNA (700 nt) in the GST–PHAX pull-down assay (Figures 3B–E and 5B–D) and IP assay (Figures 3G–J and 5F–H) are summarized. PHAX-inhibitory activity was calculated from pull-down efficiency and normalized to the buffer control. RNA-binding activity was calculated from IP efficiency. '+++', '++' and '+' indicate strong, middle and weak activities, respectively. '–' indicates almost no activity.

little inhibition of PHAX binding to longer RNAs (Figure 5A, lanes 6–11, C, D). Their RNA-binding activities were weaker than that of the WT (Figure 5E, lanes 6–11, G, H). These results suggested that oligomer formation strengthened the RNA-binding activity of hnRNP C, thereby supporting PHAX-inhibitory activity. Notably, PHAX binding to RNA was not affected even when these mutants efficiently bound to RNA (compare lanes 3, 7 and 10 in Figure 5A and E), suggesting that not only RNA binding but also tetramer formation of hnRNP C was required for effective PHAX-inhibitory activity (Table 1).

The ZIPPER domain strengthens the RNA-binding activity of hnRNP C

RNA co-IP assays showed that the RNA binding of ZIPPER domain mutants was weaker than that of WT hnRNP C (Figure 5E). To quantitatively assess the contribution of the ZIPPER domain to RNA-binding activity, we performed an EMSA using *in vitro* transcribed RNAs (DHFR 300 nt and 120 nt) and purified recombinant C1 proteins. As shown in Figure 6, the RNA binding of ZIP-MT4 was weaker than that of the WT. The WT preferentially bound to longer RNA, as previously reported (18), whereas this binding preference was not observed for ZIP-MT4. These results strongly suggested that oligomer formation via the ZIPPER domain strengthens RNA binding of hnRNP C and is important for the classification of transcripts in accordance with their lengths.

The BASIC-ZIPPER domain of hnRNP C confers PHAX-inhibitory activity on Raly, the closest relative of hnRNP C

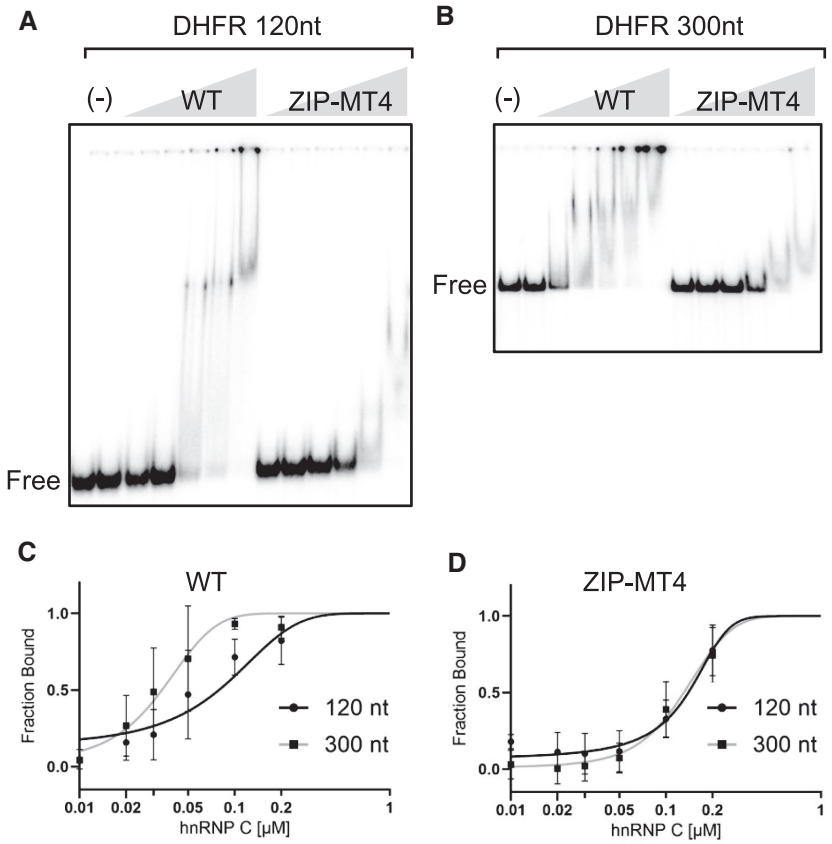
Raly, a member of the hnRNP family, shares high amino acid sequence similarity with hnRNP C, except for the C-terminal region (Figure 7A; Supplementary Figure S5). This prompted us to investigate whether Raly also exhibits PHAX-inhibitory activity. In GST–PHAX pull-down assays, Raly did not inhibit PHAX binding to RNAs (Figure 7B–D). RNA co-IP results showed that the RNA-binding activity of Raly was markedly weaker than that of C1 (Figure 7G–I). To further confirm the importance of the BASIC and ZIPPER domains of hnRNP C on PHAX-inhibitory activity, we used a chimeric mutant of Raly in which the BASIC-ZIPPER domains were substituted with those of hnRNP C1 (R-C-R) (Figure 7A). GST–PHAX pull-down assay showed that the PHAX-inhibitory activity of R-C-R increased, but it was still less than that of C1 (Figure 7B, E). In contrast, when we used another chimeric mutant in which the BASIC-ZIPPER domains of C1 were substituted with those of Raly (C-R-C), C-R-C almost completely lost PHAX-inhibitory activity (Figure 7B, F). These results strongly support the importance of the BASIC and ZIPPER domains of hnRNP C in the inhibition of PHAX binding to longer RNAs.

