*13788–13798* *Nucleic Acids Research, 2014, Vol. 42, No. 22* *Published online 20 November 2014*
*doi: 10.1093/nar/gku1159*

# Alanine repeats influence protein localization in splicing speckles and paraspeckles

**Shuo-Hsiu Chang[1,2,3], Wei-Lun Chang[2], Chia-Chen Lu[2] and Woan-Yuh Tarn[1,2,*]**

[1]Program in Molecular Medicine, National Yang-Ming University and Academia Sinica, Taipei, Taiwan, [2]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan and [3]Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei, Taiwan

Received March 11, 2014; Revised October 29, 2014; Accepted October 29, 2014

## ABSTRACT

**Mammalian splicing regulatory protein RNA-binding motif protein 4 (RBM4) has an alanine repeat-containing C-terminal domain (CAD) that confers both nuclear- and splicing speckle-targeting activities. Alanine-repeat expansion has pathological potential. Here we show that the alanine-repeat tracts influence the subnuclear targeting properties of the RBM4 CAD in cultured human cells. Notably, truncation of the alanine tracts redistributed a portion of RBM4 to paraspeckles. The alanine-deficient CAD was sufficient for paraspeckle targeting. On the other hand, alanine-repeat expansion reduced the mobility of RBM4 and impaired its splicing activity. We further took advantage of the putative coactivator activator (CoAA)-RBM4 conjoined splicing factor, CoAZ, to investigate the function of the CAD in subnuclear targeting. Transiently expressed CoAZ formed discrete nuclear foci that emerged and subsequently separated—fully or partially—from paraspeckles. Alanine-repeat expansion appeared to prevent CoAZ separation from paraspeckles, resulting in their complete colocalization. CoAZ foci were dynamic but, unlike paraspeckles, were resistant to RNase treatment. Our results indicate that the alanine-rich CAD, in conjunction with its conjoined RNA-binding domain(s), differentially influences the subnuclear localization and biogenesis of RBM4 and CoAZ.**

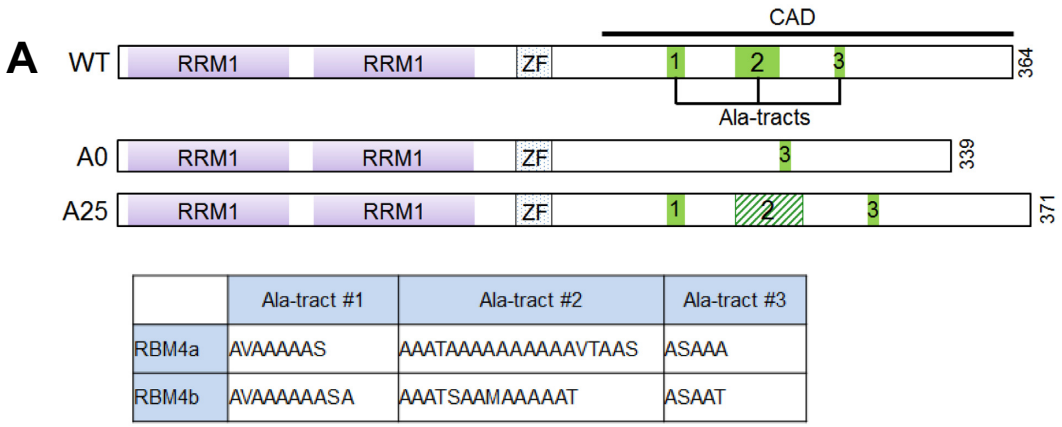
## INTRODUCTION

Eukaryotic mRNA biogenesis involves numerous RNA-binding proteins that generally function in ribonucleoprotein complexes and exhibit different subcellular localization. The human RNA-binding motif protein 4 (RBM4) has multiple roles in mRNA metabolism including alternative splicing regulation and translational control (1). RBM4 localizes to the nucleus and, in part, is concentrated in splicing speckles (2). RBM4 and its homologs have a conserved N-terminal domain that contains two RNA recognition motifs and a zinc knuckle. The C-terminal domain of RBM4 homologs is less conserved but contains one or more single amino acid repeats or dipeptide repeats. Alanine repeats are present in RBM4 of mammals and fowl but not in other species (Figure 1A) (3). Moreover, unlike mammalian RBM4, *Danio rerio* RBM4 and *Drosophila melanogaster* Lark, when transiently expressed in HeLa cells, do not localize to splicing speckles (3). We previously reported that the C-terminal alanine-rich domain (CAD) of human RBM4 can function as a speckle targeting signal (2). Therefore, we assume that the alanine-rich tracts (Ala-tracts) in the CAD may be important for the localization of mammalian RBM4 in splicing speckles.

Polyalanine sequences exist in ∼1.5% of human proteins, and alanine expansions have been implicated in neurodegenerative diseases and neurological malfunction (4,5). Among these, oculopharyngeal muscular dystrophy (OPMD)-associated polyalanine expansion in the nuclear polyadenylate-binding protein (PABPN1) is particularly noteworthy (5). PABPN1 stimulates polyadenylation processivity and suppresses alternative polyadenylation at proximal/weak sites (6). A mutant PABPN1 with an expanded polyalanine tract induces proximal cleavage and polyadenylation and thus shortens the 3′ untranslated region of target transcripts, which may relieve miRNA-mediated repression, similar to what is seen in some cancer cells (6). Polyalanine expansions may lead to protein misfolding and subsequent aggregation (7). Moreover, PABPN1 mutants may sequester mRNAs in intranuclear domains (8). The Ala-tracts of RBM4 homologs vary in length (3), and thus the Ala-tracts may affect RBM4 function and subcellular localization.

In this study, we found that truncation of the Ala-tracts shifted localization of RBM4 to paraspeckles, which are mammalian-specific nuclear bodies that often lie adjacent to splicing speckles (9). Paraspeckles contain several multifunctional proteins, including PSF (PTB-associated splic-

*To whom correspondence should be addressed. Tel: +886 2 2652 3052; Fax: +886 2 2782 9142; Email: wtarn@ibms.sinica.edu.tw
Present address: Wei-Lun Chang, Institute of Molecular and Genomic Medicine, National Health Research Institutes, Zhunan, Taiwan.




| | Ala-tract #1 | Ala-tract #2 | Ala-tract #3 |
|---|---|---|---|
| RBM4a | AVAAAAAS | AAATAAAAAAAAAAVTAAS | ASAAA |
| RBM4b | AVAAAAAASA | AAATSAAMAAAAAT | ASAAT |

