Targeting of RBM10 to S1-1 Nuclear Bodies: Targeting Sequences and its Biological Significance

Running Title: RBM10 and S1-1 nuclear bodies

Ling-Yu Wang[1,2,#], Sheng-Jun Xiao[1,3,#], Hiroyuki Kunimoto[1], Kazuaki Tokunaga[4], Hirotada Kojima[1], Masatsugu Kimura[5], Takahiro Yamamoto[6], Naoki Yamamoto[7,8], Zhao Hong[1], Koji Nishio[9], Hideo Yamane[10], Tokio Tani[4], Koichi Nakajima[1], Hiroyoshi Iguchi[10], and Akira Inoue[1,10*]

Departments of [1]Immunology, [5]Radioisotope Center,and [10]Otolaryngology, Osaka City University Graduate School of Medicine, Osaka 545-8585, Japan; Departments of [2]Human Genetics and [3]Pathology, Guilin Medical University, Guilin 541004, P. R. China; [4]Department of Biological Sciences, Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Kumamoto 860-8555, Japan; [6]Graduate School of Life Science, Hokkaido University, Sapporo 060-0808, Japan;[7]Psychiatry Department, Tokyo Metropolitan Tama Medical Center, Tokyo 183-8524, Japan;[8]Department of Biological Sciences, Tokyo Metropolitan University, Tokyo 192-0397, Japan;[9]Department of Anatomy and Neurosciences, Graduate School of Medicine, Nagoya University, Nagoya 466-8550, Japan.

*Corresponding author: Akira Inoue, Department of Otolaryngology, Osaka City University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan;

E-mail: ainoue@med.osaka-cu.ac.jp

#These authors contributed equally to this work.

Key words: RBM10/S1-1, S1-1 nuclear body, nuclear body-targeting sequence, alternative splicing, transcription, C_2H_2 Zn finger

Abstract

RBM10 is an RNA-binding protein that regulates alternative splicing (AS). This protein localizes to the extra-nucleolar nucleoplasm and S1-1 nuclear bodies (NBs). We investigated the biological significance of RBM10 localization to S1-1 NBs, which is poorly understood. Our analyses revealed that RBM10 possesses two S1-1 NB-targeting sequences (NBTSs), one in the KEKE motif region and another in the C_2H_2 Zn finger (ZnF). These NBTSs acted synergistically and were sufficient for localization of RBM10 to S1-1 NBs. Furthermore, the C_2H_2 ZnF not only acted as an NBTS, but was also essential for regulation of AS by RBM10. RBM10 did not participate in S1-1 NB formation. We confirmed the previous finding that localization of RBM10 to S1-1 NBs increases as cellular transcriptional activity decreases and vice versa. These results indicate that RBM10 is a transient component of S1-1 NBs and is sequestered in these structures via its NBTSs when cellular transcription decreases. We propose that the NB-targeting activity of the C_2H_2 ZnF is induced when it is not bound to pre-mRNA or the splicing machinery complex under conditions of reduced transcription.

Introduction

RBM10 (previously called S1-1) is an RNA-binding protein (Inoue et al., 1996) that regulates alternative splicing (AS) of pre-mRNAs (Wang et al., 2013) via exon skipping, as shown for the Drosophila protein Discs large homolog 4 (Zheng et al., 2013), the Notch pathway regulator NUMB (Bechara et al., 2013), and the apoptosis-related protein Fas (Inoue et al., 2014), or via selection of an internal 5′-splice site in the target exon, as shown for the apoptosis-related protein Bcl-x (Inoue et al., 2014). Loss of RBM10 function causes TARP syndrome (talipes equinovarus, atrial septal defect, Robin sequence, and persistent left superior vena cava), which is characterized by various developmental anomalies, such as cleft palate and malformation of the heart (Johnston et al., 2010; Gripp et al., 2011), as well as other diseases, including cancers such as lung adenocarcinoma (Imielinski et al., 2012; Cancer Genome Atlas Research Network, 2014) (for a review, see Loiselle and Sutherland, 2018).

Nuclear bodies (NBs) are membrane-less compartments of the nucleus. There are various types of NBs, and their formation, compositions, and roles in relation to the biological activities of the nucleus have been studied (for a review see Mao et al., 2011 and Staněk and Fox 2017). While some types of NBs, such as Cajal bodies, nuclear speckles, and paraspeckles, have been extensively investigated, the biological significance of S1-1 NBs and

their relationship with RBM10 remain largely unknown. RBM10 localizes to S1-1 NBs and S1-1 granules in the nucleus. S1-1 NBs are about 0.5–2.0 μm in size, with usually 10–20 per nucleus. S1-1 granules, which localize to the perichromatin fibrils that comprise transcription and splicing sites, are much smaller and more abundant than S1-1 NBs. Similar to paraspeckles, S1-1 NBs are often closely associated with nuclear speckles. The structure of S1-1 NBs changes according to transcriptional activity; when cellular transcription is globally reduced, RBM10 accumulates in S1-1 NBs and these structures become enlarged (Inoue et al., 2008).