Role of the ZIPPER domain of hnRNP C *in vivo*

The ZIPPER domain was required for tetramer formation of hnRNP C *in vitro* (Figure 2) (35,39), and our *in vitro* assays suggested that tetramer formation of hnRNP C was important for RNA-binding activity (Figures 5 and 6). To investigate whether this was also the case *in vivo*, we performed IP assays using HEK293T cell lysates. In HEK293T cells transiently transfected with plasmid expressing FLAG-tagged WT hnRNP C1 (FLAG-WT) or the ZIPPER deletion mutant (FLAG-ΔZIPPER), FLAG-ΔZIPPER has much lower expression than FLAG-WT. Endogenous hnRNP C was co-precipitated with FLAG-WT but not FLAG-ΔZIPPER in the presence of RNase A. This result suggested that FLAG-WT formed a tetramer with endogenous hnRNP C and that the ZIPPER domain was required for tetramer formation *in vivo* (Figure 8A, lanes 8, 9, 11 and 12). Two RNA-binding proteins, hnRNP A1 and ALY, were co-precipitated with FLAG-WT in the absence of RNase A (Figure 8A, lanes 2 and 5). These interactions were dependent on RNA because co-IP was not observed in the presence of RNase A (Figure 8A, lanes 8 and 11). The interaction between hnRNP C and CBP80 was also dependent on RNA (Figure 8A, lanes 2, 5, 8 and 11). In contrast, none of the RNA-binding proteins was co-precipitated with FLAG-ΔZIPPER (Figure 8A, lanes 3, 6, 9 and 12).

To directly examine *in vivo* RNA–protein interactions, we performed a UV cross-linking assay. HEK293T cells expressing FLAG-WT or FLAG-ΔZIPPER were irradiated by UV light to cross-link RNA with RNA-binding proteins in cells. Endogenous hnRNP C and FLAG-WT, but not FLAG-ΔZIPPER, were pulled down with oligo(dT) beads, indicating that endogenous hnRNP C and FLAG-WT, but not FLAG-ΔZIPPER, directly bound to mRNA in cells (Figure 8B). This suggested that tetramer formation of hnRNP C was required for RNA-binding activity *in vivo*.

The remodeling of RNA–protein export complex formation in accordance with RNA length was originally identified in experiments using a *Xenopus* oocyte microinjection system (2,16). To investigate whether tetramer formation of hnRNP C is required for PHAX-inhibitory activity *in vivo*, we analyzed RNA binding of PHAX using the *Xenopus* oocyte microinjection system (Figure 9A). A mixture of ^{32}P-labeled RNAs was injected into the nucleus either alone or together with purified recombinant hnRNP C protein. After an incubation period of 45 min, the oocyte was dissected into the nucleus and cytoplasm, and RNA



E

	K_D (nM)	
	120 nt	300 nt
WT	79.2	34.0
ZIP-MT4	139	130

Figure 6. RNA-binding activity of hnRNP C1 WT and ZIP-MT4. (**A** and **B**) An EMSA was performed using purified recombinant His-T7-C1 (A) and His-T7-ZIP-MT4 (B) proteins (10, 20, 30, 50, 100 and 200 nM) with ^{32}P-labeled DHFR mRNAs (120 or 300 nt). The mixture was subjected to 5% native PAGE. (**C** and **D**) Quantification of EMSA from three independent experiments as in (A) and (B), respectively. (**E**) The equilibrium dissociation constant (K_D) was calculated from the quantification in (C) and (D).

co-IP assay was performed using the nuclear fraction. Injection of hnRNP C1 WT, but not ΔZIPPER, inhibited RNA binding of PHAX (Figure 9B, C), suggesting the importance of tetramer formation of hnRNP C in PHAX-inhibitory activity *in vivo*. WT hnRNP C1 efficiently bound to longer RNAs, whereas ΔZIPPER bound less effectively to DHFR (700 nt) and hardly bound to other RNAs (Figure 9D, E). These findings were consistent with the observations in experiments with purified proteins and HEK293T cells.

DISCUSSION

We previously demonstrated that hnRNP C classified nascent RNA polymerase II transcripts into U snRNAs and mRNAs by inhibiting PHAX binding to long transcripts (18). In the present study, we performed *in vitro* RNA–protein binding assays using hnRNP C mutants to identify the regions in hnRNP C essential for its RNA classification function. We found that PHAX-inhibitory activity by hnRNP C largely depended on its RNA-binding activity (Figure 3). The BASIC region plays a more important role in the inhibition of PHAX than the canonical RRM because of its strong RNA-binding activity. While the ZIPPER domain *per se* did not exhibit RNA-binding activity, it contributed to the inhibition of PHAX by facilitating oligomer formation (Figures 2 and 5). The ZIPPER domain markedly contributed to the RNA-binding activity of hnRNP C *in vitro* and *in vivo* (Figures 5, 6, 8 and 9). In addition to RNA binding, oligomer formation was also necessary for PHAX-inhibitory activity (Figure 5). RRM contributed to the inhibition of PHAX when RNA contained a uridine stretch (Figure 4). Additionally, hnRNP C directly interacted with CBC on RNA (Figure 1). A long stretch of N-terminal regions may provide a structured region for CBC to interact with hnRNP C (Figure 1). The