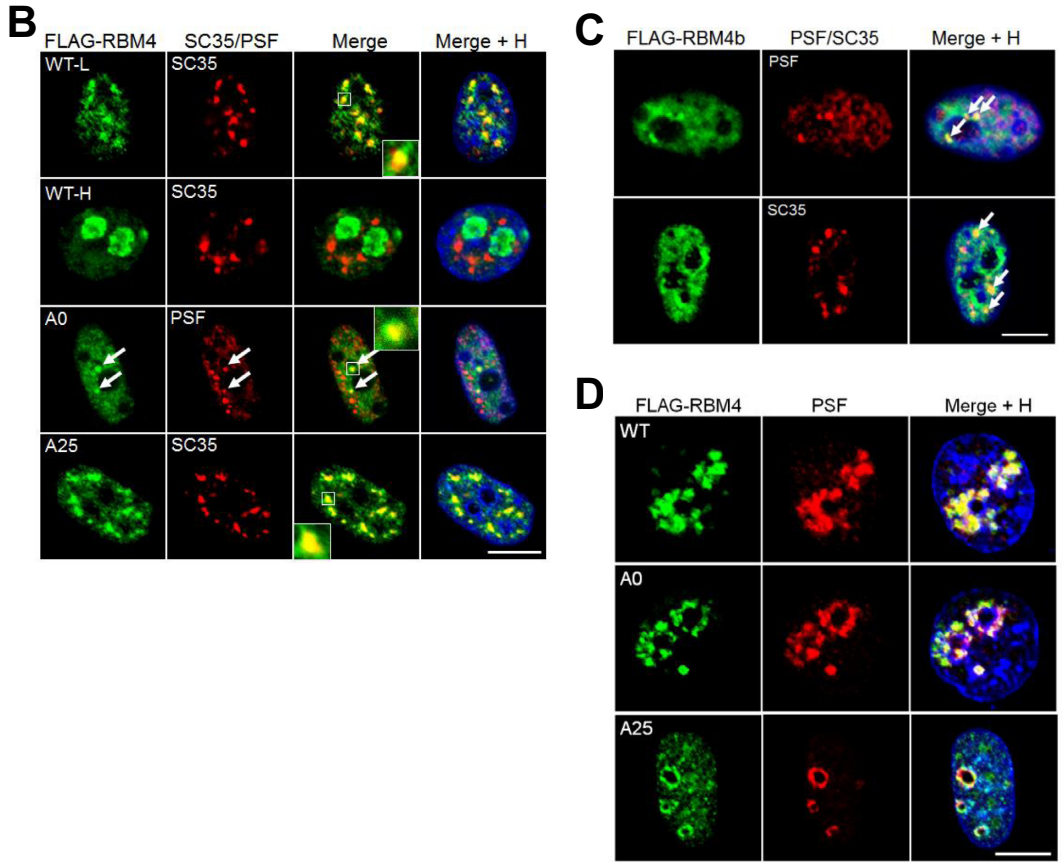



**Figure 1.** The length of the Ala-tracts affects subnuclear localization of RBM4. (**A**) Schematic diagram of RBM4 (WT) and the Ala-tract mutants (RBM4-A0 and RBM4-A25). RRM: RNA recognition motif. ZF: zinc finger. The sequences of the three Ala-tracts of human RBM4 and RBM4b are listed in the table. (**B**) FLAG-RBM4, RBM4-A0 and RBM4-A25 were transiently expressed in HeLa cells. Immunofluorescence was performed using anti-FLAG and anti-SC35 or anti-PSF. Representative images in the upper two panels show the localization patterns of FLAG-RBM4 at low (WT-L) and high (WT-H) expression levels, respectively. Merged images show RBM4 or its Ala-tract mutants (green), and PSF or SC35 (red) superimposed on Hoechst nuclear counterstaining (blue; +H). Insets show 3× magnification of colocalization (yellow) of FLAG-RBM4 and SC35 or PSF. (**C**) FLAG-RBM4b was transiently expressed in HeLa cells, and immunofluorescence with Hoechst staining was performed as in panel (B). (**D**) Transiently transfected HeLa cells as in panel (B) were treated with 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) for 4 h before fixation and immunofluorescence. Scale bar represents 10 μm.

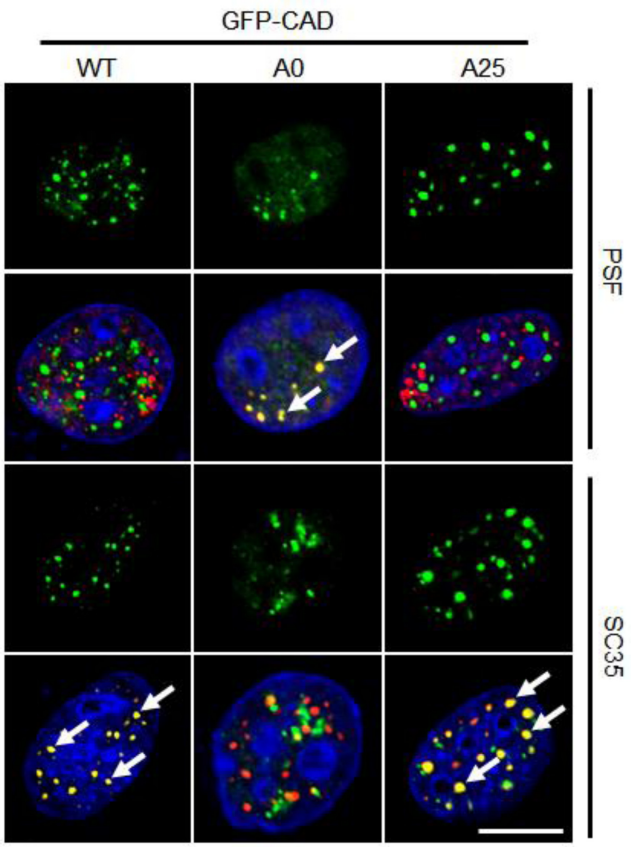



**Figure 2.** The length of the Ala-tracts determines the subnuclear targeting properties of the RBM4 CAD. GFP-CAD (WT), -CAD-A0 and -CAD-A25 were transiently expressed in HeLa cells. Representative images show the GFP-CAD and GFP-Ala-tract mutants (green) alone or merged with PSF or SC35 (red) immunofluorescence and Hoechst (blue) staining. Arrows indicate colocalization of the GFP constructs with PSF or SC35 (yellow). Scale bar represents 10 μm.

ing factor), p54nrb, PSP1 and RBM14/ CoAA/PSP2 (hereafter referred to CoAA) (9,10). The long noncoding RNA *NEAT1* assembles these protein components and is critical for paraspeckle formation (11–13). *NEAT1* is absent in embryonic stem cells, but is expressed in differentiated cells, suggesting that paraspeckles function in cell differentiation (13,14). Moreover, paraspeckles harbor extensively edited RNAs and may modify and release these sequestered transcripts during cell stress, suggesting a role for paraspeckles in the stress response (15).

The CoAA gene is located immediately upstream of and in the same transcriptional orientation as the RBM4a gene in the human and mouse genomes (see the text). The CoAA-RBM4 conjoined (CoAZ) transcripts, including the noncoding CoAZ RNA and the CoAZ mRNA, have been detected in various human tissues and cell lines (16). These transcripts may be generated by *trans*-splicing of the CoAA and RBM4a transcripts or by transcriptional read-through followed by *cis*-splicing (16,17). CoAZ contains one and a half RNA recognition motifs of CoAA and the C-terminal portion of RBM4a, including the zinc finger and the CAD. CoAZ stimulates transcription in a hormone-dependent manner and also regulates alternative splicing (16). RBM4 expression is initially increased during neuronal cell differentiation and appears to undergo a switch to CoAZ at later stages of differentiation (16). Therefore, we were interested in exploring the functional relationship between RBM4 and CoAZ in cells.

Here we evaluated whether the alanine-rich motifs of RBM4 are critical for its subcellular localization and function. We also took advantage of CoAZ to explore how the CAD functions in this heterologous RNA-binding protein.

## MATERIALS AND METHODS

### Plasmids

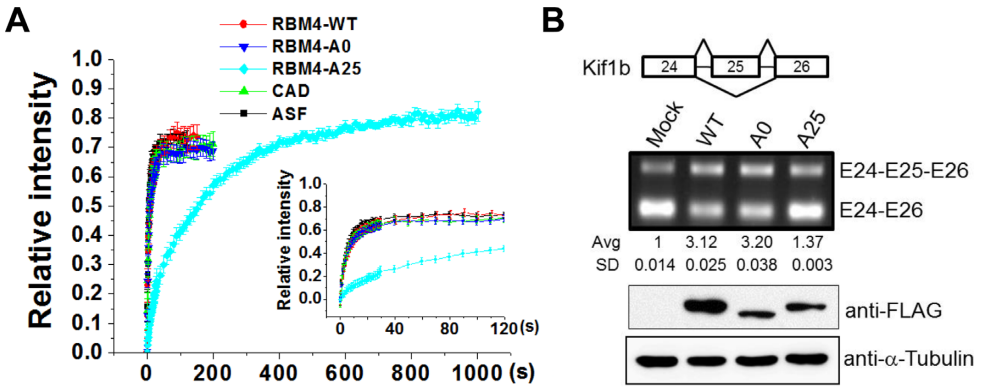
The plasmids pcDNA-FLAG-RBM4(a) and pEGFP-CAD (encoding residues 196–364 of RBM4a) have been described (2). Using a polymerase chain reaction (PCR)-based gene deletion strategy, Ala-tract #1 (residues 233–239) and #2 (residues 282–299) of pcDNA-FLAG-RBM4(a) were removed to generate pFLAG-RBM4-A0. The pFLAG-RBM4-A25 plasmid was constructed by substituting three nonalanine residues of Ala-tract #2 with alanines and inserting an additional seven alanines. The RBM4b-coding region was cloned into pcDNA3.1 in frame with the FLAG sequence. The coding regions of CAD-A0, CAD-A25, PSF and SRSF1/ASF were each inserted into vector pEGFP-C1 (Invitrogen) to generate the respective C-terminal green fluorescent protein (GFP) fusions. The pcDNA-FLAG-CoAZ was a kind gift of Lan Ko (Medical College of Georgia). The red fluorescent protein (RFP)-coding sequence was inserted into pcDNA-FLAG-RBM4(a) and pcDNA-FLAG-CoAZ to generate vectors RFP-RBM4(a) and RFP-CoAZ, respectively. The splicing reporter plasmids have been described (3).