In the present study, we identified the S1-1 NB-targeting sequences (NBTSs) within RBM10 that participate in the targeting and localization of this protein to S1-1 NBs. In addition, we elucidated the biological significance of this localization in relation to the AS regulatory function of RBM10. We discuss that RBM10 is reversibly regulated by sequestration in S1-1 NBs in response to changes in cellular transcriptional activity.

Results

Specific sequences are responsible for targeting of RBM10 to S1-1 NBs

Wild type and deletion mutants of the 852-amino acid (aa) isoform of RBM10 linked in-frame to the N-terminus of EGFP (S1-1/EGFP-N1) were expressed in ARL (adult rat liver epithelial) cells. The subcellular localizations of these proteins were assessed by examining EGFP fluorescence and by staining with anti-S1-1 antiserum, which recognized an epitope within aa 8–63 of RBM10 (Supporting inform. Fig. S1). In contrast with the aa 4–414 deletion, the aa 4–742 deletion completely abolished localization to S1-1 NBs (Fig. 1A and B), suggesting that an NBTS is present within aa 415–742.

The location of the tentative NBTS was further defined as being within aa 411–680 using deletion constructs 1–4 [Fig. 2A and 2C (a)]. The product of construct 3 localized weakly to S1-1 NBs, suggesting that an NBTS is present within aa 543–586 [Fig. 2C (a)] and that another NBTS is present in the unexamined aa 681–742 region. The NBTS within aa 543–586 was further analyzed using constructs 5–8 [Fig. 2B and 2C (b)]. The products of constructs 5 and 6 localized weakly to S1-1 NBs, indicating that the NBTS is present within aa 532–577 [Fig. 2C (b)]. Based on these results, an upstream NBTS (NBTS1) was concluded to be present within aa 543–577 [Fig. 2C (c)]. A KEKE motif, which is characterized by

alternating positive and negative aa residues (Realini et al., 1994), is located at aa 549–568 [Fig. 3 (1)].

RBM10 contains another NBTS

The other NBTS located within aa 681–742 was examined using the constructs pNB1/EGFP-C3 and pNB2/EGFP-C3, where *NB1* and *NB2* encoded aa 476–645 and aa 476–760 of RBM10, respectively (Fig. 4A). NB1 and NB2 lacked the epitope recognized by the anti-S1-1 antiserum; thus, the antiserum was used to detect endogenous intact RBM10 localized to S1-1 NBs as a control.

NB1 contained the KEKE region/NBTS1 and a weak nuclear localization sequence (NLS3) (Xiao et al., 2013). Consequently, NB1 was not only detected in S1-1 NBs, but also in the nucleoplasm and cytoplasm (Fig. 4A). By contrast, NB2 harbored a strong NLS (NLS1) in its C-terminus and consequently localized predominantly to the nucleus. Moreover, similar to endogenous RBM10, NB2 localized to distinct S1-1 NBs owing to an additional NBTS within aa 681–742.

This region (aa 681–742) harboring the second putative NBTS contains a C_2H_2-type Zn finger (ZnF) at aa 681–706 [Fig. 3 (2)] and a downstream sequence (designated after-ZnF) at aa 707–742. These regions and a control region (aa 644–670) were examined by similar

deletion analysis. Only the ZnF deletion mutant localized weakly to S1-1 NBs (Fig. 4B). To confirm that the ZnF possesses NB-targeting activity, the cysteine residues critical for ZnF function at aa 683 and 686 were substituted with alanine or the histidine residue at aa 699 was substituted with serine [Figs 3 (2) and 4C]. The localizations of these two ZnF mutants to S1-1 NBs were weaker than those of parental NB2 and the after-ZnF deletion mutant (Δ after-ZnF). However, the ZnF mutants still localized weakly to S1-1 NBs due to the presence of NBTS1. The second NBTS in the ZnF region was called NBTS2 [Fig. 3 (2)].

Verification of NBTS1 and NBTS2

As shown above, the NB-targeting activity of NB2 was retained upon single deletion of NBTS1 (KEKE region) or NBTS2 (C_2H_2 ZnF). Importantly, NB2 did not clearly localize to S1-1 NBs upon simultaneous deletion of both NBTSs [Fig. 5A (1)], demonstrating their synergistic action in NB localization.

Using ImageJ Plot Profile, NB localization was analyzed in a more quantitative manner. RBM10 does not localize to the nucleolus (Inoue et al., 2008); thus, analysis was performed using lines that traversed the extra-nucleolar region [Fig. 5A (2)]. In contrast with NB2 and Δ after-ZnF, the NB2 mutants lacking the KEKE region (NBTS1) or the C_2H_2 ZnF region (NBTS2) showed blunt peaks, and the double deletion mutant showed no distinct

peaks. Consistently, the number of S1-1 NBs larger than 0.5 μm decreased upon deletion of NBTSs [Fig. 5A (3)]. The results indicate that NBTS1 and NBTS2 synergistically enhance NB localization.