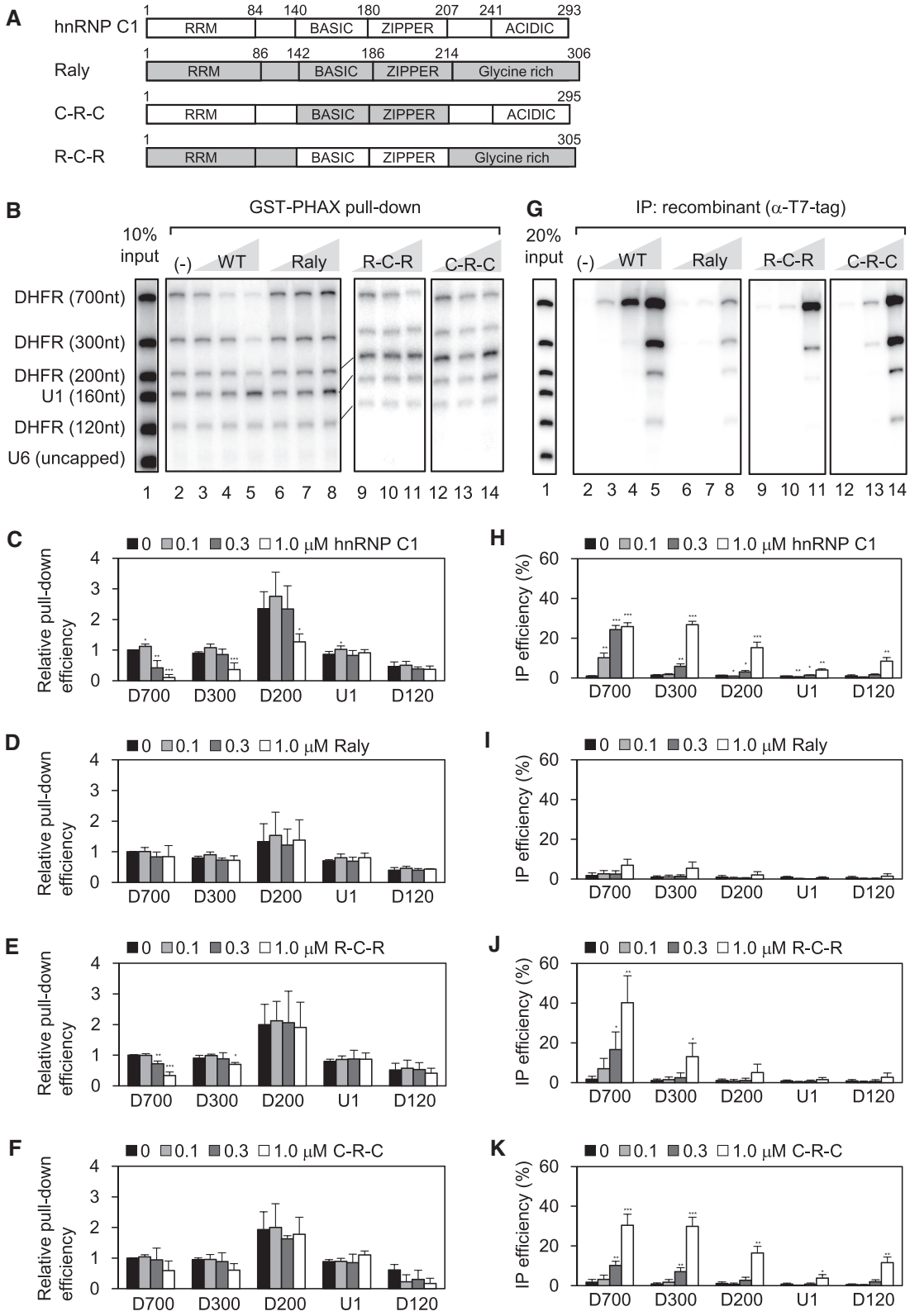


Figure 7. Raly does not exhibit PHAX inhibitory activity. (**A**) Schematic representation of hnRNP C1, Raly and their chimeric mutants. (**B**) and (**G**) GST–PHAX pull-down and RNA co-IP assays were performed as in Figure 3. (**C**–**F**) and (**H**–**K**) Quantitation of three independent experiments as in (**B**) and (G). Averages and standard deviations are denoted. Statistically significant differences compared with each buffer control are shown. **P* <0.05, ***P* <0.01, ****P* <0.001.