### Cell culture and transfection

HeLa and HEK293 cells were cultured in Dulbecco's Modified Eagle's medium (Invitrogen) containing 10% fetal bovine serum, 100 U/ml penicillin, 100 μg/ml streptomycin and 2 mM L-glutamine at 37°C and 5% $CO_2$. Transfections were performed using Lipofectamine 2000 (Invitrogen). HeLa cells (~8 × $10^4$) grown in 24-well culture plates were transfected with 0.6 μg of each expression plasmid for 24 h or cotransfected with 20 pmol of siRNA (Stealth siRNA, Invitrogen) targeting luciferase (siLuc) (18), PSF (sense ACUUGCCCAGAAGAAUCCAAUGUAU and antisense AUACAUUGGAUUCUUCUGGGCAAGU) or hnRNP K (sense UCAUCAGAGUCUAGCAGGAGGAAUU and antisense AAUUCCUCCUGCUAGACUCUGAUGA) and 0.5 μg of each protein expression vector for 48 h. To inhibit RNA polymerase II–mediated transcription, HeLa cells were treated with 100 μM 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) for 4 h. For RNase treatment, cells were incubated with 200 μg/ml RNase A (Sigma) in phosphate-buffered saline (PBS) at 37°C for 30 min.

### Antibodies

Mouse monoclonal antibodies used in the study included those against PSF (for immunofluorescence, P2860, Sigma-Aldrich; and for immunoblotting, sc-101137, Santa Cruz),




| | $t_{1/2}$(s) | mobile fraction(%) |
|---|---|---|
| GFP-RBM4 | 5.28 ±0.86 | 70.99±4.36 |
| GFP-RBM4-A0 | 4.00±0.65 | 67.77±2.80 |
| GFP-RBM4-A25 | 110.04±13.17 | 80.4±4.39 |
| GFP-CAD | 4.68±0.81 | 66.38±4.10 |
| GFP-ASF | 4.16±0.33 | 71.68±1.35 |

**Figure 3.** Polyalanine expansion reduces nuclear mobility of RBM4 and impairs its splicing regulatory function. (**A**) HeLa cells were transiently transfected with GFP-fusion proteins as indicated. Selected GFP foci were subjected to fluorescence recovery after photobleaching (FRAP) analysis. The graph shows the fluorescence recovery of each GFP-fusion up to 1000 s after photobleaching; the average and standard deviation (SD) were obtained from three independent experiments. The inset shows the recovery curve of the GFP-fusion proteins within the initial 120 s. The table shows the half-time of fluorescence recovery ($t_{1/2}$) and mobile fraction (%) of each GFP-fusion. (**B**) The expression vector encoding FLAG-tagged RBM4 or Ala-tract mutants was cotransfected with the Kif1b minigene into HEK293 cells. The splicing products were amplified by reverse transcriptase-PCR (RT-PCR), and the exon inclusion efficiencies relative to the mock (empty vector) are indicated below the gel. The averages (Avg) and SD were obtained from three independent experiments. Immunoblotting was performed using anti-FLAG and anti-α-Tubulin.

SC35 (S4045, Sigma-Aldrich), FLAG tag (F1804, Sigma-Aldrich) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH, NB300-221, NOVUS). Rabbit polyclonal antibodies included those against CoAA (A300-331A, Bethyl), FLAG tag (F7425, Sigma-Aldrich) and RBM4 (2). The anti-hnRNP K rabbit monoclonal antibody was from Abcam (ab52600). Alexa Fluor 488- and 568-conjugated secondary antibodies were from Invitrogen.

### Indirect immunofluorescence

HeLa cells cultured on coverslips in 24-well culture plates were transfected with an expression vector or siRNA as described above. Cells were fixed with 3% formaldehyde in PBS for 15 min and subsequently permeabilized with 0.5% Triton X-100 for 10 min. Cells were blocked with 3% nonfat dried milk in PBS for 30 min and then sequentially incubated with primary antibodies in the blocking solution for 2 h and Alexa Fluor 488- or 568-conjugated secondary antibody for 1 h. Nuclei (DNA) were counterstained with Hoechst 33258 (Sigma-Aldrich). Cells were mounted to slides with the ProLong Gold antifade reagent (Invitrogen). Images were acquired on a laser scanning confocal microscope (LSM 700; Carl Zeiss, Jena, Germany) with a Plan-Apochromat 63/1.40 oil DIC M27 objective and processed with ZEN 2009 software (Carl Zeiss). For Figure 5B and Supplemental Figure S5D, the master gain values of laser used were 650 (low-power) and 750 (high-power).

### Fluorescence in situ hybridization

The *NEAT1* noncoding RNA was detected by fluorescence in situ hybridization (FISH) as described (19). In brief, HeLa cells cultured on coverslips in 24-well culture plates were mock transfected or transfected with siRNA. Cells were fixed with 3% formaldehyde in PBS for 15 min at room temperature. After wash with PBS, cells were permeabilized with 0.5% Triton X-100 in PBS for 10 min at 4°C. Cells were then briefly washed with PBS three times and 2× saline-sodium citrate (SSC) buffer once prior to hybridization. The cDNA probes were digoxigenin (DIG)-labeled by using the DIG-Nick Translation kit (Roche). Hybridization was performed in 20 μl of 2× SSC containing 10% dextran sulphate, 100 ng of each probe and 20 μg of yeast tRNA overnight at 37°C. After wash with 50% formamide-containing 2× SSC followed by 1× SSC, cells were incubated with anti-DIG-fluorescein antibody for 1 h. After FISH, cells were fixed with 3% formaldehyde in PBS for 5 min and then subjected to indirect immunofluorescence as described above.

### *NEAT1* knockdown and expression of FLAG-CoAZ

HeLa cells were transiently transfected with control (siLuciferase) or *NEAT1*-targeting siRNA (siNEAT; Integrated DNA Technologies) using Lipofectamine 2000 (Invitrogen) for 24 h. After first transfection, cells were transfected with siRNA for the second time as well as the FLAG-CoAZ expression vector for another 24 h. For each transfection, 100 nM siRNA was used. The sequences of the siNEAT1 duplex are 5'-/5Phos/rGrUrGrArGrArArGrUr