The NBTSs were further verified by employing clones that expressed the KEKE or C_2H_2 ZnF region linked to the FLAG sequence. In contrast with the control octamer repeat sequence (OCRE), KEKE and C_2H_2 ZnF localized to S1-1 NBs (Fig. 5B). Based on these results, we conclude that both the KEKE region (NBTS1) and the C_2H_2 ZnF region (NBTS2) possess NB-targeting activity.

The C_2H_2 ZnF is essential for the AS function of RBM10

RBM10 regulates AS. We assessed whether the identified NBTSs are associated with this function of RBM10. Wild-type RBM10 and RBM10 harboring mutations in NBTS1 (KEKE region) or NBTS2 (C_2H_2 ZnF) were expressed in RBM10-knockout HepG2 cells and their AS activities in pre-mRNA splicing of *fas* were examined. In this system, wild-type RBM10 causes skipping of exon 6 in *fas* pre-mRNA (Inoue et al., 2014). pre-mRNA splicing of *fas* was not affected by mutation of the critical KEKKE sequence in the KEKE region/NBTS1 to AAKAA, but was almost completely abolished by mutation of the CLLC sequence in the C_2H_2 ZnF/NBTS2 to ALLA (Fig. 6, see also Fig. 3), demonstrating that the C_2H_2 ZnF is essential

for regulation of AS by RBM10. Importantly, these results also indicate that the C_2H_2 ZnF has two functions: regulation of AS and localization of RBM10 to S1-1 NBs.

RBM10 does not participate in the structural organization of S1-1 NBs

The possibility that RBM10 localizes to S1-1 NBs and participates in the formation of these structures was examined. We previously showed that RNA is stored in nuclear domains called TIDRs (transcription-inactivation-dependent RNA domains) when HeLa cells are injected with Cy3-labeled *ftz* model pre-mRNA and successively incubated with actinomycin D (Act D) (Tokunaga et al., 2006) and that TIDRs and S1-1 NBs are the same structural entities (Inoue et al., 2008). Using this experimental system, we examined whether S1-1 NBs/TIDRs formed in the absence of RBM10. For this purpose, RBM10 was knocked down using specific siRNAs in HeLa cells (Fig. 7a), Cy3-labeled *ftz* pre-mRNA was introduced into cells, and cells were incubated with Act D. S1-1 NBs/TIDRs formed even in the absence of RBM10, as visualized by Cy3 fluorescence of *ftz* RNA (Fig. 7b). Thus, we conclude that RBM10 does not participate in the structural organization of S1-1 NBs.

RBM10 is transiently sequestered in S1-1 NBs when transcription is reduced

We previously showed that localization of RBM10 to S1-1 NBs is closely associated with reduced transcription (Inoue et al., 2008). When cells were incubated with Act D, the localization of RBM10 to S1-1 NBs increased in a time-dependent manner. These NBs became larger and fewer in number [Fig. 8A (1)]; however, the cellular level of RBM10 does not change upon incubation with Act D (Inoue et al., 2008). Enlarged S1-1 NBs also form upon heat shock, serum starvation, and confluent cell culture, where cellular transcriptional activity is commonly decreased, and upon treatment with the RNA polymerase II (pol II) inhibitor α-amanitin (Inoue et al., 208). We verified these previous results using other pol II inhibitors, namely, flavopiridol and 5,6-dichloro-1-β-D-ribo-furanosylbenzimidazole (DRB) [Fig. 8A (2)]. These reagents inhibit CDK9, the kinase subunit of the positive transcription elongation factor P-TEFb, which is essential for productive transcription (Chao and Price, 2001). These two inhibitors were examined because flavopiridol also inhibits the cell cycle kinases CDK1 and CDK4, while DRB also inhibits casein kinase II (CK2) (Meggio et al., 1990, Bensaude, 2011). RBM10 accumulated in S1-1 NBs upon treatment with either inhibitor [Fig. 8A (2)], suggesting that this accumulation was due to inhibition of P-TEFb. In support of this, the effects of DRB [Fig. 8A (2)] were not replicated by treatment with the potent CK2 inhibitor LY294002 (Gharbi et al., 2007) or knockdown of CK2 with specific siRNAs (data not shown). Thus, the formation of enlarged S1-1 NBs upon treatment with

flavopiridol and DRB is due to their inhibition of P-TEFb, and RBM10 accumulates in S1-1 NBs when transcription is arrested.

The changes in S1-1 NBs were reversible. When cells were washed with fresh medium, the enlarged S1-1 NBs formed in the presence of α-amanitin or DRB returned to their initial state within 4 and 1 h, respectively [Fig. 8A (2)]. The difference in the amount of time required to reverse the effect is likely due to differences in the affinities of these inhibitors for their binding sites or because α-amanitin also triggers pol II degradation (Nguyen et al, 1996).