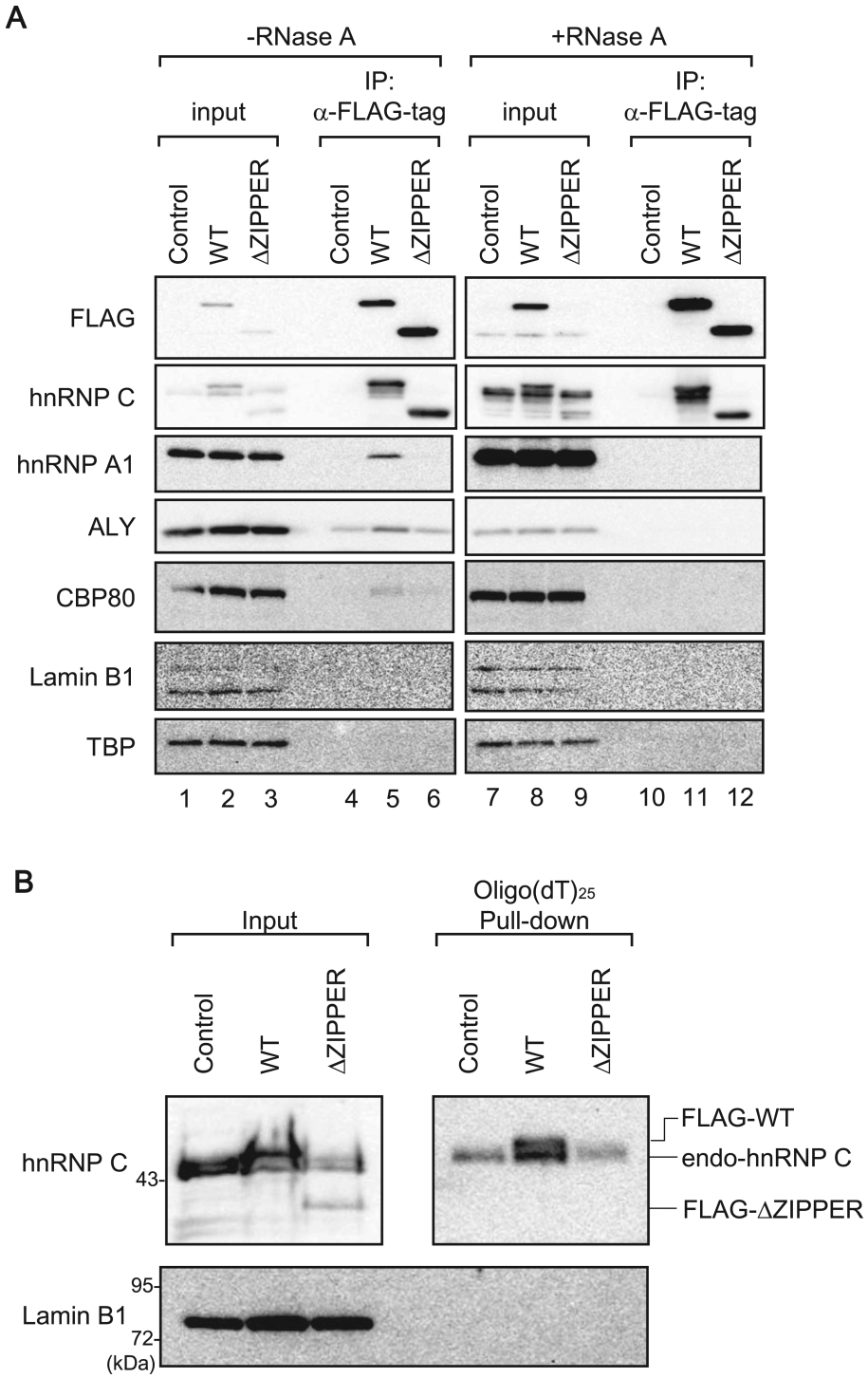


Figure 8. Effects of the ZIPPER domain on RNA-binding activity of hnRNP C in HEK293T cells. (**A**) HEK293T cell lysates were subjected to IP assay using the anti-FLAG antibody M2 in the absence or presence of RNase A. FLAG-tagged proteins, endogenous hnRNP C, hnRNP A1, ALY, CBP80, Lamin B1 and TATA-binding protein (TBP) were detected by WB. (B) HEK293T cells were irradiated by UV light and cell lysates were subjected to pull-down assay using oligo(dT)$_{25}$ beads. HnRNP C and Lamin B1 were detected by WB.