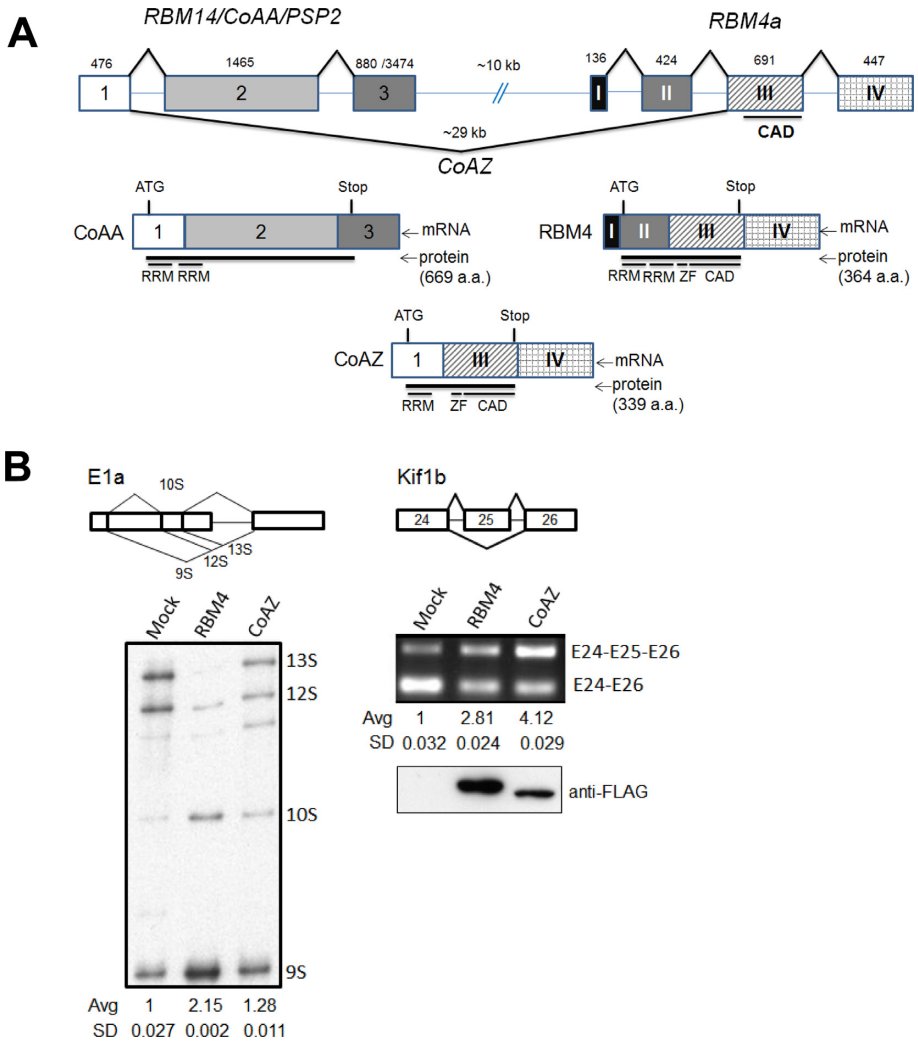



**Figure 4.** CoAZ acts as a splicing regulator with activity distinct from that of RBM4. (**A**) Schematic diagram of the exon composition of RBM4a, CoAA and CoAZ. (**B**) The mock, FLAG-tagged RBM4 or CoAZ expression vectors were cotransfected with the E1a splicing reporter into HeLa cells or with the Kif1b into HEK293 cells. E1a splicing products were amplified by RT-PCR, fractionated on polyacrylamide gels and detected with $^{32}$P-labeled-specific probes in immunoblots (2). Kif1b splicing was analyzed as in Figure 3B. The relative efficiencies of distal 5′ splice site utilization (9S/total) in E1a and exon inclusion of Kif1b are the averages (Avg) and SD of three independent experiments.

UrGrCrUrUrArGrArArArCrUrUrUCC-3' (sense) and 5'-rGrGrA rArArGrUrUrUrCrUrArArGrCrArArCrUrUrC rUrCrArCrUrU-3' (antisense). Cells were then subjected to immunofluorescence and FISH or lysates were harvested for immunoblotting and reverse transcription-PCR (RT-PCR).

### Live-cell confocal imaging

HeLa cells cultured on coverslips were transiently transfected as described above with the GFP and/or RFP fusion-protein vectors for 6 h. After replacing the Lipofectamine 2000 transfection medium (see above) with complete medium, time-lapse cell images were acquired using a laser scanning confocal microscope (LSM 780; Carl Zeiss) at 37°C and 5% $CO_2$. Images were taken every 15 min and processed with ZEN 2012 software.

### Fluorescence recovery after photobleaching

HeLa cells grown to 70% confluence on glass coverslips were transfected with the GFP-fusion expression vectors encoding SRSF1/ASF, CAD, CoAZ or RBM4. Fluorescence recovery after photobleaching (FRAP) experiments were performed 18 h after transfection on the LSM 780 confocal microscope with a 100× oil immersion objective (Plan-Apochromat 100×/1.40 oil DIC M27) at 37 or 25°C. A circular region 2 μm in diameter was photobleached by the 488-nm line of a 25-mW argon laser at full power for 100 iterations. Images were taken before and after photobleaching with low laser power for monitoring the fluorescence recovery. The immobile population and half-time of fluorescence recovery ($t_{1/2}$) were analyzed with ZEN 2011 software.

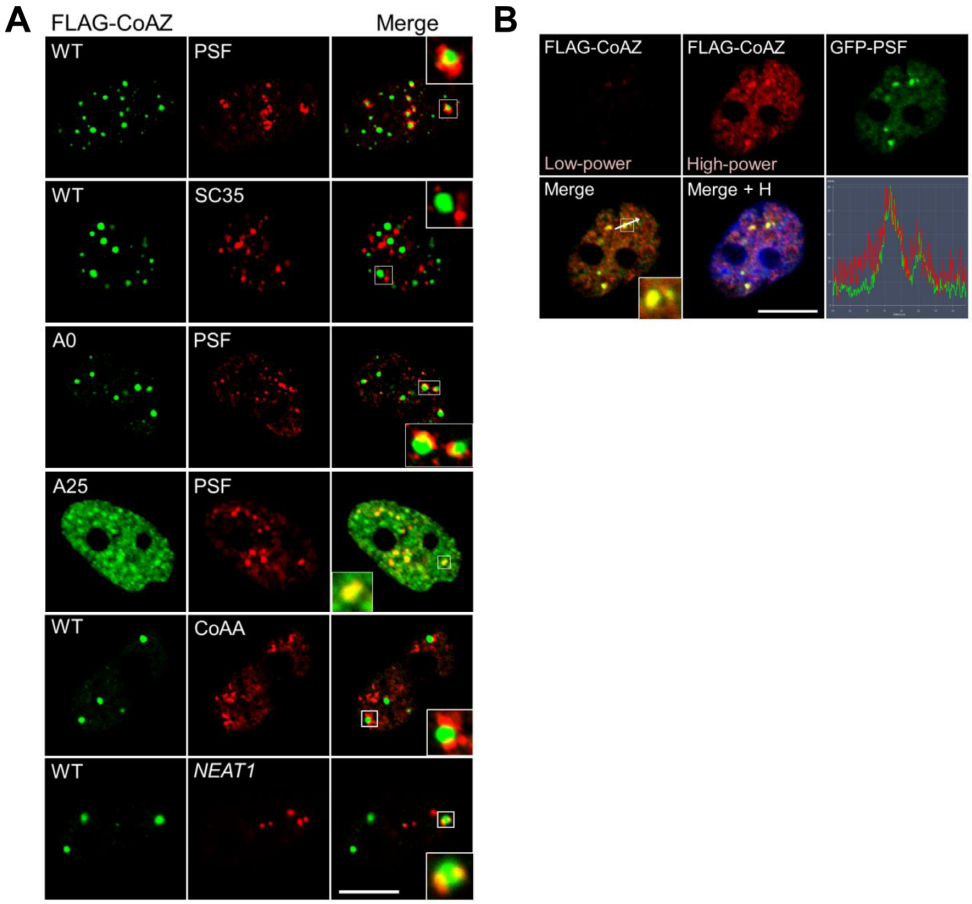



**Figure 5.** CoAZ is localized in close proximity to paraspeckles. (**A**) HeLa cells were transiently transfected with the vector expressing FLAG-tagged CoAZ, CoAZ-A0 or CoAZ-A25. Immunofluorescence was performed using anti-FLAG (green) and anti-PSF, anti-SC35 or anti-CoAA (red). Fluorescence in situ hybridization (FISH) was performed with the *NEAT1* probe (red). Insets show higher magnification (3×) of selected FLAG-CoAZ foci. Merged images are shown in the right-most column. (**B**) FLAG-CoAZ (red) and GFP-PSF (green) were transiently expressed in HeLa cells. FLAG-CoAZ was detected by indirect immunofluorescence using anti-FLAG at low and high laser power. The merge of the high-power FLAG-CoAZ (red) and GFP-PSF (green) is shown in the lower-left panel, and a further merge with the Hoechst (blue) image is shown in the lower-middle panel. In merge, the line (with an arrowhead) indicates the fluorescence intensity scanning direction. The inset shows a higher magnification (3×) of the selected CoAZ/PSF foci.