Figure 8B (1) shows the effect of DRB on cellular transcriptional activity. Transcriptional activity decreased almost to the basal level upon treatment with DRB for 45 min. By contrast, the protein level of RBM10 remained constant upon treatment with DRB for 3 h [Fig. 8B (2)], as previously shown upon treatment with Act D for 7 h (Inoue et al., 2008). In addition, following a chase with fresh medium, transcriptional activity returned close to the initial level after 1 h, and accordingly enlarged S1-1 NBs returned to their original state [Fig. 8B (1) and A (2)]. The RBM10 protein level remained constant during the chase period [Fig. 8B (2)]. Thus, the localization of RBM10 to S1-1 NBs changes in a dynamic and reversible manner according to the level of cellular transcriptional activity, without a change in the protein level of RBM10. The dynamic and reversible localization of RBM10 to S1-1

NBs, as well as the finding that RBM10 does not participate in the formation of S1-1 NBs (Fig. 7), indicate that RBM10 is a transient component of S1-1 NBs and is sequestered in these structures when cellular transcription decreases. These results also suggest that S1-1 NB sequesters proteins, in addition to other possible functions.

Discussion

The present study demonstrates that RBM10 has two NBTSs, called NBTS1 and NBTS2, which are located in the KEKE motif and C_2H_2 ZnF regions, respectively. Each NBTS alone can target and localize to S1-1 NBs; however, both are required for increased localization of RBM10 to S1-1 NBs. Thus, NBTS1 and NBTS2 act in a synergistic manner. NBTS1 and NBTS2 are separated by approximately 110 aa residues and have no sequence homology (Figs 1 and 3). Thus, these NBTSs likely interact with spatially close but distinct structures within S1-1 NBs, the geometry of which seems to enable their synergistic action.

Under reduced transcription conditions, RBM10 localizes to S1-1 NBs and these structures become enlarged. When transcriptional activity is restored, enlarged S1-1 NBs return to their initial state and RBM10 leaves from these structures (Inoue et al, 2008, and Fig. 8). The present study showed that RBM10 does not participate in the formation of S1-1 NBs and that the cellular level of RBM10 is maintained upon inhibition of transcription and

during the subsequent chase period. These findings indicate that RBM10 is a transient component of S1-1 NBs and is stored or sequestered in these structures when transcription decreases.

The identified NBTSs were sufficient for localization of RBM10 to S1-1 NBs. All known NBs, including S1-1 NBs (Inoue et al., 2008), are membrane-less compartments (for a review, Mao et al., 2011); therefore, we suggest that targeting and localization of transient NB proteins to NBs are based on the association of specific sequences of the proteins, here NBTSs, with acceptor sites in NBs. This may be obvious, but it is worth noting.

The C_2H_2 ZnF/NBTS2 was bi-functional: it was not only important for NB-targeting/localization, but was also essential for regulation of AS by RBM10. Hernández et al. (2016) reported that deletions from the C-terminus, including the G-patch and C_2H_2 ZnF, result in progressive and partial loss of the AS regulatory activity of RBM10, indicating that the integrity of RBM10 is required for its full function. Our results support this previous study and further demonstrate that substitution of the two critical cysteine residues of the C_2H_2 ZnF with alanine was sufficient to completely abolish the AS regulatory activity of RBM10 (Fig. 6).

How the ZnF performs the dual functions of NB-targeting and AS regulation is an intriguing question. When transcription is suppressed, the level of nucleoplasmic pre-mRNAs

decreases and the C_2H_2 ZnF no longer binds to its substrates or the splicing machinery complex, which functions together with RBM10. This unoccupied state seems to induce the NB targeting activity of the C_2H_2 ZnF, thereby increasing the localization of RBM10 to S1-1 NBs. Thus, the C_2H_2 ZnF likely acts as a sensor to enhance sequestration of RBM10 in S1-1 NBs.

The *RBM10* gene resides on the X chromosome, and one of its copies is inactivated by X-inactivation in females (Coleman, et al., 1996, Thiselton, et al., 2002, Goto and Kimura, 2009). This implies that a double dose of RBM10 is unfavorable to cells and that the cellular level of active RBM10 must be controlled. In this regard, sequestration of RBM10 in S1-1 NBs may be a mechanism to rapidly and efficiently regulate the active nucleoplasmic level of RBM10. Specifically, RBM10 is promptly released from S1-1 NBs and localizes to transcription/splicing sites in the nucleoplasm when transcription increases, while it is stored in S1-1 NBs when transcription decreases. This sequestration may prevent excessive RBM10 from engaging in unfavorable AS of irrelevant pre-mRNAs, which would produce products that elicit unfavorable effects on cells.

Dysfunction of RBM10 leads to various diseases. Specific defects in NBTS activity may also cause RBM10-related diseases. Further studies are required to elucidate the molecular mechanisms that regulate sequestration of RBM10 in S1-1 NBs, including the

identification and characterization of the acceptor sites for RBM10 in S1-1 NBs and the resident core proteins that comprise S1-1 NBs.