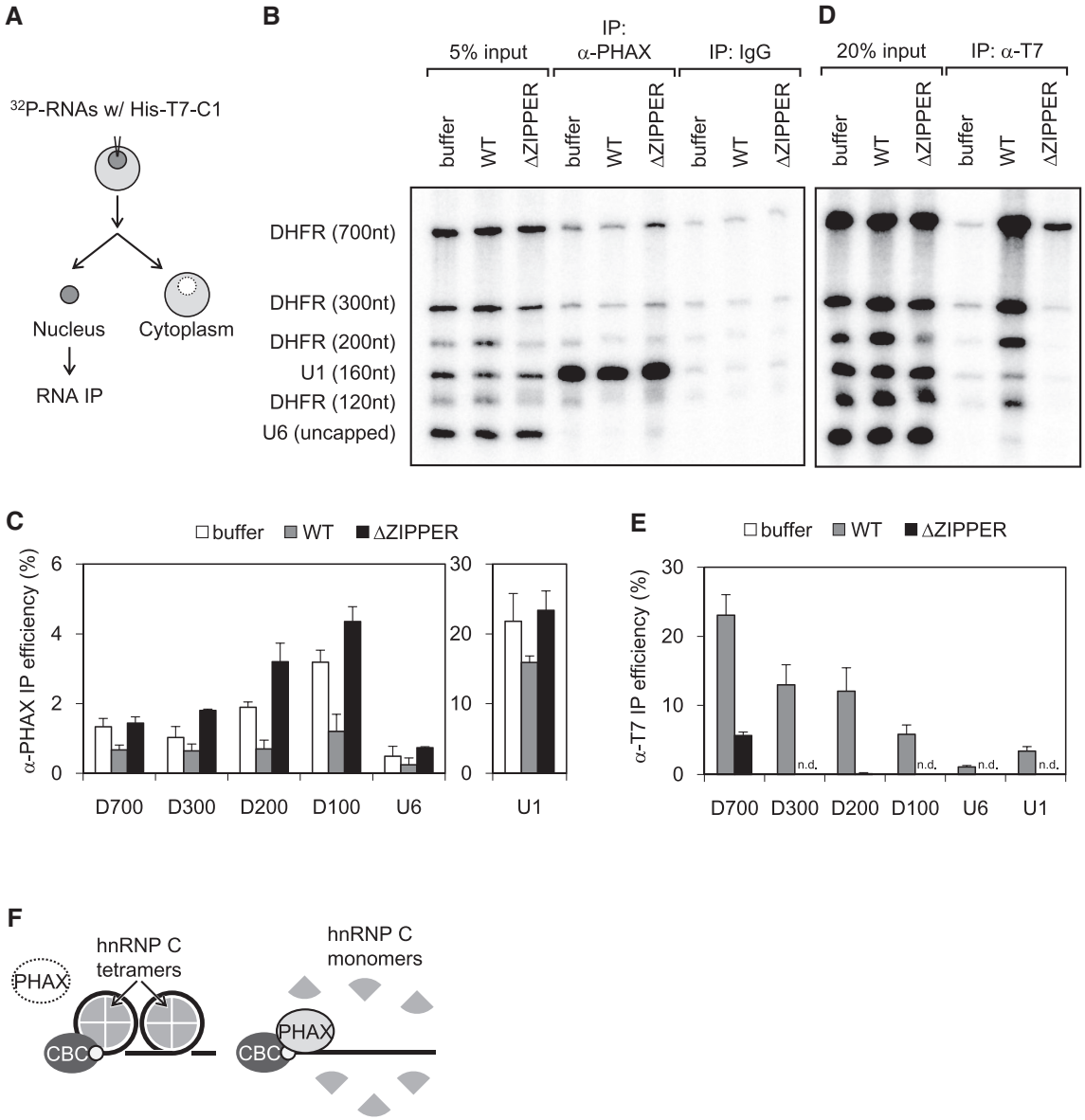


Figure 9. Effects of the ZIPPER domain of hnRNP C on PHAX-inhibitory activity in *Xenopus* oocytes. (**A**) A ^{32}P-labeled RNA mixture was injected alone or with either purified recombinant T7-hnRNP C1 WT or ΔZIPPER (300 fmol/oocyte) into the nucleus of *Xenopus* oocytes. After an incubation period of 45 min, the oocyte was dissected into the nucleus and the cytoplasm. The nuclear fraction was subjected to RNA co-IP. (**B**) IP was performed with either anti-PHAX or anti-mouse IgG antibodies. (**C**) Quantitation of IP efficiency from three independent experiments performed as in (B). (**D**) IP was performed with anti-T7 antibody. (**E**) Quantitation of IP efficiency from three independent experiments performed as in (D). Averages and standard deviations are denoted. n.d: not detected. (**F**) Refer to the Discussion for details.

interaction of hnRNP C with CBC may be relevant to the inhibition of PHAX, which binds to m^{7}G-capped RNA in a CBC-dependent manner.

Model for the inhibitory effects of hnRNP C on the recruitment of PHAX to mRNA

A model of the mechanism by which hnRNP C inhibits the recruitment of PHAX to mRNA is shown in Figure 9F. Tetramer formation of hnRNP C via the ZIPPER domain and ACIDIC region presents four sequence-independent RNA-binding BASIC regions. When RNA is sufficiently long, RNA tightly winds around the tetramer, and all four BASIC regions bind to RNA, thus stabilizing the RNP. If the RNA contains a uridine stretch, RNP formation is further strengthened, with the interaction between the uridine stretch and the RRM. In the cap-proximal region, RNA is wrapped around the hnRNP C tetramer without a gap from the cap through a direct interaction with CBC via N-terminal regions, thereby inhibiting the access of PHAX. In contrast, if the RNA is short or hnRNP C is deficient in tetramer formation, the tetramer or monomer cannot stably bind to RNA, thereby allowing PHAX recruitment. Monomers bound to RNA may provide PHAX with the necessary space for recruitment.

While the ZIPPER and ACIDIC regions of hnRNP C, both of which are required for its tetramer formation, do not exhibit RNA-binding activity, tetramer formation

strengthens the RNA-binding activity of hnRNP C. RNA binding of hnRNP C was previously shown to be weakened by the phosphorylation of serine residues mainly within the ACIDIC region (45–47). These findings suggested that phosphorylation of the ACIDIC region induced the disassembly of the tetramer, thereby decreasing RNA binding. A previous study indicated that the hnRNP complex, in which hnRNP C is a core component, was assembled and disassembled during cell cycle progression as a result of mitosis-specific phosphorylation of hnRNP C (48). The phosphorylation-induced disassembly of the hnRNP C tetramer may explain hnRNP disassembly from mRNA.