### *In vivo* splicing assay

Cells were cultured and transfected using Lipofectamine 2000 as described (2). Briefly, 2 μg of the effector expression vector and 0.5 μg of the splicing reporter plasmid were co-transfected into $7 \times 10^5$ HEK 293 or $2.5 \times 10^5$ HeLa cells for 24 h. The splicing reporters used were E1a and Kif1b (exons 24–25–26). The splicing products were analyzed by RT-PCR as described (3). PCR bands were quantified by densitometry using ImageJ software (National Institutes of Health, USA).

## RESULTS

### The alanine-rich sequences of RBM4 modulate its subnuclear localization

We evaluated whether and how the Ala-tracts of human RBM4a (hereafter referred to as RBM4) influence its subnuclear localization and function in cultured human cells. As previously observed (2), FLAG-tagged RBM4 expressed at moderate levels was distributed similarly to endogenous RBM4 (Figure 1B, WT-L), whereas highly expressed FLAG-RBM4 was mainly concentrated in nucleoli (WT-H, Figure 1B). Next, we deleted the two longest Ala-tracts (#1 and #2, Figure 1A) of RBM4. The resulting truncated protein, FLAG-RBM4-A0, was primarily localized in the nucleoplasm and also concentrated in some small nuclear spots (A0, Figure 1B). Using antibodies against various nuclear body markers, we found that a significant portion of these RBM4-A0 spots contained PSF (Figure 1B) or CoAA but not the splicing factor SC35/SRSF2 (Supplemental Figure S1A). This observation was reminiscent of recent findings of RBM4b in paraspeckles (10,20). Therefore, we examined the localization of transiently expressed FLAG-RBM4b. Indirect immunofluorescence showed that FLAG-RBM4b was more diffuse throughout the nucleoplasm than FLAG-RBM4(a), but still formed foci and that some of the RBM4b foci overlapped with PSF and some overlapped with SC35 (Figure 1C, arrows). We deduced that the length of the Ala-tracts accounts for such a difference. To evaluate this possibility, we replaced the second Ala-tract of RBM4(a) with that of RBM4b or with the proline-rich tracts of zebrafish RBM4.2 (3). Proline-rich zebrafish RBM4.2 could modulate alternative splicing,

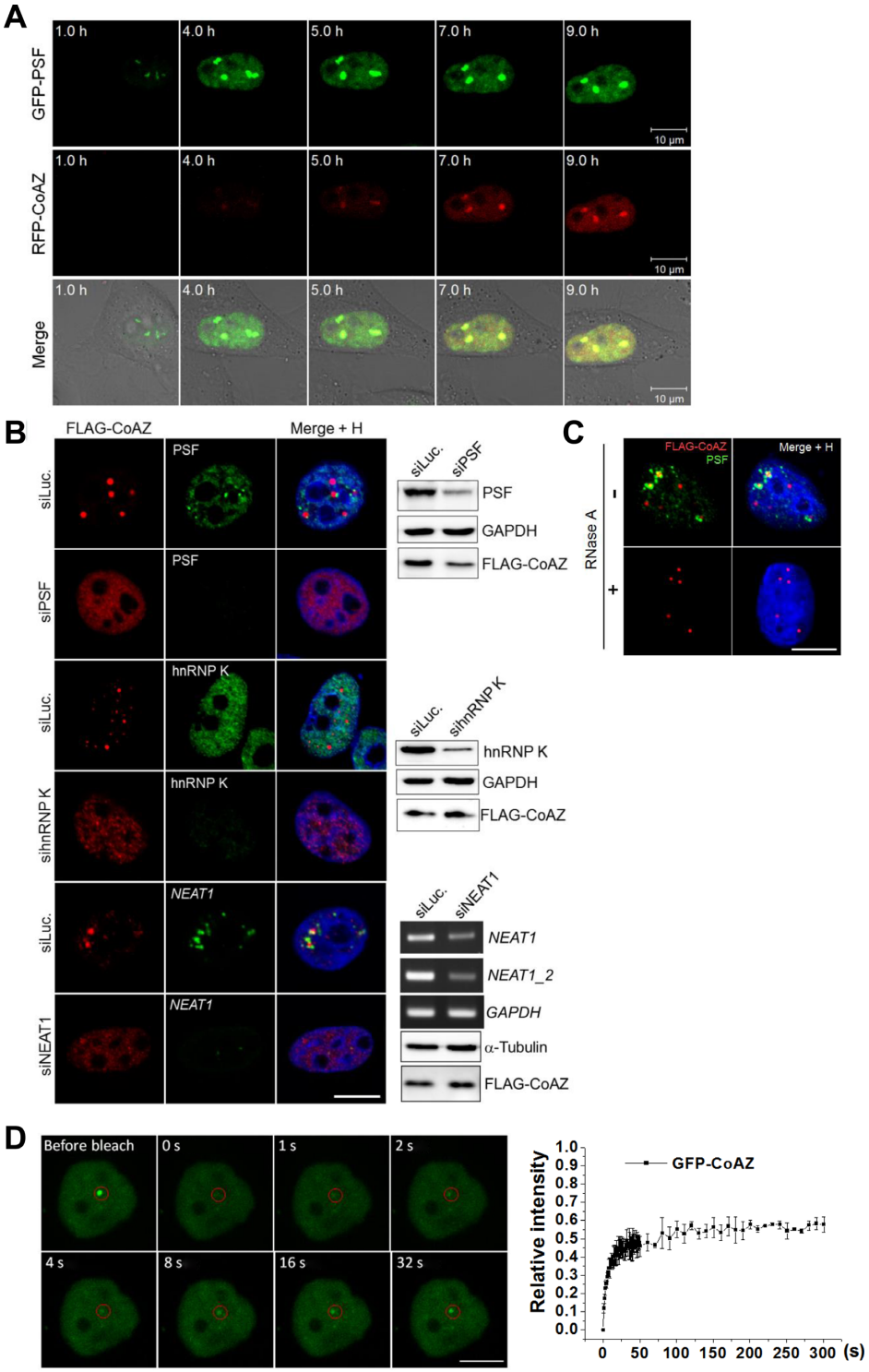



**Figure 6.** Formation of CoAZ foci depends on paraspeckles but not RNA. (**A**) HeLa cells were transiently transfected with the RFP-CoAZ and GFP-PSF expression vectors. Images were recorded at the indicated times after medium change (6 h post-transfection) using time-lapse fluorescence microscopy. Merged images are shown in the bottom row. (**B**) HeLa cells were transiently cotransfected with siRNA targeting PSF (siPSF), hnRNP K (sihnRNP K), *NEAT1* (siNEAT1) or luciferase (siLuc, control) and the FLAG-CoAZ expression vector. Immunoblotting analysis was performed using antibodies against PSF, hnRNP K, FLAG, α-Tubulin and GAPDH. RT-PCR was performed to detect *NEAT1*, *NEAT1_2* and *GAPDH*. Indirect immunofluorescence was performed using antibodies against FLAG, PSF and hnRNP K. FISH was performed with the *NEAT1* probe. Merged images including Hoechst staining are shown at right. (**C**) HeLa cells were transfected with the FLAG-CoAZ expression vector for 24 h and then mock treated or treated with RNase for 30 min before cell fixation. Indirect immunofluorescence using anti-FLAG and anti-PSF was performed as in panel (B). Merged images including Hoechst (blue) staining are shown. (**D**) FRAP was performed in HeLa cells transiently expressing GFP-CoAZ at 25°C. Images show a representative GFP-CoAZ focus (circled in red) before bleaching and at the indicated times post-bleaching. The graph shows the fluorescence recovery of GFP-CoAZ as in Figure 3A. Scale bar represents 10 μm.