Experimental procedures

Cells and cell culture. HeLa cells derived from human cervical carcinoma, HepG2 cells derived from human hepatocellular carcinoma, and HEK293T cells derived from human embryonic kidney and transformed with the SV40 large T-antigen were cultured in Dulbecco's Modified Eagle Medium (Nissui). ARL cells (rat liver epithelial cells; Japan Health Sci. Found., JCRB0248) were cultured in Williams' Medium E (Gibco BRL). The media were supplemented with 5% fetal calf serum (Gibco BRL), 100 μg/mL streptomycin, and 100 IU/mL penicillin. Fetal calf serum was used at a concentration of 10% for culture of HepG2 cells. In inhibition experiments, ARL cells were incubated with 5 μg/mL Act D, 50 μg/mL α-amanitin, 0.5 mM DRB, or 0.2 μM flavopiridol.

Immunostaining. Immunofluorescence analysis was performed as described previously (Inoue et al., 2008). RBM10 and S1-1 NBs were stained with anti-S1-1 antiserum (Inoue et al., 2008; 1:250 working dilution) and rhodamine-conjugated goat anti-rabbit IgG (1:200 dilution; ICN, 55662). Stained cells were examined using a fluorescence microscope (Olympus, type IX70; numerical aperture, 0.4; and objective lens, ×20). Data were captured

using image acquisition software (DP Controller). Images were processed for contrast and/or brightness.

Immunoblotting. Western blotting was performed as previously described (Inoue et al., 2008). Total proteins (equivalent to 0.2 μg of cellular DNA) were electrophoresed on 10% sodium dodecyl sulfate-polyacrylamide gels (e-PARGEL; Atto).

Primers, RBM10 mutant plasmids, and transfection. The PCR primers used to prepare the various constructs are listed in Table S1. Each primer contained a restriction sequence at one end. The PCR conditions were as follows: elongation (30 cycles, 5 s/kb) at 72°C, denaturation (10 s) at 95°Cor 98°C, and annealing (15 s) at 3.5–5.0°C below the lower T_m of each primer set. PCR products were digested with the corresponding restriction enzymes, purified, and ligated into pEGFP-N1, pEGFP-C3 (Clontech), or p3xFLAG-Myc-CMV-24 (Sigma-Aldrich). Competent JM109 *Escherichia coli* cells (Takara) were transformed with the constructs, the resulting colonies were cultured, and recombinant plasmids were isolated using the Pure Yield Plasmid Miniprep System (Promega). The sizes and DNA sequences of the purified plasmids were verified by agarose gel electrophoresis and sequencing with a 3130x1 DNA Sequencer (ABI), respectively. Transfection was performed using Lipofectamine LTX and PLUS Reagent (Invitrogen) according to the manufacturer's recommendations.

Constructs and mutagenesis of RBM10. The constructs ND1, ND11, and ND15, and that encoding aa 1–469 were produced by sequential deletion from pS1-1/EGFP-N1 (Inoue et al., 2008) as described previously (Xiao et al., 2013). To generate pNB1/EGFP-C3 and pNB2/EGFP-C3, sequences encoding aa 476–645 and aa 476–760 of RBM10 were PCR-amplified and ligated into the *Sac*I and *Pst*I sites of pEGFP-C3. Deletion or substitution mutagenesis of these plasmids was performed using a PrimeSTAR Mutagenesis Basal Kit (Takara) and the primers listed in Table S1. The sequences encoding aa 475–533, the KEKE region (aa 529–590 or aa 529–571), and the C_2H_2 ZnF region (aa 676–711) were similarly amplified (Table S1) and ligated into the *Hin*dIII and *Kpn*I sites of the p3xFLAG-Myc-CMV-24 vector. The two KEKE constructs yielded similar results.

RBM10 knockdown. HeLa cells were transfected with 40 nM specific human RBM10-siRNA5 or control AllStar siRNA (Qiagen) using Lipofectamine 2000 (Invitrogen) and incubated for 3 days, in accordance with the manufacturer's protocol. A second specific siRNA, which was designed to target a different region of RBM10 mRNA, yielded similar results, indicating that the results obtained were not due to an off-target effect.

Nuclear injection of *ftz* pre-mRNA. Cy3-labeled *ftz* pre-mRNA was microinjected into the nuclei of HeLa cells at approximately 80% confluency using a Microinjector 5170 (Eppendorf) as previously described (Tokunaga et al., 2006). Cells were incubated for 2 h in

the presence of Act D (5.0 μg/mL). Enlarged S1-1 NBs/TIDRs were observed by fluorescence microscopy (Tokunaga et al., 2006).

RBM10 knockout. RBM10-knockout HepG2 cells were generated using the CRISPR/Cas9-mediated genome editing system. Two guide RNAs (cgttcatatcctcgcgagta and cagccgagaccacgactacc) designed to target exon 3 (14-bp apart) were cloned into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene) according to the manufacturer's instructions. HepG2 cells were transfected with the plasmids together with pEF-puromycin using Lipofectamine 3000 (Thermo Fisher Scientific Inc.) and incubated with 1 μg/mL puromycin for 2–3 days to obtain cell colonies. Knockout of RBM10 in the cloned cells was confirmed by Western blotting using anti-S1-1 antiserum and by DNA sequencing.