RNP remodeling by hnRNP C

The strong RNA-binding activity of the BASIC region and the tetramer formation activity of the ZIPPER domain are important for the PHAX-inhibitory activity of hnRNP C. The mechanism underlying the inhibition of PHAX via the BASIC and ZIPPER domains may also play a role in RNP remodeling by hnRNP C. For example, hnRNP C inhibits the splicing reaction by competing with the splicing factor U2AF, thereby protecting the transcriptome from abnormal exonization, such as uridine-rich Alu elements (26). The RRM of hnRNP C should play an important role in this competition because it exhibits a sequence preference for the uridine stretch. In addition to the RRM, the BASIC and ZIPPER domains wrap RNA around the hnRNP C tetramer, allowing hnRNP C to tightly occupy a long RNA region centering on the uridine stretch. This binding mode may inhibit the recruitment of not only U2AF but also other splicing factors, which effectively suppresses exonization. Similar RNP remodeling by hnRNP C may be important for telomerase RNP and/or telomere complex formation through the direct interaction between hnRNP C and the uridine-rich tract of telomerase RNA (49). The hnRNP C tetramer, together with hnRNP A/B proteins, binds to RNA as a core component of 40S hnRNP particles (20–23). Further studies are warranted to clarify whether hnRNP C-mediated RNP remodeling is executed by hnRNP C or 40S hnRNP.

ARS2 is recruited to CBC on nascent m^7G-capped RNA (50). ARS2 stimulates 3′-end processing of m^7G-capped RNA (50–52) and RNA degradation (53). Our *in vitro* RNA-binding assays using purified recombinant CBC, hnRNP C and ARS2 (54) showed that hnRNP C inhibited RNA binding of ARS2 (Supplementary Figure S6). This observation might suggest two functions of hnRNP C. First, hnRNP C navigates nascent transcripts to the mRNA export pathway more effectively by inhibiting the recruitment of both ARS2 and PHAX to long RNA. Second, hnRNP C regulates the stability of capped RNA through suppression of ARS2-mediated degradation and 3′-processing. HnRNP C may broadly control the composition of RNPs near the cap that influences the fate of RNA.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary data.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Dr Yoshinori Akiyama for technical advice and to Drs Tetsuro Hirose, Kensuke Ninomiya and Tomohiro Yamazaki for valuable discussions. We also thank the members of our laboratory, especially Drs Makoto Kitabatake and Mitsuhiro Machitani, for their critical comments on this manuscript. We thank Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.
Author contributions: I.T. and M.O. obtained funding and designed the study. S.D. and I.T. performed the experiments and analyzed data. S.D., M.O. and I.T. wrote the manuscript.

FUNDING

This work was supported by JSPS KAKENHI [25251004 to M.O.]; JSPS KAKENHI [21K06016 to I.T.]; and the Japan Foundation for Applied Enzymology (I.T.).
Conflict of interest statement. None declared.

REFERENCES

1. Jarmolowski,A., Boelens,W.C., Izaurralde,E. and Mattaj,I.W. (1994) Nuclear export of different classes of RNA is mediated by specific factors. *J. Cell Biol.*, **124**, 627–635.
2. Ohno,M., Segref,A., Kuersten,S. and Mattaj,I.W. (2002) Identity elements used in export of mRNAs. *Mol. Cell*, **9**, 659–671.
3. Jin,L., Guzik,B.W., Bor,Y., Rekosh,D. and Hammarskjöld,M.-L. (2003) Tap and NXT promote translation of unspliced mRNA. *Genes Dev.*, **17**, 3075–3086.
4. Wiegand,H.L., Lu,S. and Cullen,B.R. (2003) Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. *Proc. Natl Acad. Sci. USA*, **100**, 11327–11332.
5. Kuersten,S., Segal,S.P., Verheyden,J., LaMartina,S.M. and Goodwin,E.B. (2004) NXF-2, REF-1, and REF-2 affect the choice of nuclear export pathway for tra-2 mRNA in *C. elegans*. *Mol. Cell*, **14**, 599–610.
6. Li,M.W., Sletten,A.C., Lee,J., Pyles,K.D., Matkovich,S.J., Ory,D.S. and Schaffer,J.E. (2017) Nuclear export factor 3 regulates localization of small nucleolar RNAs. *J. Biol. Chem.*, **292**, 20228–20239.
7. Lei,H., Dias,A.P. and Reed,R. (2011) Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. *Proc. Natl Acad. Sci. USA*, **108**, 17985–17990.
8. Izaurralde,E., Lewis,J., McGuigan,C., Jankowska,M., Darzynkiewicz,E. and Mattaj,I.W. (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*, **78**, 657–668.
9. Hamm,J. and Mattaj,I.W. (1990) Monomethylated cap structures facilitate RNA export from the nucleus. *Cell*, **63**, 109–118.
10. Ohno,M., Segref,A., Bachi,A., Wilm,M. and Mattaj,I.W. (2000) PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell*, **101**, 187–198.
11. Cheng,H., Dufu,K., Lee,C.-S., Hsu,J.L., Dias,A. and Reed,R. (2006) Human mRNA export machinery recruited to the 5′ end of mRNA. *Cell*, **127**, 1389–1400.
12. Huang,Y. and Steitz,J.A. (2005) SRprises along a messenger's journey. *Mol. Cell*, **17**, 613–615.
13. Le Hir,H., Izaurralde,E., Maquat,L.E. and Moore,M.J. (2000) The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *EMBO J.*, **19**, 6860–6869.
14. Grüter,P., Tabernero,C., von Kobbe,C., Schmitt,C., Saavedra,C., Bachi,A., Wilm,M., Felber,B.K. and Izaurralde,E. (1998) TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell*, **1**, 649–659.

