Published online 28 May 2020

Nucleic Acids Research, 2020, Vol. 48, No. 12 ***6943–6953***
doi: 10.1093/nar/gkaa445

Mapping domains of ARS2 critical for its RNA decay capacity

Mireille Melko, Kinga Winczura, Jérôme Olivier Rouvière, Michaela Oborská-Oplová, Pia K. Andersen and Torben Heick Jensen*

Department of Molecular Biology and Genetics, Aarhus University, C.F. Møllers Allé 3, Building 1130, 8000 Aarhus C, Denmark

Received September 10, 2019; Revised May 11, 2020; Editorial Decision May 13, 2020; Accepted May 14, 2020

ABSTRACT

ARS2 is a conserved protein centrally involved in both nuclear RNA productive and destructive processes. To map features of ARS2 promoting RNA decay, we utilized two different RNA reporters, one of which depends on direct ARS2 tethering for its degradation. In both cases, ARS2 triggers a degradation phenotype aided by its interaction with the poly(A) tail exosome targeting (PAXT) connection. Interestingly, C-terminal amino acids of ARS2, responsible for binding the RNA 5′cap binding complex (CBC), become dispensable when ARS2 is directly tethered to the reporter RNA. In contrast, the Zinc-finger (ZnF) domain of ARS2 is essential for the decay of both reporters and consistently co-immunoprecipitation analyses reveal a necessity of this domain for the interaction of ARS2 with the PAXT-associated RNA helicase MTR4. Taken together, our results map the domains of ARS2 underlying two essential properties of the protein: its RNP targeting ability and its capacity to recruit the RNA decay machinery.

INTRODUCTION

RNA metabolism in mammalian nuclei is a complex process involving both productive and destructive activities. Several factors and distinct macromolecular complexes engage in the synthesis, processing and degradation of a myriad of different RNA species (1–4). From the RNA destructive side, the evolutionarily conserved 3′-5′ exo- and endo-nucleolytic RNA exosome is the major nuclear machinery carrying out both the processing and complete decay of various RNA substrates (4–6). In human nuclei the exosome consists of a catalytically inactive core that associates with the exonuclease RRP6 (EXOSC10) and the exo- and endo-nuclease RRP44 (DIS3). Besides these enzymatic co-factors, RNA-binding adaptor proteins are required to direct the nuclear exosome to its substrates (5). Here, the RNA helicase MTR4 (SKIV2L2/MTREX) is crucial for exosome activity through, on the one hand, its unwinding of RNAs, enabling their access to the central channel of the exosome core, and, on the other hand, by associating with specific RNA binding adaptor proteins, providing target specificity (7,8). In the nucleoplasm, human MTR4 has been found associated with two adaptors, the nuclear exosome targeting (NEXT) complex (9) and the poly(A) tail exosome targeting (PAXT) connection (10). Within NEXT, the zinc-knuckle (ZnK) protein ZCCHC8 and the RNA-binding protein RBM7 associate with MTR4 to promote the exosomal degradation of rather short and immature RNAs such as some promoter-upstream transcripts (PROMPTs), enhancer RNAs (eRNAs) and 3′extended snRNAs and snoRNAs (Supplementary Figure S1A) (9–14). In the PAXT connection, the zinc-finger (ZnF) protein ZFC3H1 interacts tightly with MTR4, and more loosely with the nuclear poly(A)-binding protein (PABPN1), to target more polyadenylated transcripts (Supplementary Figure S1A) (10). Interestingly, the destructive NEXT and PAXT assemblies both appear to interact physically with factors involved in RNA production; most notably the 5′ m^7G-cap-binding complex (CBC) and its associated factors (10,12,15), and it has been suggested that such intersection between destructive and productive factors helps to facilitate the proper sorting of nuclear transcripts (4,16–19).

The 5′cap structure is added to all nascent RNA polymerase II (RNAPII) transcripts 20–50 nucleotides after their transcription initiation. Shortly thereafter, the CBC, consisting of cap-binding proteins 20 and 80

*To whom correspondence should be addressed. Tel: +45 6020 2705; Fax: +45 8619 6500; Email: thj@mbg.au.dk
Present addresses:
Kinga Winczura, School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.
Michaela Oborská-Oplová, Institute of Medical Microbiology, University of Zurich, CH-8006 Zurich, Switzerland.
Pia K. Andersen, Nationalt Genom Center, Department of Molecular Medicine, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark.

(CBP20/NCBP2 and CBP80/NCBP1), binds tightly to the cap to prevent its decapping and also serving to recruit various proteins that will eventually determine the fate of the transcript (20–22). A critical factor that interacts early with the CBC, and its capped nascent transcript, is the highly conserved ARS2 (SRRT) protein (12,15,23). This predominantly nuclear protein is essential for cellular proliferation and early mammalian development, most probably due to its role in RNA metabolism (23,24). ARS2 acts as a scaffold, binding both RNA and protein, linking the CBC to other factors and complexes, which facilitate transcription termination as well as RNA 3′-end processing, maturation, degradation and export (25). Indeed, the CBC-ARS2 (CBCA) complex stimulates cap-proximal transcription termination of short transcripts such as snRNAs, replication-dependent histone (RDH) RNAs and PROMPTs (12,15,26). CBCA may then connect with the ZnF protein ZC3H18 to target these RNAs for exosomal trimming, or complete decay, by mediating interactions with the NEXT complex (forming the CBC-NEXT (CBCN) complex) or the PAXT connection (10,12). Alternatively, ARS2 may stimulate the binding of the RNA transport factor PHAX to the CBC, resulting in the formation of the CBCA-PHAX (CBCAP) complex and the intra-nuclear transport of snoRNAs or the nuclear export of snRNAs (15). Instead, mRNA-bound CBCA complexes also recruit the mRNA export factor ALYREF (27). Collectively, biochemical, structural and cell biological analyses of CBCA complexes suggest that they offer dynamic platforms of protein-protein interactions, which eventually determine RNA fate (16,28,29). As an example, the CBCA complex interacts, in a mutually exclusive manner, with PHAX or ZC3H18, directing the transcript for either nuclear transport or degradation, respectively (16).

Despite its central position in nuclear RNA metabolism, relatively little is known about which domains of ARS2 perform its molecular interactions. Recent structural studies showed that ARS2 folds into a structured core, that comprises a helical 'body' projecting N- and C- terminal 'legs', constituted by a C-terminal ZnF domain, suggested to be involved in protein-protein interactions and an N-terminal domain of unknown function (DUF) (28). In addition, ARS2 harbors three unstructured stretches, the N- and C-terminal tails that hold, respectively, the nuclear localization signal (NLS) and the CBC-interacting domain, and a glutamate-rich region of unknown function positioned in the middle of the protein (29,30). Finally, ARS2 harbors an RNA recognition motif (RRM) that binds single stranded RNA (28).

To better understand the domain requirements of ARS2 in RNA metabolism, we assessed its impact on the expression of two distinct RNA reporters, both of which are targeted for exosomal decay by the PAXT connection. We demonstrate that the CBC-binding ability of ARS2 becomes dispensable when ARS2 is directly tethered to the RNA. Moreover, the ARS2 ZnF domain is essential for RNA degradation of both reporters and for the interaction of ARS2 with the RNA decay machinery. Collectively, our analyses highlight essential contacts for recruitment of RNA decay activity to nuclear RNAs.

MATERIALS AND METHODS

Cell lines and culture

HEK293 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Invitrogen) at 37°C, 5% CO_2. For transient transfection, TurboFect Transfecting Reagent (Thermo Scientific) was used according to the manufacturer's recommendations. GFP-fused ARS2 FL, fragments F1-F6 and mutants DUF, ZnF, M and ΔC were stably expressed in a HEK293 Flp-In T-REx tetracyclin-inducible cell line. Fusion proteins were introduced using a modified pcDNA5/FRT/TO vector that contains an N-terminal GFP-tag followed by the protein of interest. HEK293 0-LpA and Histone cell lines were described in (31). All HEK293 Flp-In T-REx tetracyclin-inducible cell lines were maintained in DMEM containing 10% tetracycline-free FBS (Invitrogen), 1% penicillin/streptomycin, 5 μg/ml Blasticidin (Invitrogen) and 100 μg/ml Hygromycin B (Invitrogen).