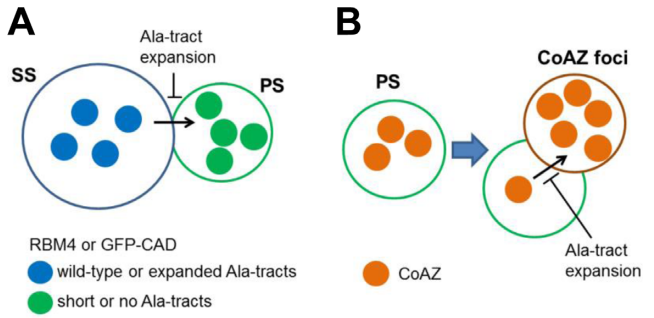



**Figure 7.** Model for subnuclear localization and biogenesis of RBM4 and CoAZ. (**A**) The CAD (GFP fusion) of RBM4 directs it primarily to splicing speckles (SS), and this localization is not affected by Ala-tract expansion. Truncation of the Ala-tracts appears to direct the RBM4 or GFP CAD-fusion to paraspeckles (PS), likely when SS and PS are transiently located close to each other. (**B**) CoAZ appears to be initially deposited in PS, and a portion of CoAZ remains localized close to PS, and subsequently separated from PS to form distinct foci. Ala-tract truncation does not affect this localization pattern, but Ala-tract expansion appears to impair separation of CoAZ foci from PS so that some CoAZ-A25 foci continue to completely overlap with PS.

but it was not located in splicing speckles (3). The resulting chimeric RBM4(a) proteins, RBM4-Ala2b and RBM4-CP, exhibited a nuclear distribution pattern more similar to that of RBM4-A0 and RBM4b (Supplemental Figure S1B). This observation emphasized that truncation of the Ala-tract(s) led RBM4 to paraspeckles.

Furthermore, we replaced the second Ala-tract of RBM4(a) with a stretch of 25 alanine residues, the resulting FLAG-RBM4-A25 exhibiting a clear speckled pattern that overlapped with the SC35 signal (A25, Figure 1B). This observation indicated that the alanine-rich tracts of RBM4 strongly influence its subnuclear localization.

RBM4 is translocated to perinucleolar regions after inhibition of RNA polymerase II–mediated transcription (20,21). Therefore, we examined whether the length of the Ala-tracts influences this property of RBM4. When RNA synthesis was blocked with DRB, FLAG-RBM4 and the two Ala-tract mutants, A0 and A25, exhibited perinucleolar localization except that the shift of FLAG-RBM4-A25 from the nucleoplasm to nucleoli was relatively less complete (Figure 1D and Supplemental Figure S1C). Although our initial test showed that Ala-tract expansion did not specifically enhance the interaction of RBM4 with SC35 or SRSF1/ASF (Supplemental Figure S2), the possibility that such an expansion increases the affinity of RBM4 to other splicing speckle components, or reduces protein dynamics of RBM4 still remains.

### The truncated-Ala-tract CAD is a paraspeckle targeting signal

We previously showed that the CAD of RBM4 functions as a splicing speckle targeting signal (2). In this study, we further examined whether the Ala-tracts affect this subnuclear localization property of the CAD. We generated GFP fusions of CAD-A0 and -A25. Fluorescence microscopy showed that GFP-CAD-A25, similar to GFP-CAD, essentially (>90%) localized to splicing speckles, indicating that alanine repeat expansion did not alter the subnuclear localization of the CAD (Figure 2). GFP-CAD-A0 also exhibited a speckled pattern in the nucleus, but surprisingly these foci colocalized with PSF in paraspeckles (Figure 2). This result was reminiscent of the localization of FLAG-RBM4-A0 and FLAG-RBM4b (Figure 1B and C) and suggested that the Ala-tract-deleted CAD acquired the ability to localize to paraspeckles.

We next examined whether the subcellular localization of the GFP-CAD fusions is altered in response to transcription inhibition. After DRB treatment, both GFP-CAD and CAD-A0 could translocate to perinucleolar regions (Supplemental Figure S3), indicating that the RNA-binding domains are not essential for transcription inhibition-induced subnuclear translocation. However, GFP-CAD-A25 was retained in nuclear speckles after DRB treatment (Supplemental Figure S3). This result coincided with the partial retention of FLAG-RBM4-A25 in the nucleoplasm and speckles after transcription inhibition (Figure 1D), suggesting that alanine repeat expansion may induce self or protein–protein interactions of RBM4 that prevent translocation between different subnuclear domains.

### Polyalanine expansion reduces the intranuclear mobility of RBM4 and impairs its splicing regulatory function

We next examined whether truncation or expansion of the Ala-tracts influences RBM4 dynamics. To perform a FRAP experiment, we fused GFP to the full-length/wild-type (WT) or Ala-tract-truncated or -expanded RBM4. GFP-RBM4 and the A25 mutant formed nuclear foci that colocalized with SC35 (Supplemental Figure S3). As observed with FLAG-RBM4-A0, the GFP-RBM4-A0 signal, although diffused in the nucleoplasm, was still detected in the foci that overlapped strongly with PSF and weakly with SC35 (Supplemental Figure S4). Individual foci of the various GFP-RBM4 fusions were selected for photobleaching, and the rate of fluorescence recovery was determined. GFP-RBM4 foci reformed at a rate similar to that of GFP-tagged alternative splicing factor (SRSF1/ASF), a splicing speckle-localized factor (Figure 3A, graph). GFP-RBM4-A0 and WT had similar recovery rates (Figure 3A). In contrast, the motility of GFP-RBM4-A25 was markedly reduced, although the mobile fraction increased slightly (Figure 3A). Therefore, truncation of the Ala-tracts did not significantly affect the rate of foci formation, whereas polyalanine expansion reduced RBM4 mobility, perhaps owing to protein aggregation or increased interactions with other factors. Finally, we observed that GFP-CAD formed similar foci to GFP-RBM4, suggesting that the RNA-binding domain is not a critical determinant of protein mobility (Figure 3A).

We next examined whether the Ala-tracts influence the splicing regulatory function of RBM4. Using the human kinesin family member 1B (Kif1b) minigene splicing reporter (3), we observed that RBM4-A0 functioned similarly to the WT RBM4 in promoting exon 25 inclusion of Kif1b, whereas RBM4-A25 was unable to affect this splicing event (Figure 3B). Therefore, Ala-tract expansion, but not truncation, impaired the splicing regulatory activity of RBM4.

### The CAD-containing splicing regulator CoAZ is a paraspeckle factor

Our intriguing observations that Ala-tract-truncated RBM4 partially localized in paraspeckles and that the CAD lacking the Ala-tract functions as a paraspeckle targeting signal prompted us to characterize the CAD in more detail. We took advantage of CoAZ (Figure 4A) to explore how the CAD of RBM4 influences the function and cellular localization of the heterologous RNA-binding protein.

CoAZ functions in an opposing manner to RBM4 in splicing of the Tau1 transcript (16). Here we compared the alternative splicing activity of CoAZ and RBM4 with two splicing reporters. FLAG-RBM4 or FLAG-CoAZ was coexpressed with the adenovirus E1a or Kif1b reporter in HeLa or HEK293 cells. Splicing analyses revealed that CoAZ, unlike RBM4, only minimally promoted distal 5′ splice site usage in E1a, whereas it was more potent than RBM4 in promoting exon 25 inclusion in Kif1b (Figure 4B). Therefore, as a splicing regulator, CoAZ may have different mechanisms of action from RBM4.