Generation of lentiviruses expressing wild-type and mutant RBM10. RBM10 mutants harboring mutations of KEKKE (to AAKAA) or CLLC (to ALLA), as well as a double mutant (to AAKAA/ALLA), were prepared by double-strand plasmid mutagenesis using the primer pairs listed in Table S1 and a Quikchange Site-Directed Mutagenesis Kit (Agilent Tech.). After verification by DNA sequencing, each RBM10 cDNA was ligated into the *Xho*I and *Not*I sites of the lentiviral vector pCSII-EF-MCS-IRES-Venus (a gift from H. Miyoshi, RIKEN, Tsukuba, Japan). Lentivirus expressing wild-type or mutant RBM10 was produced

by co-transfection of 293T cells with pCSII-EF-3xFLAG-wild type or -mutated RBM10, pCMV-VSV-G-RSV-Rev, and pCAG-HIVgp (Miyoshi et al., 1997).

Assessment of transcriptional activity. ARL cells at ~80% confluency in 3-cm plates were incubated with or without DRB (0.5 mM). [^{3}H]-uridine was added to the cultures 30 min before reaction termination. Cells were lysed in 0.2 mL of 0.2 M NaCl, 0.1% SDS, and Tris-HCl, pH 8.0. Plates were rinsed with 0.3 mL of 0.2 M NaCl and Tris-HCl, pH 8.0. The combined cell lysates were mixed with three volumes of ethanol and incubated at −30°C for 30 min. Precipitates were collected by centrifugation and digested in 0.2 mL of 2 mM $MgCl_2$, 10 mM Tris-HCl, pH 8.0, first with DNase I (10 μg) and RNase A (5 μg), and then with proteinase K (10 μg), each for 12 h. Aliquots of the digests were mixed with scintillation cocktail to measure radioactivity.

Acknowledgments

S.X. and L.W. are recipients of Human Resource Development Scholarships from the Japan International Cooperation Center (JICE), Japan. This work was supported by grants-in-aid for scientific research from the Japan Society for the Promotion of Science (JSPS) (KAKENHI grant numbers 21550162, J152640005, and 15K05569 to A.I., and 21590317 to K.N.).

References

Aravind, L. & Koonin, E. V. (1999). G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem. Sci. 24*, 342-344. https://doi.org/10.1016/S0968-0004(99)01437-1

Bechara, E. G., Sebestyén, E., Bernardis, I., Eyras, E. & Valcárcel, J. (2013). RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol. Cell 52*, 720-733. doi: 10.1016/j.molcel.2013.11.010

Bensaude, O. (2011). Inhibiting eukaryotic transcription. *Transcription 2*, 103-108. DOI: 10.4161/trns.2.3.16172

Cancer Genome Atlas Research Network. (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature 511*, 543-50. DOI: 10.1038/nature13385.

Chao, S. H. & Price, D. H. (2001). Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. *J. Biol. Chem. 276*, 31793-31799. DOI 10.1074/jbc.M102306200

Coleman, M. P., Ambrose, H. J., Carrel, L., Németh, A. H., Willard, H. F. & Davies, K. E. (1996). A novel gene, DXS8237E, lies within 20 kb upstream of UBE1 in Xp11.23 and has a different X inactivation status. *Genomics 31*, 135-138. https://doi.org/10.1006/geno.1996.0022

Gharbi, S. I., Zvelebil, M. J., Shuttleworth, S. J., Hancox, T., Saghir, N., Timms, J. F. & Waterfield, M. D. (2007). Exploring the specificity of the PI3K family inhibitor LY294002. *Biochem. J. 404*, 15-21. DOI: 10.1042/BJ20061489

Goto, Y. & Kimura, H. (2009). Inactive X chromosome-specific histone H3 modifications and CpG hypomethylation flank a chromatin boundary between an X-inactivated and an escape gene. *Nucl. Acids Res. 37*, 7416-7428. DOI:10.1093/nar/gkp860

Gripp, K. W., Hopkins, E., Johnston, J. J., Krause, C., Dobyns, W. B. & Biesecker, L. G. (2011). Long-term survival in TARP syndrome and confirmation of RBM10 as the disease-causing gene. *Am. J. Med. Genet. A. 155A*, 2516-2520. DOI: 10.1002/ajmg.a.34190

Hernández, J., Bechara, E., Schlesinger, D., Delgado, J., Serrano, L. & Valcárcel, J. (2016) Tumor suppressor properties of the splicing regulatory factor RBM10. *RNA Biol. 13*, 466-72. doi: 10.1080/15476286.2016.1144004