Inducible reporter and protein expression

Expressions of β-globin-MS2 (32) and 0-LpA/Histone (31) reporters were induced by adding tetracycline (Sigma) to the media at a final concentration of 250 ng/ml, respectively 4 and 20 h before the cell harvest. Expression of GFP-tagged fusion proteins was induced by replacing cell growth media with fresh media containing tetracycline. For immunoprecipitation experiments, cells were induced with 5 ng/ml of tetracycline.

Cloning

ARS2 fragments were cloned using Gateway cloning technology (Invitrogen). Fragment sequences were first amplified by PCR using primers listed in Supplementary Table S1, then transferred to the pDONR223 vector by recombination reaction (Invitrogen) and then finally cloned by recombination into pBS-L30-MS2-GFP (kindly provided by E. Bertrand lab) or pgLAP1 (pcDNA5/FRT/TO-N-ter GFP). ARS2 DUF and ZnF mutants were made by site-directed mutagenesis using QuikChange II site-directed mutagenesis (Agilent Technologies) and the primers listed in Supplementary Table S2. ARS2 M and ΔC mutants were subcloned into pBS-L30-MS2-GFP and pgLAP1 plasmids using primers listed in Supplementary Table S1.

RNAi and plasmids transfections

For RNAi assays, cells were seeded in Collagen (Sigma) pre-coated dishes and grown for 24 h before the first transfection. siRNA transfections were performed using siLentFect Lipid Reagent (Bio-Rad) according to the manufacturer's instructions and siRNAs at a final concentration of 20 nM. Forty eight hours after the first transfection, media was changed and the transfection was repeated. At least 5 h after each transfection, media was replaced with fresh media. Cells were harvested 24 h after the second transfection. siRNA sequences are listed in Supplementary Table

S3. Plasmids transfections were performed using Turbofect Transfecting Reagent (Thermo Scientific) according to the manufacturer's recommendations, 48 hours before the cell harvest. Prior to cell harvest, plasmid transfection efficiencies were monitored by fluorescence microscopy and found to be constant between 70 and 80%.

RNA isolation, northern blotting analysis and reverse transcription analysis

RNA was purified with TRIzol reagent (Invitrogen) according to the manufacturer's instructions. For northern blotting analyses, 5–10 μg of RNA per sample was loaded and separated on a 1.2 % agarose gel, containing 1× MOPS and 6% formaldehyde, and transferred to a positively charged nylon membrane. After membrane blocking with the corresponding hybridization buffer, hybridization was performed in RapidHyb buffer (Ambion) at 68°C for β-globin and GAPDH RNAs and in ULTRAHyb buffer at 42°C for 0-LpA, Histone and 18S RNAs and according to the manufacturer's guidelines. The Strip-EZ RNA kit (Ambion) was used to transcribe ^{32}P-labelled riboprobes (β-globin and GAPDH). ^{32}P-5′labelled oligoprobes were used to detect 0-LpA, Histone, 18S and GAPDH RNAs in the 0-LpA experiments. RNA was visualized and quantified using phosphorimager analysis and Quantity One software (Bio-Rad). For RT-qPCR analyses, RNA was DNAse I treated using PureLink DNase kit (Invitrogen). 1 μg of RNA was subsequently used for cDNA synthesis with random hexamers and anchored oligo-dT primers. Reverse transcription was performed using SuperscriptIII (Invitrogen) according to the manufacturer's instructions. qPCR reactions were set up in total volumes of 15 μl with 5 μl cDNA sample, 2,5 μM primer pair and 7.5 μl 2× Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen) and run as standard short PCR programs with an annealing temperature of 60°C. qPCR was performed with the primers listed in Supplementary Table S4. Data were processed with the ddCt method, with normalization to both GAPDH mRNA levels and eGFP siRNA control samples. Error bars represent standard deviations calculated from at least three independent experiments.

Fluorescence microscopy

HeLa cells were grown on cover slips in 24-well plates in DMEM with 10% FBS. Cells were transfected with 1 μg plasmid constructs using 2 μl Turbofect (Thermo Scientific). After 24 h of transfection, cells were washed twice with cold PBS, then fixed in 4% paraformaldehyde for 20 min at RT. Cells were then permeabilized with PBS + 0.5% Triton X-100 for 15 min at room temperature (RT) and then incubated with 1 μg/μl 4′,6-diamidino-2-phenylindole (DAPI) for 10 min at RT. Cover slips were then washed three times in PBS at RT and mounted on glass slides using Slow Fade Gold Antifade Mountant (Invitrogen). Cells were observed under a Zeiss 40× objective.

Immunoprecipitation analysis

Cells from p10 plates were lysed in Protease Inhibitor Cocktail (Roche) containing lysis buffer (20 mM Hepes/NaOH pH 7.4, 150 mM NaCl, 0.5% Triton X-100), and sonicated 3 × 10 s. Cell lysates were clarified by centrifugation 13 000 rpm, for 10 min at 4°C. Immunoprecipitations were performed using pre-washed anti-GFP nanobodies coupled to magnetic beads for 1 h in a cold room. The beads were then washed three times with ice-cold lysis buffer and eluted with NuPAGE LDS sample buffer for 10 min at 70°C. The samples were finally analyzed by SDS-PAGE and western blotting analysis.

Western blotting analysis

Protein expression or depletion efficiencies were monitored by western blotting analysis. Cell pellets were resuspended in RSB100 (10 mM Tris pH 7.5, 100 mM NaCl, 2.5 mM $MgCl_2$, 0.5% Triton X-100) and sonicated. Cell lysates were cleared by centrifugation at 10 000 rpm, 4°C for 15 min. Protein extracts were separated on NuPage 4–12% Bis–Tris gels (Invitrogen) and transferred onto PVDF membrane (Millipore) by wet transfer with XCell II Blot Module (Invitrogen) for 2,5 h at 4°C. After transfer, membranes were blocked for 1 h with 5% skimmed milk in PBS buffer (10 mM Tris–HCl, pH 8.0; 0.15 M NaCl). Membranes were then incubated overnight at 4°C with the primary antibodies diluted in PBS/5% milk/0.05% Tween 20 as described in Supplementary Table S5. The next day, membranes were washed with PBS/0.05% Tween 20 and then incubated with horse-radish-peroxidase (HRP) conjugated goat-anti-rabbit or -mouse secondary antibody (Dako) diluted in PBS/5% milk/0.05% Tween 20 as described in Supplementary Table S5. Membranes were washed again and exposed using Supersignal West Femto reagent (Thermo Scientific).

RESULTS

Establishing ARS2-dependent RNA reporter constructs

We previously showed that ARS2 connects the RNA 5′cap to the nuclear exosome, promoting RNA processing or complete degradation via the CBCN complex (10,12,33). To conduct a more comprehensive analysis of the domains of ARS2 involved in RNA degradation, we first evaluated the direct impact of tethering ARS2 fused to the MS2 phage coat protein to a tetracycline inducible and intron-containing β-globin mini-gene reporter RNA, harboring six MS2 binding sites ('β-globin-MS2', Figure 1A). HEK293 Flp-In T-REx cells were used to transiently co-express an MS2-GFP-ARS2 fusion protein (Figure 1A and B, lane 2) together with the β-globin-MS2 reporter induced for 4 h. As assessed by northern blotting analysis, tethering of MS2-GFP-ARS2 considerably lowered β-globin-MS2 reporter levels when compared to a MS2-GFP control (Figure 1C, left panel compare lanes 1 and 2). This effect was MS2-tethering dependent, since GFP-ARS2 protein expressed without the MS2 epitope (Figure 1B, lane 3) only marginally affected reporter RNA levels (Figure 1C, left panel compare lanes 1 and 3) and since decreased reporter RNA levels depended on its presence of MS2 sites (Figure 1C, right panel).