Next, to examine the subcellular localization of CoAZ, FLAG-CoAZ was transiently expressed in HeLa cells. Immunofluorescence showed that FLAG-CoAZ formed nuclear foci (Figure 5A). Further double immunofluorescence and in situ hybridization revealed that some of the CoAZ foci were located in close proximity to paraspeckles, as represented by either PSF or CoAA protein or the *NEAT1* RNA (Figure 5A, WT: PSF/CoAA/*NEAT1*), but none colocalized with splicing speckles (Figure 5A, WT: SC35, and Supplemental Figure S5A for costaining of FLAG-CoAZ and SC35 and fluorescence imaging of GFP-PSF). Notably, a considerable fraction of the paraspeckle signal was seen as crescent-shaped patches or smaller speckles at the periphery of FLAG-CoAZ foci (Figure 5A, WT: PSF/CoAA/*NEAT1*). Nevertheless, overexpression of FLAG-CoAZ did not affect the expression level of PSF, CoAA or *NEAT1* (Supplemental Figure S5B). Similar to WT CoAZ, the Ala-tract-truncated CoAZ (FLAG-CoAZ-A0), formed discrete foci, some of which were surrounded by PSF patches (Figure 5A, A0), but none overlapped with splicing speckles (Supplemental Figure S5C). Surprisingly, FLAG-CoAZ-A25 was primarily distributed in the nucleoplasm and formed foci that overlapped with paraspeckles (Figure 5A, A25) but not with splicing speckles (Supplemental Figure S5C), suggesting an intimate relationship between CoAZ biogenesis and paraspeckles.

We then overexpressed exogenous PSF to further examine its localization relationship with CoAZ. When FLAG-CoAZ and GFP-PSF were coexpressed in HeLa cells, we observed FLAG-CoAZ localization adjacent to GFP-PSF in ∼70% of cells (Supplemental Figure S5D). Interestingly, in the remaining cells, FLAG-CoAZ levels were relatively much lower and were detectable only by high-power laser excitation (Figure 5B). Under these conditions, in addition to the signal observed in the nucleoplasm, FLAG-CoAZ still formed some foci that overlapped with GFP-PSF foci (Figure 5B). A similar observation was detected using polyclonal FLAG-CoAZ expressing HeLa cells, in which the expression level of ectopically expressed CoAZ was largely reduced (Supplemental Figure S5E) This result confirmed the relationship between CoAZ and the paraspeckle component PSF and indicated that when CoAZ is expressed at lower levels, presumably near the levels of the endogenous protein (16), it appears to almost entirely colocalize with PSF.

### Formation of CoAZ foci depends on paraspeckles

To explore the mechanism of CoAZ foci formation, we examined the expression of RFP-CoAZ and GFP-PSF in HeLa cells using time-lapse microscopy. The result showed that the newly synthesized RFP-CoAZ accumulated at sites where GFP-PSF was located (Figure 6A). To exclude the possibility that early-expressed GFP-PSF sequestered RFP-CoAZ, we switched the fluorescent protein tag between these two proteins. GFP-CoAZ was still colocalized with late-expressed RFP-PSF (Supplemental Figure S6A), suggesting the colocalization of newly expressed CoAZ and PSF. Although only a small number of cells survived after extended periods of recording, the separation of GFP-CoAZ and RFP-PSF foci was detected around 13–14 h after sample examination (Supplemental Figure S6B). This result indicated that CoAZ initially accumulated in paraspeckles and later left from paraspeckles to adjacent and even distant locations. Unlike CoAZ, newly synthesized RFP-RBM4 colocalized with GFP-tagged splicing speckle protein SRSF1/ASF instead of GFP-PSF (Supplemental Figure S6C).

To confirm the correlation between CoAZ foci formation and paraspeckles, we depleted the paraspeckle essential factor PSF using small-interfering RNA (siRNA) and evaluated the localization of FLAG-CoAZ. PSF expression was suppressed by ∼80% in PSF siRNA-transfected HeLa cells, and the FLAG-CoAA signal was widely dispersed (Supplemental Figure S7), indicating disruption of paraspeckles (Figure 6B, siPSF). PSF knockdown also led to FLAG-CoAZ redistribution to the nucleoplasm and loss of discrete foci (Figure 6B). Moreover, siRNA-mediated depletion of the heterogeneous nuclear ribonucleoprotein K (hnRNP K), another paraspeckle protein (10), or *NEAT1* RNA (11–13) also resulted in FLAG-CoAZ diffusion in the nucleoplasm (Figure 6B). Therefore, it is possible that formation of CoAZ foci requires intact paraspeckles. Interestingly, RNase treatment did not affect CoAZ foci, but the PSF signal that surrounded regions of CoAZ foci in mock-treated cells completely disappeared (Figure 6C). This result suggested that once CoAZ foci are formed, they become independent of paraspeckles and their constituents.

To further our understanding of CoAZ foci, we examined the dynamics of GFP-CoAZ using FRAP. We initially observed that GFP-CoAZ foci moved away from the plane of focus rapidly at 37°C. Therefore, we performed the FRAP analysis at 25°C. The result revealed that the GFP-CoAZ foci were dynamic with a half-time of fluorescence recovery ($t_{1/2}$) of 18 s (Figure 6D), which was not exceptionally high. The detailed mechanism for CoAZ focus formation and movement warrants future investigation.

## DISCUSSION

RBM4 and CoAZ share a common C-terminal domain, and both proteins function as splicing regulatory factors, albeit with different activities (Figure 4). In this study, we highlighted the different subnuclear localization and biogenesis patterns of these two proteins and the influence of the Ala-tracts on these patterns.

### The Ala-rich tracts determine subnuclear destination of RBM4

RBM4 localizes to splicing speckles, and its CAD serves as a splicing speckle targeting signal (2) (Figure 7A). Here, we show that a small subpopulation of Ala-tract-deleted RBM4 (RBM4-A0) instead localized to paraspeckles (Figure 1), similar to RBM4b, likely owing to the naturally relatively short Ala-tract of RBM4b. This possibility was strengthened by the subnuclear localization pattern of two other chimeric RBM4 proteins, RBM4-Ala2b and RBM4-CP (Supplemental Figure S1). Moreover, our observation is consistent with recent reports that RBM4b is a constituent of paraspeckles (10,20). Therefore, Ala-tract truncation substantially altered the subnuclear targeting properties of the CAD. More importantly, the Ala-tract-deleted CAD (CAD-A0) was sufficient to direct protein to paraspeckles (Figure 2). It has been reported that the NONO/paraspeckle (NOPS) domain, a coiled-coil domain and RNA-binding domains are required for paraspeckle protein localization (22). CAD-A0, which neither contains a NOPS domain nor possesses RNA-binding activity, may represent a new paraspeckle targeting signal. We, therefore, propose that RBM4 occasionally associates with paraspeckles during the period when they transiently lie adjacent to splicing speckles (22) and that Ala-tract presence or length influences RBM4 retention in either subnuclear domain (Figure 7A).

CoAZ localized to distinct subnuclear foci that, intriguingly, often appeared immediately adjacent to or partially overlapped with paraspeckles (Figure 5). In contrast to RBM4, Ala-tract truncation appeared to have no effect on the subnuclear localization of CoAZ. Considering that CoAZ initially emerged in paraspeckles (Figures 6 and 7B) and that the Ala-tract-lacking CAD tended to localize to paraspeckles, the observation that CoAZ-A0 exhibited the same localization pattern as WT CoAZ was not surprising. Together, our results suggest that the presence or absence of Ala-tracts determines the subnuclear localization properties of the CAD of RBM4 and CoAZ.

### Biogenesis of CoAZ-containing foci

Owing to the low abundance of endogenous CoAZ (16), we examined the subnuclear localization of transiently overexpressed CoAZ. Time-lapse microscopy and paraspeckle protein knockdown experiments revealed that newly synthesized CoAZ initially accumulates in paraspeckles and that CoAZ foci formation is highly dependent on paraspeckle components, both proteins and noncoding RNA (Figure 6). This result suggests that CoAZ may not form foci in paraspeckle-deficient cells, such as undifferentiated stem cells (14). When exogenous CoAZ was expressed at very low levels, it tended to be more consistently retained in paraspeckles (Figure 5), perhaps mimicking the localization behavior of endogenous CoAZ. Nevertheless, the majority of transiently expressed CoAZ formed foci that eventually completely or partially separated from paraspeckles, and mature CoAZ foci were essentially RNase resistant, which was distinct from paraspeckles (Figures 5, 6 and 7B). Furthermore, we found that although CoAZ foci were highly dynamic, unlike RBM4 and paraspeckle proteins, CoAZ did not relocate to perinucleolar regions after transcription inhibition (Supplemental Figure S8). Therefore, it remains to be investigated which protein factors or even associated RNA provide the physiological driving force for the movement of CoAZ foci and for the unique localization and biochemical properties of CoAZ.