15. Le Hir,H., Gatfield,D., Izaurralde,E. and Moore,M.J. (2001) The exon–exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.*, **20**, 4987–4997.
16. Masuyama,K., Taniguchi,I., Kataoka,N. and Ohno,M. (2004) RNA length defines RNA export pathway. *Genes Dev.*, **18**, 2074–2085.
17. Fuke,H. and Ohno,M. (2008) Role of poly (A) tail as an identity element for mRNA nuclear export. *Nucleic Acids Res.*, **36**, 1037–1049.
18. McCloskey,A., Taniguchi,I., Shinmyozu,K. and Ohno,M. (2012) hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science*, **335**, 1643–1646.
19. Choi,Y.D. and Dreyfuss,G. (1984) Isolation of the heterogeneous nuclear RNA–ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. *Proc. Natl Acad. Sci. USA*, **81**, 7471–7475.
20. Samakina,O.P., Lukanidin,E.M., Molnar,J. and Georgiev,G.P. (1968) Structural organization of nuclear complexes containing DNA-like RNA. *J. Mol. Biol.*, **33**, 251–263.
21. Beyer,A.L., Christensen,M.E., Walker,B.W. and LeStourgeon,W.M. (1977) Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell*, **11**, 127–138.
22. Karn,J., Vidali,G., Boffa,L.C. and Allfrey,V.G. (1977) Characterization of the non-histone nuclear proteins associated with rapidly labeled heterogeneous nuclear RNA. *J. Biol. Chem.*, **252**, 7307–7322.
23. Huang,M., Rech,J.E., Northington,S.J., Flicker,P.F., Mayeda,A., Krainer,A.R. and LeStourgeon,W.M. (1994) The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles. *Mol. Cell. Biol.*, **14**, 518–533.
24. Venables,J.P., Koh,C.-S., Froehlich,U., Lapointe,E., Couture,S., Inkel,L., Bramard,A., Paquet,E.R., Watier,V., Durand,M. *et al.* (2008) Multiple and specific mRNA processing targets for the major human hnRNP proteins. *Mol. Cell. Biol.*, **28**, 6033–6043.
25. Choi,Y.D., Grabowski,P.J., Sharp,P.A. and Dreyfuss,G. (1986) Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. *Science*, **231**, 1534–1539.
26. Zarnack,K., König,J., Tajnik,M., Martincorena,I., Eustermann,S., Stévant,I., Reyes,A., Anders,S., Luscombe,N.M. and Ule,J. (2013) Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell*, **152**, 453–466.
27. Merrill,B.M., Barnett,S.F., LeStourgeon,W.M. and Williams,K.R. (1989) Primary structure differences between proteins C1 and C2 of HeLa 40S nuclear ribonucleoprotein particles. *Nucleic Acids Res.*, **17**, 8441–8449.
28. Dreyfuss,G., Matunis,M.J., Piñol-Roma,S. and Burd,C.G. (1993) hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.*, **62**, 289–321.
29. Swanson,M.S., Nakagawa,T.Y., LeVan,K. and Dreyfuss,G. (1987) Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. *Mol. Cell. Biol.*, **7**, 1731–1739.
30. König,J., Zarnack,K., Rot,G., Curk,T., Kayikci,M., Zupan,B., Turner,D.J., Luscombe,N.M. and Ule,J. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.*, **17**, 909–915.
31. Cieniková,Z., Jayne,S., Damberger,F.F., Allain,F.H.-T. and Maris,C. (2015) Evidence for cooperative tandem binding of hnRNP C RRMs in mRNA processing. *RNA*, **21**, 1931–1942.
32. Görlach,M., Wittekind,M., Beckman,R.A., Mueller,L. and Dreyfuss,G. (1992) Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. *EMBO J.*, **11**, 3289–3295.
33. Soltaninassab,S.R., McAfee,J.G., Shahied-Milam,L. and LeStourgeon,W.M. (1998) Oligonucleotide binding specificities of the hnRNP C protein tetramer. *Nucleic Acids Res.*, **26**, 3410–3417.
34. McAfee,J.G., Soltaninassab,S.R., Lindsay,M.E. and LeStourgeon,W.M. (1996) Proteins C1 and C2 of heterogeneous nuclear ribonucleoprotein complexes bind RNA in a highly cooperative fashion: support for their contiguous deposition on pre-mRNA during transcription. *Biochemistry*, **35**, 1212–1222.
35. McAfee,J.G., Shahied-Milam,L., Soltaninassab,S.R. and LeStourgeon,W.M. (1996) A major determinant of hnRNP C protein binding to RNA is a novel bZIP-like RNA binding domain. *RNA*, **2**, 1139–1152.
36. Jiang,W., Guo,X. and Bhavanandan,V.P. (1998) Four distinct regions in the auxiliary domain of heterogeneous nuclear ribonucleoprotein C-related proteins. *Biochim. Biophys. Acta*, **1399**, 229–233.