ARS2 facilitates the transcription termination and exosomal degradation of short RNAs, such as PROMPTs

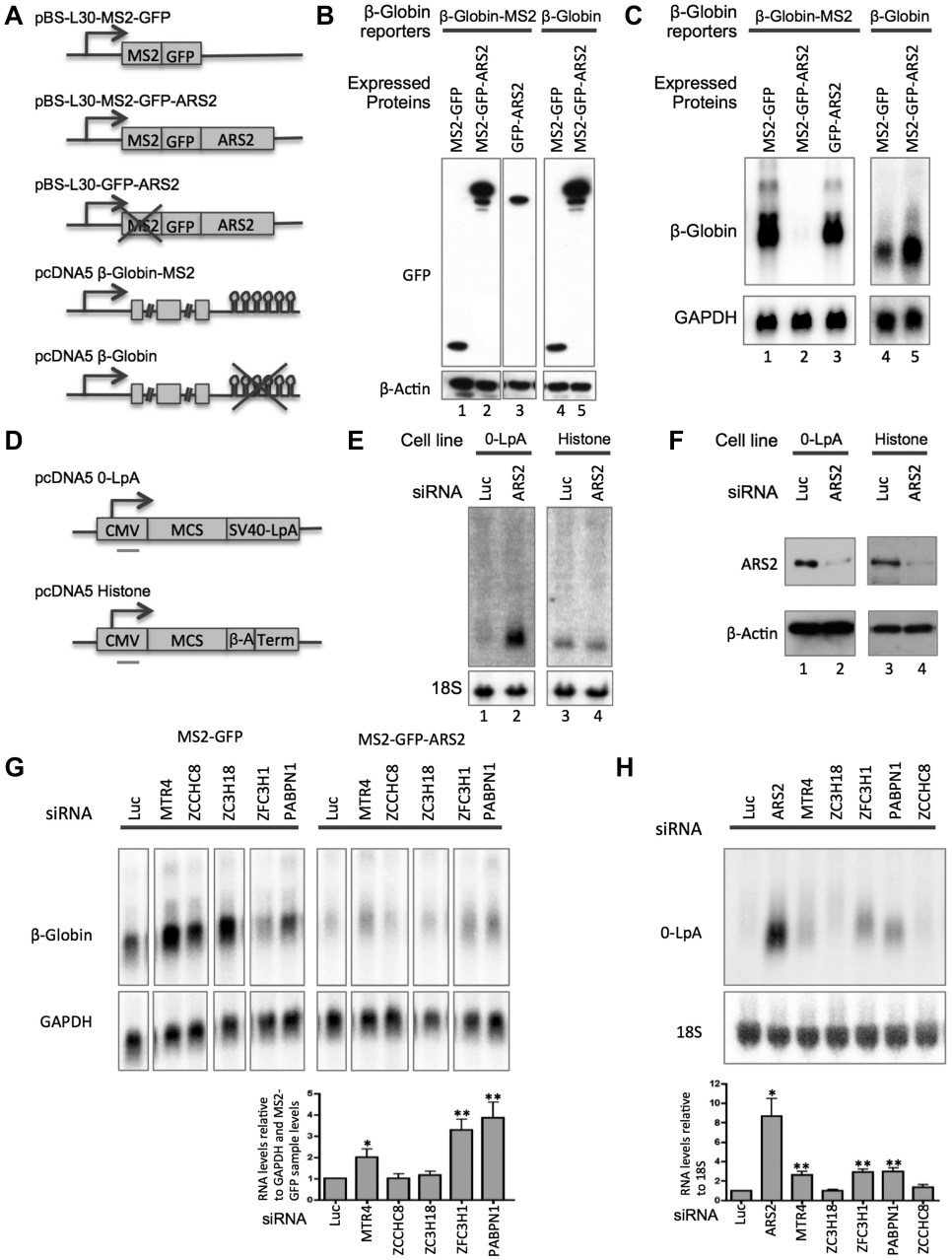

Figure 1. Establishing ARS2-dependent RNA reporter constructs. (**A**) Schematic representation of the utilized pBS-L30-MS2(or ΔMS2)-GFP-ARS2(or ΔARS2) expression constructs and pcDNA5-β-globin-MS2(or ΔMS2) reporters. (**B**) Western blotting analysis of extracts obtained from HEK293 cells transfected with β-globin reporters and MS2-GFP, MS2-GFP-ARS2 or GFP-ARS2 expression constructs as indicated. GFP-tagged proteins were detected by an α-GFP antibody while β-actin was probed as a loading control. (**C**) Northern blotting analysis of total RNA purified from cells from (B) and subjected to 4hours of β-globin-MS2 (or ΔMS2) gene induction. Membranes were hybridized with probes directed against β-globin (top) and GAPDH (bottom) RNA. (**D**) Schematic representation of 0-LpA and Histone terminator constructs (31). MCS: Multiple Cloning Site of pcDNA5 FRT/TO. β-A: Spacer consisting of 119 reverse complement bp of the ß-Actin gene. Positioning of the utilized probe for northern blotting analysis is indicated. (**E**) Northern blotting analysis of total RNA purified from HEK293 cells stably integrated with LpA- or Histone-terminator constructs from (D) and subjected to ARS2 depletion or control Luciferase siRNA treatment ('Luc') as indicated. Cells were subjected to 20 h of CMV-directed induction. Membranes were hybridized with the CMV-directed probe from (D) and 18S RNA as a loading control. (**F**) Western blotting analysis of extracts obtained from (E). ARS2 depletion was verified by an α-ARS2 antibody while β-actin was probed as a loading control. (**G**) Northern blotting analysis of total RNA purified from siRNA depleted HEK293 cells, that were transfected with β-globin reporter and MS2-GFP or MS2-GFP-ARS2 expression constructs, as indicated. β-Globin-MS2 gene was induced for 4 h. Membranes were probed against β-globin-MS2 (top) and GAPDH (bottom) RNA. The lower graphic displays quantification of β-globin-MS2 RNA from MS2-GFP-ARS2 samples normalized first to levels of GAPDH RNA, and subsequent normalization to its control sample within the MS2-GFP group. (**H**) Northern blotting analysis of total RNA purified from siRNA depleted HEK293 0-LpA cells as indicated, and after 20 h of CMV-directed induction. Membranes were probed against 0-LpA (top) and 18S (bottom) RNAs. The lower graphic displays quantification of the 0-LpA transcript signal normalized to 18S RNA. For (G) and (H), an unpaired *t*-test was used to determine the significance levels between control (Luc) vs. decay factor-depleted samples. Error bars represent standard deviations of three experiments. Asterisks correspond to the following two-tail *P*-values: $*P \leq 0.05$; $**P \leq 0.01$.

(12). Hence, to interrogate a PROMPT-like and ARS2-dependent reporter, devoid of splicing signals, we employed a previously utilized exosome sensitive construct (34), consisting of a short DNA sequence derived from the pcDNA5 multiple cloning site stably integrated into HEK293 Flp-In T-REx cells under the control of a tetracycline-inducible CMV promoter and harboring an SV40-late polyadenylation (LpA) signal (Figure 1D, (31)). Northern blotting analysis revealed low steady-state levels of this '0-LpA' RNA, which were strongly increased upon depletion of endogenous ARS2 protein (Figure 1E, compare lanes 1 and 2, and see Figure 1F for ARS2 depletion efficiency). In contrast, ARS2 depletion did not alter the expression levels of a similar reporter RNA, harboring an RDH terminator signal in place of the SV40 LpA site (Figure 1E, lanes 3 and 4).