Imielinski, M., Berger, A. H., Hammerman, P. S., Hernandez, B., Pugh, T. J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., et al. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell 150*, 1107-1120. DOI: 10.1016/j.cell.2012.08.029

Inoue, A., Takahashi, K. P., Kimura, M., Watanabe, T. & Morisawa, S. (1996). Molecular cloning of an RNA binding protein. S1-1. *Nucl. Acids Res. 24*, 2990-2997. https://doi.org/10.1093/nar/24.15.2990

Inoue, A., Tsugawa, K., Tokunaga, K., Takahashi, K. P., Uni, S., Kimura, M., Nishio, K., Yamamoto, N., Honda, K., Watanabe, T., Yamane, H. & Tani, T. (2008). S1-1 nuclear domains: characterization and dynamics as a function of transcriptional activity. *Biol. Cell 100*, 523-535. DOI: 10.1042/BC20070142.

Inoue, A., Yamamoto, N., Kimura, M., Nishio, K., Yamane, H. & Nakajima K. (2014). RBM10 regulates alternative splicing. *FEBS Lett. 588*, 942-947. https://doi.org/10.1016/j.febslet.2014.01.052

Johnston, J. J., Teer, J. K., Cherukuri, P. F., Hansen, N. F., Loftus, S. K., NIH Intramural Sequencing Center (NISC), Chong, K., Mullikin, J. C. & Biesecker, L. G. (2010). Massively parallel sequencing of exons on the X chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate. *Am. J. Hum. Genet. 14*, 743-748. doi: 10.1016/j.ajhg.2010.04.007

Loiselle, J. J. & Sutherland, L. C. (2018) RBM10: Harmful or helpful-many factors to consider. *J. Cell. Biochem. 119*, 3809-3818. doi: 10.1002/jcb.26644

Mao, Y. S., Zhang, B. & Spector, D. L. (2011) Biogenesis and function of nuclear bodies. *Trends in Genetics 27*, 295-306. doi:10.1016/j.tig.2011.05.006

Martin, B. T., Serrano, P., Geralt, M. & Wüthrich, K. (2016). Nuclear magnetic resonance structure of a novel globular domain in RBM10 containing OCRE, the Octamer repeat sequence motif. *Structure 24*, 158-164. https://doi.org/10.1016/j.str.2015.10.029

Meggio, F., Shugar, D. & Pinna, L. A. (1990). Ribofuranosyl-benzimidazole derivatives as inhibitors of casein kinase-2 and casein kinase-1. *Eur. J. Biochem. 187*, 89-94.

Miyoshi, H., Takahashi, M., Gage, F. H. & Verma, I. M. (1997). Stable and efficient gene transfer into the retina using an HIV-based lentiviral vector. *Proc. Natl. Acad. Sci. USA 94*, 10319-10323.

Nguyen, V. T., Giannoni, F., Dubois, M. F., Seo, S. J., Vigneron, M., Kédinger, C., et al. (1996). In vivo degradation of RNA polymerase II largest subunit triggered by α-amanitin. *Nucl. Acids Res. 24*, 2924-2929.

Realini, C., Rogers, S. W. & Rechsteiner, M. (1994). KEKE motifs: Proposed roles in protein- protein association and presentation of peptides by MHC Class I receptors. *FEBS Letters 348*, 109-113. https://doi.org/10.1016/0014-5793(94)00569-9

Sloan, K. E. & Bohnsack, M. T. (2018) Unravelling the Mechanisms of RNA Helicase Regulation. *Trends in Biochemical Sciences 43*, 237-250. doi.org/10.1016/j.tibs.2018.02.001.

Staněk, D. & Fox, A. H. (2017). Nuclear bodies: news insights into structure and function. *Curr Opin Cell Biol. 46*, 94-101. doi: 10.1016/j.ceb.2017.05.001.

Thiselton, D. L., McDowall, J., Brandau, O., Ramser, J., d'Esposito, F., Bhattacharya, S. S., Ross, M. T., Hardcastle, A. J. & Meindl, A. (2002). An integrated, functionally annotated gene map of the DXS8026–ELK1 interval on human Xp11. 3–Xp11. 23: Potential hotspot for neurogenetic disorders. *Genomics 79*, 560-572. https://doi.org/10.1006/geno.2002.6733

Tokunaga, K., Shibuya, T., Ishihama, Y., Tadakuma, H., Ide, M., Yoshida, M., Funatsu, T., Ohshima, Y. & Tani, T. (2006). Nucleocytoplasmic transport of fluorescent mRNA in living mammalian cells: nuclear mRNA export is coupled to ongoing gene transcription. *Genes Cells 11*, 305-317. https://doi.org/10.1111/j.1365-2443.2006.00936.x

Wang, Y., Gogol-Döring, A., Hu, H, Fröhler, S., Ma, Y., Jens, M., Maaskola, J., Murakawa, Y., Quedenau, C., Landthaler, M., et al. (2013). Integrative analysis revealed the molecular mechanism underlying RBM10-mediated splicing regulation. *EMBO Mol. Med. 5*, 1431-1442. doi: 10.1002/emmm.201302663

Xiao, S., Wang, L., Kimura, M. Kojima, H., Kunimoto, H., Nishiumi, F., Yamamoto, N., Nishio, K., Fujimoto, S., Kato, T., et al. (2013). S1-1/RBM10: Multiplicity and cooperativity

of nuclear localization domains. *Biol. Cell 105*, 162-174.

https://doi.org/10.1111/boc.201200068

Zheng, S., Damoiseaux, R., Chen, L. & Black, D. L. (2013) A broadly applicable high-

throughput screening strategy identifies new regulators of Dlg4 (Psd-95) alternative splicing.