37. Barnett,S.F., Friedman,D.L. and LeStourgeon,W.M. (1989) The C proteins of HeLa 40S nuclear ribonucleoprotein particles exist as anisotropic tetramers of (C1)3 C2. *Mol. Cell. Biol.*, **9**, 492–498.
38. Whitson,S.R., LeStourgeon,W.M. and Krezel,A.M. (2005) Solution structure of the symmetric coiled coil tetramer formed by the oligomerization domain of hnRNP C: implications for biological function. *J. Mol. Biol.*, **350**, 319–337.
39. Shahied,L., Braswell,E.H., LeStourgeon,W.M. and Krezel,A.M. (2001) An antiparallel four-helix bundle orients the high-affinity RNA binding sites in hnRNP C: a mechanism for RNA chaperonin activity. *J. Mol. Biol.*, **305**, 817–828.
40. Kershaw,C.J. and O'Keefe,R.T. (2013) Splint ligation of RNA with T4 DNA ligase. *Methods Mol. Biol*., **941**, 257–269.
41. Castello,A., Fischer,B., Eichelbaum,K., Horos,R., Beckmann,B.M., Strein,C., Davey,N.E., Humphreys,D.T., Preiss,T., Steinmetz,L.M. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, **149**, 1393–1406.
42. Baltz,A.G., Munschauer,M., Schwanhäusser,B., Vasile,A., Murakawa,Y., Schueler,M., Youngs,N., Penfold-Brown,D., Drew,K., Milek,M. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell*, **46**, 674–690.
43. Wang,D.O., Ninomiya,K., Mori,C., Koyama,A., Haan,M., Kitabatake,M., Hagiwara,M., Chida,K., Takahashi,S.-I., Ohno,M. *et al.* (2017) Transport granules bound with nuclear cap binding protein and exon junction complex are associated with microtubules and spatially separated from eIF4E granules and P bodies in human neuronal processes. *Front. Mol. Biosci*., **4**, 93.
44. Barnett,S.F., LeStourgeon,W.M. and Friedman,D.L. (1988) Rapid purification of native C protein from nuclear ribonucleoprotein particles. *J. Biochem. Biophys. Methods*, **16**, 87–97.
45. Stone,J.R., Maki,J.L. and Collins,T. (2003) Basal and hydrogen peroxide stimulated sites of phosphorylation in heterogeneous nuclear ribonucleoprotein C1/C2. *Biochemistry*, **42**, 1301–1308.
46. Mayrand,S.H., Dwen,P. and Pederson,T. (1993) Serine/threonine phosphorylation regulates binding of C hnRNP proteins to pre-mRNA. *Proc. Natl Acad. Sci. USA*, **90**, 7764–7768.
47. Kattapuram,T., Yang,S., Maki,J.L. and Stone,J.R. (2005) Protein kinase CK1alpha regulates mRNA binding by heterogeneous nuclear ribonucleoprotein C in response to physiologic levels of hydrogen peroxide. *J. Biol. Chem.*, **280**, 15340–15347.
48. Piñol-Roma,S. and Dreyfuss,G. (1993) Cell cycle-regulated phosphorylation of the pre-mRNA-binding (heterogeneous nuclear ribonucleoprotein) C proteins. *Mol. Cell. Biol.*, **13**, 5762–5770.
49. Ford,L.P., Suh,J.M., Wright,W.E. and Shay,J.W. (2000) Heterogeneous nuclear ribonucleoproteins C1 and C2 associate with the RNA component of human telomerase. *Mol. Cell. Biol.*, **20**, 9084–9091.
50. Hallais,M., Pontvianne,F., Andersen,P.R., Clerici,M., Lener,D., Benbahouche,N.E.H., Gostan,T., Vandermoere,F., Robert,M.-C., Cusack,S. *et al.* (2013) CBC–ARS2 stimulates 3′-end maturation of multiple RNA families and favors cap-proximal processing. *Nat. Struct. Mol. Biol.*, **20**, 1358–1366.
51. Gruber,J.J., Zatechka,D.S., Sabin,L.R., Yong,J., Lum,J.J., Kong,M., Zong,W.-X., Zhang,Z., Lau,C.-K., Rawlings,J. *et al.* (2009) Ars2 links the nuclear cap-binding complex to RNA interference and cell proliferation. *Cell*, **138**, 328–339.
52. Gruber,J.J., Olejniczak,S.H., Yong,J., La Rocca,G., Dreyfuss,G. and Thompson,C.B. (2012) Ars2 promotes proper replication-dependent histone mRNA 3′ end formation. *Mol. Cell*, **45**, 87–98.
53. Giacometti,S., Benbahouche,N.E.H., Domanski,M., Robert,M.-C., Meola,N., Lubas,M., Bukenborg,J., Andersen,J.S., Schulze,W.M., Verheggen,C. *et al.* (2017) Mutually exclusive CBC-containing complexes contribute to RNA fate. *Cell Rep.*, **18**, 2635–2650.
54. Schulze,W.M., Stein,F., Rettel,M., Nanao,M. and Cusack,S. (2018) Structural analysis of human ARS2 as a platform for co-transcriptional RNA sorting. *Nat. Commun.*, **9**, 1701.