To uncover the molecular pathway(s) associated with the observed ARS2-dependent alterations of RNA reporter levels, we measured their abundances upon depletion of selected RNA degradation factors. To this end, HEK293 cells, transiently co-expressing the β-globin-MS2 reporter with either MS2-GFP or MS2-GFP-ARS2 fusion protein, were treated with siRNAs targeting either factors involved in the NEXT or PAXT decay pathways, respectively, or a Firefly Luciferase (Luc) RNA, as a control (Supplementary Figure S1B). Using northern blotting analysis to measure β-globin-MS2 RNA levels in cells, expressing either MS2-GFP or MS2-GFP-ARS2, revealed that depletion of components involved in the PAXT pathway (MTR4, ZFC3H1 and PABPN1) partly relieved the RNA suppressive effect of MS2-GFP-ARS2 tethering (Figure 1G, right top panel). This was also the case when normalizing to general depletion effects as measured in MS2-GFP samples (Figure 1G, left top panel and right bottom panel). Similarly, HEK293 0-LpA cells depleted for the same factors (Supplementary Figure S1C) displayed moderately increased reporter RNA levels when treated with siRNAs targeting MTR4, ZFC3H1 and PABPN1 (Figure 1H), whereas no effect on neither of the reporters could be detected upon depletion of ZC3H18 or the NEXT component ZCCHC8 (Figure 1G and H). We conclude that the two employed reporters constitute a good starting point for assessing RNA metabolic functions of ARS2 and that at least a part of the observed effects are based on the implication of ARS2 in PAXT-mediated RNA decay.

The ARS2 C-terminus is involved in RNA decay

To identify which part of the multi-domain ARS2 protein engages in RNA decay, we started a mutational analysis by dissecting the protein into different fragments (F1-F6), encompassing individual, or multiple, of the previously reported ARS2 domains (Figure 2A, (28,30)). Full length ARS2 and its constructed variants were all fused to an N-terminal MS2-GFP tag to facilitate cellular localization analysis and subsequent interrogation of possible biological activity. Subcellular localization was investigated by transient transfection of HeLa cells with plasmids expressing the various MS2-GFP-ARS2 variants and revealed, that, like full-length ARS2 protein, F1 and F5 variants appeared nuclear, confirming the presence of an NLS in the N-terminus of ARS2 (Figure 2B, (30)). Consistently, F2, F3 and F6 variants all appeared strictly cytoplasmic. Surprisingly, however, the F4 variant, although lacking the N-terminal NLS, localized to both the cytoplasm and the nucleus (Figure 2B). Hence, some NLS activity also resides in the ARS2 C-terminus and/or this fragment enters the nucleus in complex with other factors, e.g. the CBC.

We next aimed to test the functional relevance of the produced ARS2 variants. To ensure their proper nuclear localization, each of these was fused to a classic NLS sequence, PKKKRKV, derived from the SV40 Large T-antigen (Supplementary Figure S2A). Additionally, GFP western blotting analysis was conducted, which demonstrated the expression of all variants, albeit to varying extents (Figure 2C and D). The ability of each fragment to affect the expression levels of both of our established ARS2-reporters was then tested. In HEK293 cells transiently expressing the β-globin-MS2 reporter, tethering of F4, the smallest C-terminal fragment of ARS2, partially decreased reporter RNA levels (Figure 2E, compare lanes 1 and 6). A similar result was observed when the largest C-terminal fragment, F6, was tethered. These results map the ability of ARS2 to engage in RNA decay to the C-terminal quarter of the protein, encompassing the ZnF domain. Employing the LpA reporter in HEK293 LpA cells, we then tested the ability of the ARS2 variants to suppress the effect of endogenous ARS2 depletion. As expected, exogenous full-length ARS2 rescued the ARS2-depletion phenotype by lowering LpA RNA reporter levels (Figure 2F, compare lanes 9 and 10). Among the ARS2 fragments, only the minimal C-terminal fragment, F4, displayed ARS2 activity on this reporter. While it is unclear to us why the F6 variant did not display a similar activity in this assay, our collective analyses define the C-terminal domain as essential for the function of ARS2.

Defining ARS2 C-terminal requirements

To further investigate the role of the relevant functional domains of ARS2 in RNA decay, we focused on three distinct features of this region of the protein: i) The ZnF domain, which was previously suggested to be required for RNA– and protein–protein interactions (28,30), ii) The DUF domain, which was previously shown to bind to miRNAs and histone mRNAs (30) and iii) The last ∼27 amino acids of the protein, which are engaged in binding to the CBC (28). Introducing substitution of three conserved lysine residues, K698A, K719A and K743A, within the ZnF domain (Figure 3A, (30)) revealed that this nuclear (Figure 3B, top panel), and well expressed (Supplementary Figure S3A, lane 3), ARS2 variant did not repress β-globin-MS2 RNA levels when tethered to this reporter construct (Figure 3C, compare lanes 2 and 3). Moreover, the mutant did not support degradation of the 0-LpA RNA after depletion of endogenous ARS2 protein (Figure 3D, compare lanes 2 and 3). We therefore conclude that the ZnF domain is central to ARS2 function. As for the DUF domain, the substitution of the three conserved tyrosines, Y172F, Y175F and Y201F (Figure 3A and B, bottom panels) did not disrupt ARS2 function in the β-globin-MS2 reporter context (Figure 3C, compare lanes 5 and 6), and it did not substantially recover ARS2 activity in the 0-LpA reporter assay, despite a mild decrease in the reporter level (Figure 3D, compare