### Ala-repeat expansion restricts protein redistribution

Alanine-expanded PABPN1 accumulates in intranuclear inclusions in muscle of OPMD patients (23). A study reports that alanine-expansion prolongs the aggregation process of PABPN1 in the nucleoplasm and thus causes a delay in the inclusion formation of PABPN1 (24). Here we show that overexpressed RBM4-A25 exhibited a prominent splicing speckle localization pattern, similar to endogenous WT RBM4, but Ala-tract expansion significantly reduced the mobility of RBM4 (Figures 1 and 3). Moreover, transcription inhibition only relocated a small proportion of RBM4-A25 to the perinucleolar region (Figure 1). We thus hypothesized that expanded Ala-repeats induce RBM4 aggregation or increase its affinity for interacting partners, and deduced that the reduced capacity of RBM4-A25 to move within the nucleus might impair its translocation from splicing speckles to its functional sites, which may lead to relatively poor splicing activity compared with WT RBM4 (Figure 3). However, at first glance, Ala-tract expansion did not particularly increase the interaction of RBM4 with splicing speckle proteins (Supplemental Figure S2). The driving force for RBM4-A25 retention in splicing speckles remains further investigation. Nevertheless, taken together with the fact that Ala-tract expansion prevented CoAZ movement out of paraspeckles, the retention of RBM4-A25 in the splicing speckles suggests that Ala-tract expansion might restrict the relocalization or destination of RBM4/CoAZ following synthesis (Figure 7B).

In conclusion, we report the functional role of Ala-tracts in two closely related splicing factors that are located in different intranuclear subdomains. This study provides a plausible explanation for the distinct subnuclear localization of RBM4a and RBM4b and reveals that CoAZ localizes to novel nuclear foci, which, intriguingly, are closely associated but not overlapping with paraspeckles. This study also proposes a possible new mechanism for protein targeting to paraspeckles.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGMENT

We thank L. Ko for plasmids, and also the Core Facilities of the Institute of Biomedical Sciences, Academia Sinica for the assistance of microscopic studies and DNA sequencing.

## FUNDING

National Science Council of Taiwan [NSC 102-2811-B-001-066]. Funding for open access charge: National Science Council of Taiwan; Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.
*Conflict of interest statement.* None declared.

## REFERENCES

1. Lin,J.C. and Tarn,W.Y. (2012) Multiple roles of RBM4 in muscle cell differentiation. *Front. Biosci. (Schol Ed)*, **4**, 181–189.
2. Lai,M.C., Kuo,H.W., Chang,W.C. and Tarn,W.Y. (2003) A novel splicing regulator shares a nuclear import pathway with SR proteins. *EMBO J.*, **22**, 1359–1369.
3. Lu,C.C., Chen,T.H., Wu,J.R., Chen,H.H., Yu,H.Y. and Tarn,W.Y. (2013) Phylogenetic and molecular characterization of the splicing factor RBM4. *PLoS One*, **8**, e59092.
4. Lavoie,H., Debeane,F., Trinh,Q.D., Turcotte,J.F., Corbeil-Girard,L.P., Dicaire,M.J., Saint-Denis,A., Page,M., Rouleau,G.A. and Brais,B. (2003) Polymorphism, shared functions and convergent evolution of genes with sequences coding for polyalanine domains. *Hum. Mol. Genet.*, **12**, 2967–2979.
5. Brown,L.Y. and Brown,S.A. (2004) Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends Genet.*, **20**, 51–58.
6. Jenal,M., Elkon,R., Loayza-Puch,F., van Haaften,G., Kuhn,U., Menzies,F.M., Oude Vrielink,J.A., Bos,A.J., Drost,J., Rooijers,K. *et al.* (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell*, **149**, 538–553.
7. Albrecht,A. and Mundlos,S. (2005) The other trinucleotide repeat: polyalanine expansion disorders. *Curr. Opin. Genet. Dev.*, **15**, 285–293.
8. Calado,A., Tome,F.M., Brais,B., Rouleau,G.A., Kuhn,U., Wahle,E. and Carmo-Fonseca,M. (2000) Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum. Mol. Genet.*, **9**, 2321–2328.
9. Fox,A.H. and Lamond,A.I. (2010) Paraspeckles. *Cold Spring Harb. Perspect. Biol.*, **2**, a000687.
10. Naganuma,T., Nakagawa,S., Tanigawa,A., Sasaki,Y.F., Goshima,N. and Hirose,T. (2012) Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.*, **31**, 4020–4034.
11. Clemson,C.M., Hutchinson,J.N., Sara,S.A., Ensminger,A.W., Fox,A.H., Chess,A. and Lawrence,J.B. (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell*, **33**, 717–726.
12. Sasaki,Y.T., Ideue,T., Sano,M., Mituyama,T. and Hirose,T. (2009) MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl Acad. Sci. U.S.A.*, **106**, 2525–2530.
13. Sunwoo,H., Dinger,M.E., Wilusz,J.E., Amaral,P.P., Mattick,J.S. and Spector,D.L. (2009) MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.*, **19**, 347–359.
14. Chen,L.L. and Carmichael,G.G. (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell*, **35**, 467–478.
15. Prasanth,K.V. and Spector,D.L. (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev.*, **21**, 11–42.
16. Brooks,Y.S., Wang,G., Yang,Z., Smith,K.K., Bieberich,E. and Ko,L. (2009) Functional pre- mRNA trans-splicing of coactivator CoAA and corepressor RBM4 during stem/progenitor cell differentiation. *J. Biol. Chem.*, **284**, 18033–18046.
17. Prakash,T., Sharma,V.K., Adati,N., Ozawa,R., Kumar,N., Nishida,Y., Fujikake,T., Takeda,T. and Taylor,T.D. (2010) Expression of conjoined genes: another mechanism for gene regulation in eukaryotes. *PLoS One*, **5**, e13284.
18. Lee,K.M., Hsu,I.W. and Tarn,W.Y. (2010) TRAP150 activates pre-mRNA splicing and promotes nuclear mRNA degradation. *Nucleic Acids Res.*, **38**, 3340–3350.
19. Prasanth,K.V., Prasanth,S.G., Xuan,Z., Hearn,S., Freier,S.M., Bennett,C.F., Zhang,M.Q. and Spector,D.L. (2005) Regulating gene expression through RNA nuclear retention. *Cell*, **123**, 249–263.
20. Fong,K.W., Li,Y., Wang,W., Ma,W., Li,K., Qi,R.Z., Liu,D., Songyang,Z. and Chen,J. (2013) Whole-genome screening identifies proteins localized to distinct nuclear bodies. *J. Cell Biol.*, **203**, 149–164.
21. Markus,M.A. and Morris,B.J. (2006) Lark is the splicing factor RBM4 and exhibits unique subnuclear localization properties. *DNA Cell Biol.*, **25**, 457–464.
22. Fox,A.H., Bond,C.S. and Lamond,A.I. (2005) P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. *Mol. Biol. Cell*, **16**, 5304–5315.
23. Brais,B., Bouchard,J.P., Xie,Y.G., Rochefort,D.L., Chretien,N., Tome,F.M., Lafreniere,R.G., Rommens,J.M., Uyama,E., Nohira,O. *et al.* (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat. Genet.*, **18**, 164–167.
24. Raz,V., Abraham,T., van Zwet,E.W., Dirks,R.W., Tanke,H.J. and van der Maarel,S.M. (2011) Reversible aggregation of PABPN1 pre-inclusion structures. *Nucleus*, **2**, 208–218.