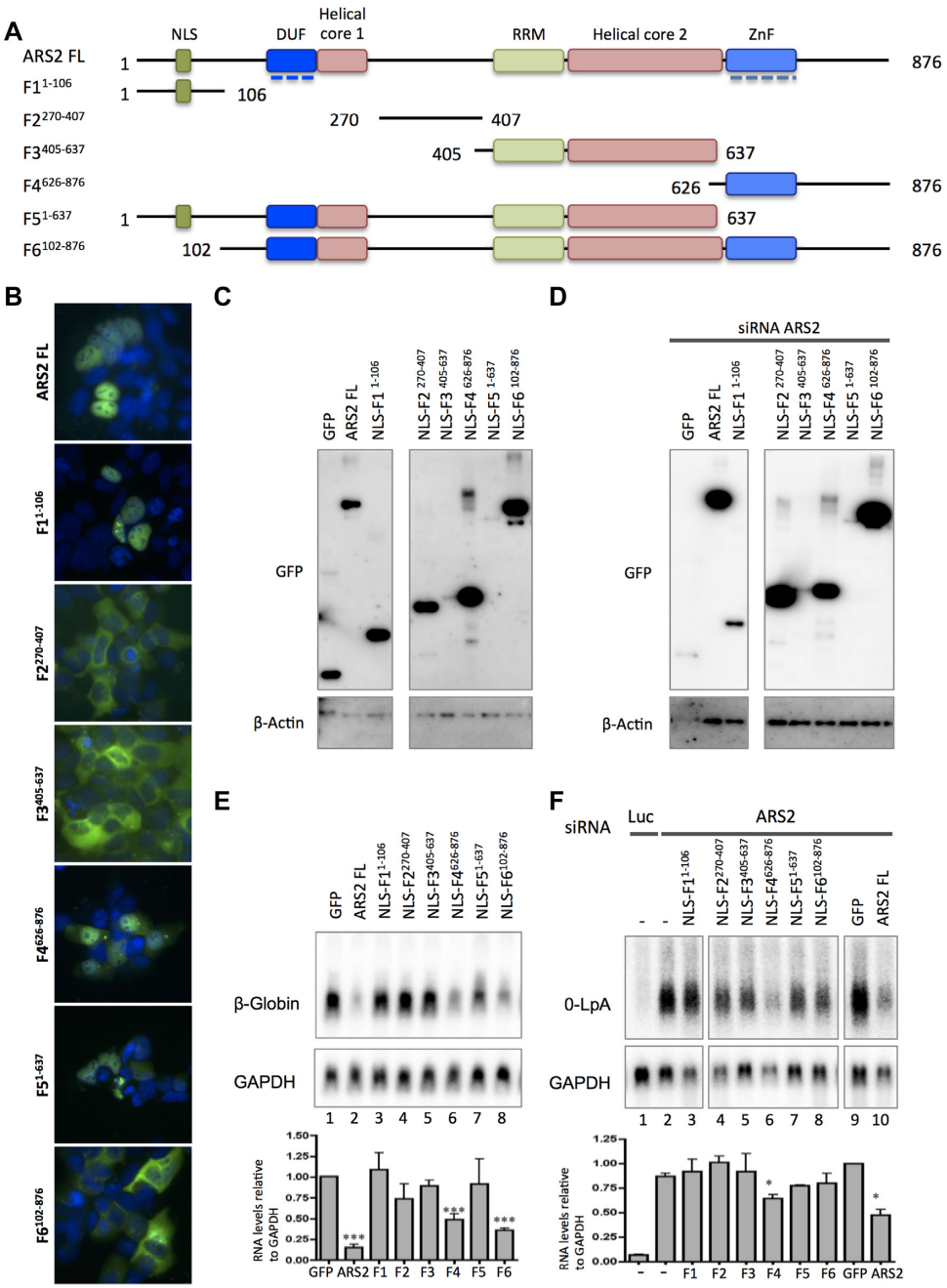

Figure 2. The ARS2 C-terminus is involved in RNA decay. (**A**) Schematic representation of full-length (FL) ARS2 and its derived F1–F6 variants. ARS2 was dissected into four distinct fragments F1–F4 harboring, respectively, the unstructured N-terminus containing a nuclear localization signal (NLS), an unstructured middle region, the RRM and the helical core 2 domains and, finally, the ZnF domain and the unstructured C-terminus. Additionally, one C-terminal truncation, F5, and one N-terminal truncation, F6, were produced. All constructs were N-terminally fused to GFP-MS2, GFP-MS2-NLS or GFP-NLS. The two dashed lines below the DUF and ZnF domains represent, respectively, the N- and C-terminal 'leg' projections of the structured protein (28). (**B**) Microscopic analysis of HeLa cells transiently expressing GFP-MS2 version of the ARS2 variants from (A). The images shown correspond to merged fluorescent GFP and DAPI channels. (**C**) Western blotting analysis of extracts obtained from HEK293 cells transfected with MS2-GFP, MS2-GFP-ARS2 or MS2-GFP-NLS-ARS2 fragment expression constructs as indicated. (**D**) Western blotting analysis of extracts obtained as in (B) but in HEK293 0-LpA cell line after depletion of endogenous ARS2. (C, D) GFP-tagged proteins were detected by an α-GFP antibody. β-Actin was probed as a loading control. All samples were run on, and blotted from, the same gel. (**E**) Northern blotting analysis of total RNA purified from HEK293 cells transfected with the β-globin-MS2 reporter and MS2-GFP, MS2-GFP-ARS2 or MS2-GFP-NLS-fused to F1–F6 expression constructs, as indicated. The β-globin-MS2 gene was induced for 4 h. Membranes were probed against β-globin-MS2 (top) and GAPDH (bottom) RNA. The lower graphic displays quantification of β-globin-MS2 transcript signal normalized to GAPDH. Error bars represent standard deviations of three experiments. Asterisks correspond to *P*-values from an unpaired *t*-test relative to the GFP control with the following two-tail *P*-values: *** $P \leq 0.001$. (**F**) Northern blotting analysis of total RNA purified from siRNA depleted HEK293 0-LpA cells transiently transfected with MS2-GFP, MS2-GFP-ARS2 or MS2-GFP-NLS-fused to F1-F6 expression constructs, as indicated. The 0-LpA reporter was induced for 20 h. Membranes were probed against 0-LpA (top) and GAPDH (bottom) RNAs. The lower graphic displays quantification of 0-LpA transcript signal normalized to GAPDH RNA. Error bars represent standard deviations of two experiments. Asterisks correspond to *P*-values from an unpaired *t*-test relative to the GFP controls with the following two-tail *P*-value: * $P \leq 0.05$.

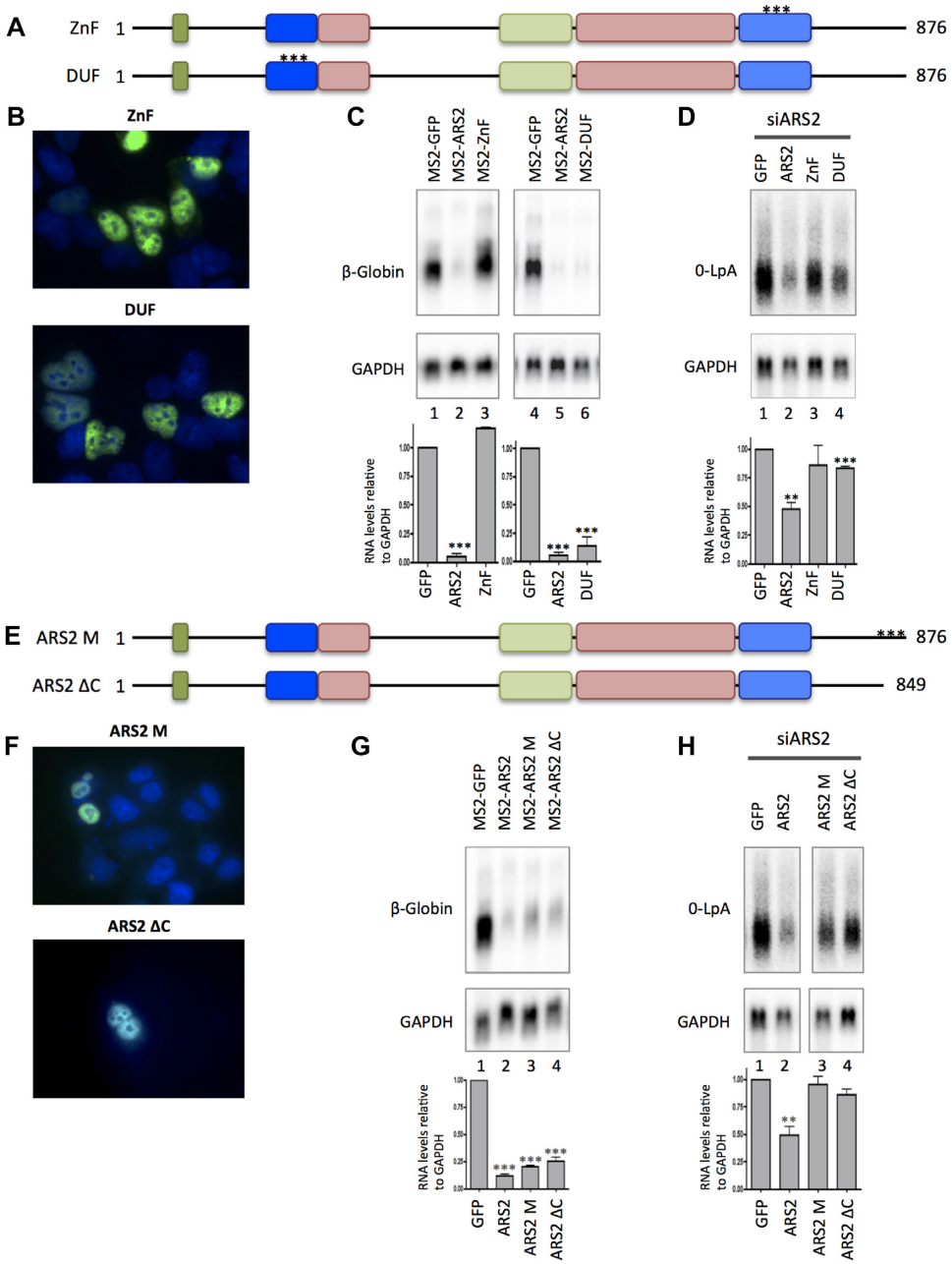

Figure 3. Defining ARS2 C-terminal requirements. (**A**) Schematic representation of ARS2 ZnF and DUF mutants, which were N-terminally fused to GFP-MS2 and GFP. *** designates positions of amino-acid substitutions. (**B**) Microscopic analysis of HeLa cells transiently expressing GFP-MS2-ZnF or GFP-MS2-DUF as indicated. Merged fluorescent images of GFP and DAPI channels are shown. (**C**) Northern blotting analysis of total RNA purified from HEK293 cells transfected with the β-globin-MS2 reporter together with MS2-GFP, MS2-ARS2, MS2-GFP-ZnF (left panel) or MS2-GFP-DUF (right panel) expression constructs, as indicated. β-globin-MS2 gene induction was performed for 4 hours. Membranes were probed against β-globin (top) and GAPDH (bottom) RNAs. (**D**) Northern blotting analysis of total RNA extracted from HEK293 0-LpA cells depleted for endogenous ARS2 and transiently transfected with GFP, GFP-ARS2, GFP-ZnF or GFP-DUF expression constructs, as indicated. The 0-LpA reporter was induced for 20 hours. Membranes were probed against 0-LpA (top) and GAPDH (bottom) RNAs. (**E**) Schematic representation of ARS2 M and ΔC mutants, which were N-terminally fused to GFP-MS2 and GFP. *** designates positions of amino-acid substitutions. (**F**) Microscopic analysis as in (B). (**G**) Northern blotting analysis as in (C) but using the indicated expression constructs. (**H**) Northern blotting analysis as in (D) but using the indicated expression constructs. For (C), (D), (G) and (H) the lower graphics display quantification of transcript signals normalized to the GAPDH loading control. Error bars represent standard deviations of 2 (D and H) or 3 (C and G) experiments. Asterisks correspond to *P*-values from an unpaired t-test relative to the GFP controls with the following two-tail *P*-values: $^{**}P \leq 0.01$; $^{***}P \leq 0.001$.

lanes 2 and 4). Thus, the ARS2 DUF domain appears to be required for full ARS2 activity, but it becomes dispensable when tethering the mutant protein to RNA.

Addressing the importance of the extreme C-terminus for ARS2 function, we employed two mutant variants: (i) the ARS2-M construct, harboring alteration of two C-terminal amino acids (R858A and Y863A) and (ii) ARS2-ΔC ($ARS2^{1-849}$), lacking the entire CBC-interacting domain (28) (Figure 3E). Both of these variants localized to the nucleus (Figure 3F) and were expressed at levels comparable to full length ARS2 (Supplementary Figure S3B). Both variants also performed like full-length ARS2 by decreasing β-globin-MS2 reporter levels in a tethering experiment (Figure 3G, compare lane 2 with lanes 3 and 4), indicating that CBC interaction is not important in this context. Instead, both of these CBC-binding mutants proved incapable of suppressing the ARS2-depletion phenotype when assayed in the context of the 0-LpA RNA reporter system (Figure 3H). We surmise that the CBC is critical for ARS2 recruitment, which is probably particularly relevant for short RNA substrates and which can be overcome when ARS2 is artificially recruited by MS2 tethering.

The ARS2 ZnF domain recruits the RNA decay machinery

Our results so far indicated that the C-terminal part of ARS2 (the F4 fragment), and more specifically its ZnF domain, is a major determinant in the triggering of RNA decay. Instead, its CBC-interaction domain is required for target recognition, presumably via binding to the CBC. To reconcile these activities with the ability, or inability, of the respective epitopes to interact with relevant co-factors, we stably integrated GFP-NLS-fused ARS2 mutant variants into HEK293 Flp-In T-REx cells and conducted GFP immunoprecipitation (IP) experiments. Considering that ARS2 interaction with the CBC, NEXT and PAXT components is RNA-independent (12), IPs were not subjected to RNase treatment. Assessed by western blotting analysis, all ARS2 variants, except for F1 and F3, were efficiently IP'ed (Figure 4A, compare left and right panels). As expected, full-length ARS2 co-purified the interrogated RNA decay factors MTR4 and ZC3H18, as well as the CBC protein CBP80 and PHAX (Figure 4A, right panel lane 2). A similar pattern was seen for the DUF mutant (Figure 4A, right panel lane 11). In contrast, the functionally inactive ZnF mutant, and constructs F2 and F5, lacking the ZnF domain, failed to co-purify MTR4 and ZC3H18 (Figure 4A, right panel lanes 10, 4 and 7 respectively), while binding to the CBC was only maintained by the ZnF mutant. Consistently, ARS2 N-terminal deletions, harboring the ZnF (constructs F4 and F6), co-purified ZC3H18 and MTR4 (Figure 4A, right panel lanes 6 and 8), which matched their ARS2-like activity in our functional assays. Finally, mutation or deletion of the CBC-binding domain (constructs M and ΔC) abolished binding to CBP80 and PHAX, as expected, but did not affect the ability to co-precipitate ZC3H18 and MTR4 (Figure 4A, lanes 11 and 12). This explains why these variants are functional when tethered to the β-globin-MS2 RNA reporter.

To assess how different domain mutations of ARS2 would affect the expression of endogenous ARS2 substrates, i.e. selected PROMPTs (26), we performed rescue experiments using tetracycline-inducible HEK293 cells, harboring stably integrated GFP-ARS2, relevant GFP-NLS-ARS2-fused fragments or GFP-ARS2 C-terminal mutants. Production of the different integrated ARS2 variants was induced subsequent to depletion of endogenous ARS2 using siRNAs to which the integrated variants were immune. A parental HEK293 cell line was employed to monitor steady-state RNA levels upon endogenous ARS2 depletion. All five tested PROMPTs, proSOD2, proZBED4, proIFNAR1, proRBM39 and proPOGZ, were stabilized upon ARS2 depletion and expression of integrated full-length ARS2 suppressed these elevated RNA levels (Figure 4B, compare 'ARS2 KD' to 'FL' columns). Similarly, expression of ARS2 F4, F6 and the DUF mutant suppressed PROMPT expression levels to a significant extent. Conversely, the ARS2 F5 variant, the CBC-binding mutants (M and ΔC) and the ZnF mutant failed to rescue normal expression of the transcripts.

Taking all of our results together, we conclude that the C-terminal part of ARS2 carries at least 2 independent protein-protein interaction domains relevant for RNA decay, engaging CBC via its extreme C-terminus while interacting with MTR4-containing complexes via the ZnF domain (Figure 4C).

DISCUSSION

Efficient quality control of nuclear RNA relies on the ability of decay factors to target a wide range of substrates. Here, ARS2 plays a central role in both productive and destructive activities (12,15,16,26). In this study, we have employed two distinct RNA reporters and the mutagenesis of ARS2 to identify two central domains of the protein linking CBC-bound transcripts to nuclear decay.

Using an MS2-tethering approach, we could show that decreased reporter RNA expression, induced by ARS2 tethering, was partly repressed by depletion of ZFC3H1 and PABPN1 from the PAXT connection, whereas depletion of the NEXT complex component ZCCHC8 had no effect. Furthermore, depletion of ZC3H18 alone was not sufficient to abolish the targeting of RNA degradation factors to the reporter. As NEXT and PAXT share the RNA helicase MTR4 in connecting these adaptor complexes to the nuclear RNA exosome (9,10), and as the ZC3H18 protein contributes to facilitating ARS2-MTR4 contacts via the NEXT- and PAXT-specific proteins ZCCHC8 and ZCF3H1, respectively (10,12,33), we do not think that the observed PAXT sensitivity of the reporter reflects a special preference of ARS2 for PAXT over NEXT interaction. Rather, we assume that efficient 3′end processing and polyadenylation of the β-globin-MS2 RNA substrate dictates its affinity for the PABPN1 component of the PAXT connection and that the effect of ARS2 tethering might well be to kinetically 'speed up' the assembly of an active PAXT-exosome, and ZC3H18 independent, interaction on the substrate. This is supposedly also the case for the 0-LpA reporter RNA, which assays the activity of ARS2 without active tethering. The abundance of this construct was also sensitive to both ARS2 and PAXT component manipulations. However, the ability of ARS2

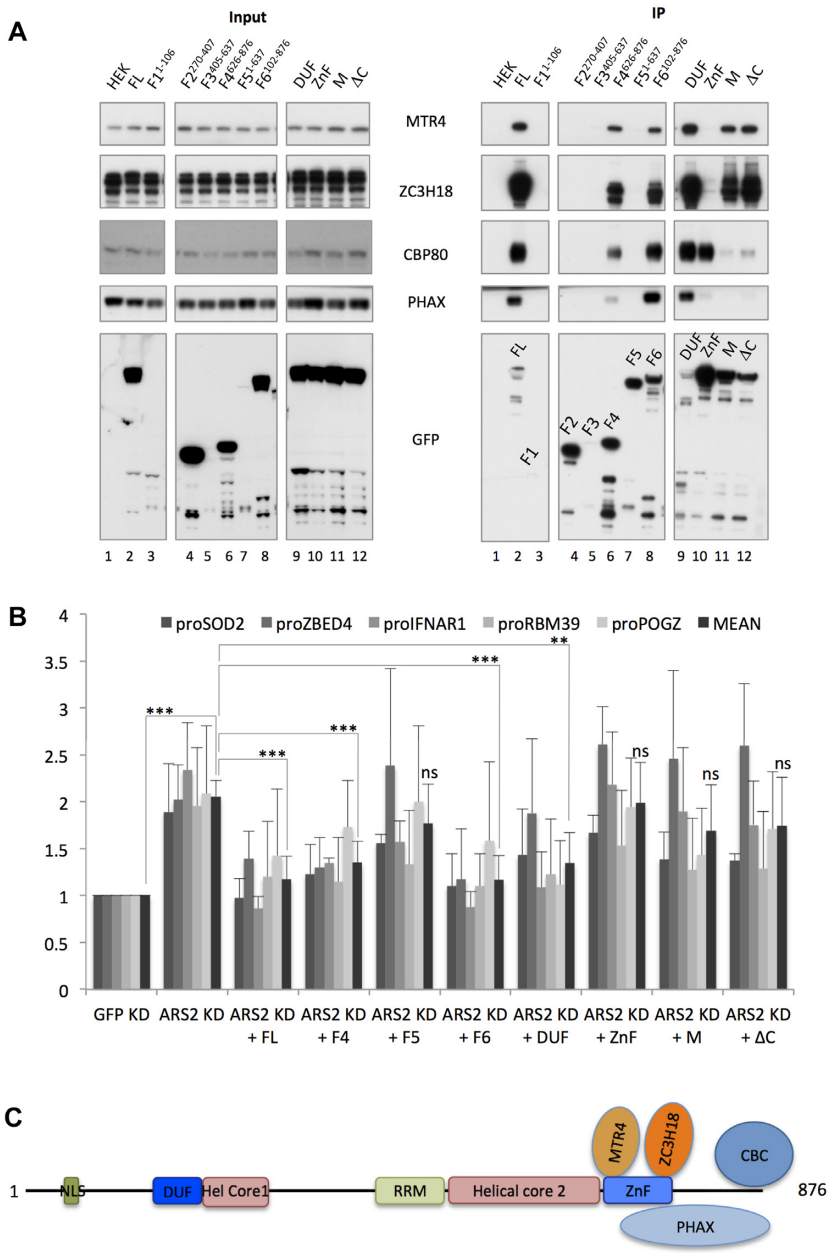

Figure 4. The ARS2 ZnF domain recruits the RNA decay machinery. (**A**) Western blotting analysis of input samples (left panel) and proteins co-IP'ed (right panel) with the indicated GFP-fused ARS2 variants expressed within their respective HEK293 inducible cell lines. Membranes were probed for MTR4, ZC3H18, CBP80, PHAX and the various GFP-ARS2 variants (α-GFP) as indicated. Additional exposure of the IP'ed GFP-fused ARS2 variants is presented in Supplementary Figure S4. FL: full length. HEK: 'Empty' cell line with no GFP-ARS2 expression. (**B**) Quantitative RT-PCR analysis of indicated RNAs purified from endogenous ARS2-depleted HEK293 cells, expressing an integrated ARS2 variant, as indicated. Data are displayed as mean ± SD of at least three biological replicates normalized to the control (eGFP) sample and GAPDH mRNA. Significance levels between the mean values of the combined total PROMPT levels in ARS2-depletion vs. 'rescued' samples were calculated using an unpaired t-test and are denoted with asterisks corresponding to the following *P*-value ranges: ** $P \leq 0.01$; *** $P \leq 0.001$. (**C**) Schematic representation of interaction domains of ARS2.

to direct the capped 0-LpA RNA to decay critically depended on the presence of the protein's CBC interaction domain. Mutation of this region of ARS2 ablated CBC binding while preserving ZC3H18/MTR4 interactions, which strongly argues that CBCA complex formation is critical for the proper processing/decay of at least some capped RNAs. This requirement for CBC binding becomes dispensable when ARS2 is directly tethered to the RNA substrate. In fact, exosomal processing of capped RNA may also completely circumvent the CBCA complex as exemplified by the NEXT-directed processing/degradation of the 3′ends of snoRNA hosting introns of protein coding mRNAs (35). Here, the RBM7 and ZCCHC8 proteins of the NEXT complex are presumably recruited to introns by the SAP49 and SAP145 splicing factors. Thus, there are alternative ways to recruit decay activity to nuclear RNAs and we suggest that

the requirement for an ARS2-CBC interaction is particularly important for shorter and less complex transcripts.

Recent structural analysis of ARS2 showed that two helical cores form a 'body' of the protein with projecting N- and C-terminal 'legs', comprising the DUF and ZnF domains respectively (28). Moreover, an NLS was shown to reside in the N-terminal extremity of the protein (30). While our domain mapping of ARS2 agreed with the presence of an N-terminal NLS, it also detected NLS capacity within the C-terminal part of ARS2 (illustrated by the F4 fragment). We suggest that this is either due to a cryptic NLS present within the last 250 amino acids or due to nuclear entry of this ARS2 fragment facilitated by its interaction with other nuclear factors in the cytoplasm, most probably the CBC. It remains to be elucidated whether this NLS activity has a functional bearing for ARS2.

Mutations of conserved residues in the DUF domain, that were previously proposed to be important for miRNA and histone RNA interactions with ARS2 (30), did neither largely affect the RNA decay mediating activity of ARS2 nor its interaction with decay factors. Consistently, fragment F5, containing the DUF but lacking the C-terminal F4 fragment, was unable to recapitulate ARS2's decay functions, both at the functional- and protein interaction-levels. It therefore remains an interesting possibility that different modules of ARS2 might explain the capability of the protein to engage in both productive and destructive RNA transactions.

The RRM domain of ARS2 binds single-stranded RNA and shares structural similarities with other RRM containing proteins (28,35). Our analysis did not allow us to associate any function to this domain. On the one hand, the F3 fragment, including the RRM domain, was too lowly expressed to assess any protein-protein interactions, and, on the other hand, the F5 variant including the RRM could not recapitulate function. The destructive role of ARS2 does therefore not seem to rely on its reported RRM.

As mentioned, the constructed C-terminal fragment of ARS2, F4, harbors the RNA decay-related properties of the protein. It encompasses the ZnF domain involved in relevant RNA- and protein-protein interactions as well as the CBC-binding domain within the extreme C-terminal 27 amino acid residues (28–30). The ZnF domain appears essential for ARS2's degradation activity by recruiting the necessary RNA decay factors. The K698A, K719A and K743A mutations within the basic patch of ZnF, forming the C-terminal foot, impaired ARS2 binding to MTR4 and ZC3H18. In previous reports, mutations of the K698A, K719A, H735A, K738A and K743A residues abolished ARS2 interactions with miRNA, RDH RNA and FLASH (30), while an ARS2 K719A K722A K734A mutant failed to interact with PHAX, NEXT and PAXT components (28). This emphasizes a functional importance of the conserved K719 residue, whose individual mutation into alanine indeed was sufficient to impair ARS2's suppressive effect on the β-globin-MS2 RNA reporter (data not shown).

All together, our results define the domains of ARS2, which serve to link nascent capped RNA to its appropriate RNA decay pathway. Further efforts will focus to address the structural underpinnings of these interactions as well as to describe their mutually exclusive relationships with RNA productive factors.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Manfred Schmid for critical comments on the manuscript and W.M. Schulze, S. Cusack and E. Bertrand for their kind provision of reagents. H.O. Pedersen, M. Vestergård, D. Riishøj and C. Scheffler are thanked for their competent technical assistance.

FUNDING

Danish Cancer Society; European Research Council Advanced Grant [339953]. Funding for open access charge: Danish Cancer Society.
Conflict of interest statement. None declared.

REFERENCES

1. Sainsbury,S., Bernecky,C. and Cramer,P. (2015) Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.*, **16**, 129–143.
2. Kuehner,J.N., Pearson,E.L. and Moore,C. (2011) Unravelling the means to an end: RNA polymerase II transcription termination. *Nat. Rev. Mol. Cell Biol.*, **12**, 283–294.
3. Schmid,M. and Jensen,T.H. (2013) Transcription-associated quality control of mRNP. *Biochim. Biophys. Acta*, **1829**, 158–168.
4. Schmid,M. and Jensen,T.H. (2018) Controlling nuclear RNA levels. *Nat. Rev. Genet.*, **19**, 518–529.
5. Kilchert,C., Wittmann,S. and Vasiljeva,L. (2016) The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.*, **17**, 227–239.
6. Chlebowski,A., Lubas,M., Jensen,T.H. and Dziembowski,A. (2013) RNA decay machines: the exosome. *Biochim. Biophys. Acta*, **1829**, 552–560.
7. Meola,N. and Jensen,T.H. (2017) Targeting the nuclear RNA exosome: poly(A) binding proteins enter the stage. *RNA Biol*, **14**, 820–826.
8. Schneider,C. and Tollervey,D. (2014) Looking into the barrel of the RNA exosome. *Nat. Struct. Mol. Biol.*, **21**, 17–18.
9. Lubas,M., Christensen,M.S., Kristiansen,M.S., Domanski,M., Falkenby,L.G., Lykke-Andersen,S., Andersen,J.S., Dziembowski,A. and Jensen,T.H. (2011) Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell*, **43**, 624–637.
10. Meola,N., Domanski,M., Karadoulama,E., Chen,Y., Gentil,C., Pultz,D., Vitting-Seerup,K., Lykke-Andersen,S., Andersen,J.S., Sandelin,A. *et al.* (2016) Identification of a nuclear exosome decay pathway for processed transcripts. *Mol. Cell*, **64**, 520–533.
11. Lubas,M., Andersen,P.R., Schein,A., Dziembowski,A., Kudla,G. and Jensen,T.H. (2015) The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. *Cell Rep.*, **10**, 178–192.
12. Andersen,P.R., Domanski,M., Kristiansen,M.S., Storvall,H., Ntini,E., Verheggen,C., Schein,A., Bunkenborg,J., Poser,I., Hallais,M. *et al.* (2013) The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat. Struct. Mol. Biol.*, **20**, 1367–1376.
13. Preker,P., Nielsen,J., Kammler,S., Lykke-Andersen,S., Christensen,M.S., Mapendano,C.K., Schierup,M.H. and Jensen,T.H. (2008) RNA exosome depletion reveals transcription upstream of active human promoters. *Science*, **322**, 1851–1854.
14. Hrossova,D., Sikorsky,T., Potesil,D., Bartosovic,M., Pasulka,J., Zdrahal,Z., Stefl,R. and Vanacova,S. (2015) RBM7 subunit of the NEXT complex binds U-rich sequences and targets 3′-end extended forms of snRNAs. *Nucleic Acids Res.*, **43**, 4236–4248.

15. Hallais,M., Pontvianne,F., Andersen,P.R., Clerici,M., Lener,D., Benbahouche,N.l.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.
16. 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.
17. Silla,T., Karadoulama,E., Mąkosa,D., Lubas,M. and Jensen,T.H. (2018) The RNA exosome adaptor ZFC3H1 functionally competes with nuclear export activity to retain target transcripts. *Cell Rep.*, **23**, 2199–2210.
18. Fan,J., Kuai,B., Wang,K., Wang,L., Wang,Y., Wu,X., Chi,B., Li,G. and Cheng,H. (2018) mRNAs are sorted for export or degradation before passing through nuclear speckles. *Nucleic Acids Res.*, **46**, 8404–8416.
19. Fan,J., Kuai,B., Wu,G., Wu,X., Chi,B., Wang,L., Wang,K., Shi,Z., Zhang,H., Chen,S. *et al.* (2017) Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *EMBO J.*, **36**, 2870–2886.
20. 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.
21. Calero,G., Wilson,K.F., Ly,T., Rios-Steiner,J.L., Clardy,J.C. and Cerione,R.A. (2002) Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. *Nat. Struct. Biol.*, **9**, 912–917.
22. Müller-McNicoll,M. and Neugebauer,K.M. (2014) Good cap/bad cap: how the cap-binding complex determines RNA fate. *Nat. Struct. Mol. Biol.*, **21**, 9–12.
23. 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.
24. Wilson,M.D., Wang,D., Wagner,R., Breyssens,H., Gertsenstein,M., Lobe,C., Lu,X., Nagy,A., Burke,R.D., Koop,B.F. *et al.* (2008) ARS2 is a conserved eukaryotic gene essential for early mammalian development. *Mol. Cell. Biol.*, **28**, 1503–1514.
25. O'Sullivan,C. and Howard,P.L. (2016) The diverse requirements of ARS2 in nuclear cap-binding complex-dependent RNA processing. *RNA Dis.*, **3**, 1–16.
26. Iasillo,C., Schmid,M., Yahia,Y., Maqbool,M.A., Descostes,N., Karadoulama,E., Bertrand,E., Andrau,J.C. and Jensen,T.H. (2017) ARS2 is a general suppressor of pervasive transcription. *Nucleic Acids Res.*, **45**, 10229–10241.
27. Gromadzka,A.M., Steckelberg,A.L., Singh,K.K., Hofmann,K. and Gehring,N.H. (2016) A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. *Nucleic Acids Res.*, **44**, 2348–2361.
28. 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.
29. Schulze,W.M. and Cusack,S. (2017) Structural basis for mutually exclusive co-transcriptional nuclear cap-binding complexes with either NELF-E or ARS2. *Nat. Commun.*, **8**, 1302.
30. O'Sullivan,C., Christie,J., Pienaar,M., Gambling,J., Nickerson,P.E., Alford,S.C., Chow,R.L. and Howard,P.L. (2015) Mutagenesis of ARS2 domains to assess possible roles in cell cycle progression and MicroRNA and replication-dependent histone mRNA biogenesis. *Mol. Cell. Biol.*, **35**, 3753–3767.
31. Andersen,P.K., Lykke-Andersen,S. and Jensen,T.H. (2012) Promoter-proximal polyadenylation sites reduce transcription activity. *Genes Dev.*, **26**, 2169–2179.
32. Damgaard,C.K., Kahns,S., Lykke-Andersen,S., Nielsen,A.L., Jensen,T.H. and Kjems,J. (2008) A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol. Cell*, **29**, 271–278.
33. Winczura,K., Schmid,M., Iasillo,C., Molloy,K.R., Harder,L.M., Andersen,J.S., LaCava,J. and Jensen,T.H. (2018) Characterizing ZC3H18, a Multi-domain protein at the interface of RNA production and destruction decisions. *Cell Rep.*, **22**, 44–58.
34. Ntini,E., Järvelin,A.I., Bornholdt,J., Chen,Y., Boyd,M., Jørgensen,M., Andersson,R., Hoof,I., Schein,A., Andersen,P.R. *et al.* (2013) Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nat. Struct. Mol. Biol.*, **20**, 923–928.
35. Falk,S., Finogenova,K., Melko,M., Benda,C., Lykke-Andersen,S., Jensen,T.H. and Conti,E. (2016) Structure of the RBM7-ZCCHC8 core of the NEXT complex reveals connections to splicing factors. *Nat. Commun.*, **7**, 13573.