Genome Res. ***23***, 998–1007. doi: 10.1101/gr.147546.112

Figure legends

Fig. 1 RBM10 is targeted to NBs.

A. Structure of the 852-aa isoform of RBM10. RRM I and II: RNA recognition motif I and II, respectively; E4: a 78-aa region encoded by exon 4 and expressed in the 930-aa isoform; C_4 ZnF and C_2H_2 ZnF: Cys_4-type and Cys_2His_2-type ZnFs, respectively; NLS1: the principal nuclear localization sequence (Xiao et al., 2013); G-patch: a motif sometimes found in eukaryotic RNA-processing proteins (Aravind and Koonin, 1999; Sloan and Bohnsack, 2018). The OCRE, KEKE motif, and amphipathic α-helix (AHH) are described in Fig. 3. **B.** Presence of an NBTS within aa 415–742. Plasmids ND1, ND11, and ND15 were derived from pS1-1(RBM10)/EGFP-N1 by deleting the regions corresponding to aa 2–4, 4–414, and 4–742, respectively, and expressed in ARL cells. Green, EGFP; red, RBM10 stained with anti-S1-1 antiserum. Some S1-1 NBs are indicated by arrows. Scale bar: 10 μm.

Fig. 2 Identification of NBTS1 within aa 543–577.

A. and **B.** The indicated regions of pS1-1(RBM10)/EGFP-N1 were deleted to yield constructs 1–4 [**C** (a)] and constructs 5–8 [**C** (b)], and expressed in ARL cells. Green, EGFP; red, RBM10 stained with anti-S1-1 antiserum. Some S1-1 NBs are indicated by arrows. Scale bar: 10 μm. **C.** Deletion diagrams. Lines in (a) and (b) indicate the deleted regions in A and B,

with the corresponding residue numbers provided at each end. Green lines: constructs whose products localized weakly to S1-1 NBs. (c) Summary of A and B. NBTS1 was located within aa 543–577 (red line). The box shows the location of the KEKE motif.

Fig. 3 Primary sequences of the KEKE, C_2H_2 ZnF, and OCRE regions.

(1) KEKE region. Alternating positive and negative aa residues in the KEKE region are indicated by boldface. "Amphi α-helix" is a region predicted to form a perfect AAH with five helical turns (Inoue et al., 1996). (2) C_2H_2 ZnF. The cysteine and histidine residues critical for ZnF function are highlighted by boldface. (3) OCRE (octamer repeat, also called NLS3). Brackets 1–5 represent the five repeating octamers (8 aa) originally assigned based on hydrophilicity/hydrophobicity profiles (Inoue et al., 1996). The aa 486–540 region (aa 564–618 of the 930-aa isoform) forms a globular domain with an anti-parallel β-sheet fold (Martin et al., 2016).

Fig. 4 Identification of NBTS2.

A. NB1 and NB2. The RBM10 sequences encoding aa 476–645 and 476–760 were ligated into pEGFP-C3 and expressed in ARL cells. Green bars in the upper panel represent NBTS1 at aa 543–577 (the KEKE region). The blue bar represents the region expected to contain

another NBTS within aa 681–742. Purple bars indicate NLS3 (aa 482–540) and NLS1 (principal NLS, aa 743–759). Lower panels show expression of pNB1/EGFP-C3 and pNB2/EGFP-C3. Green, EGFP; red, endogenous RBM10 stained with anti-S1-1 antiserum. Representative S1-1 NBs are indicated by arrows. **B.** Determination of NBTS2. The region corresponding to aa 644–670 (control) (1), aa 685–703 (C_2H_2ZnF) (2), or aa 711–742 (after-ZnF) (3) was deleted from pNB2/EGFP-C3 and these constructs were expressed in ARL cells. Green, EGFP; red, RBM10 stained with anti-S1-1 antiserum. The lower diagram indicates the location of NBTS1 (green bar), the deleted regions (1–3), and the location of the C_2H_2 ZnF (box, aa 681–706). **C.** Confirmation of NBTS2. Two cysteine residues and one histidine residue in the C_2H_2 ZnF of NB2/EGFP-C3 were substituted with alanine and serine, respectively. Green, EGFP. Scale bar: 10 μm.

Fig. 5 Verification of NBTSs.

A. NB localization of NB2 and its mutants. (1) The KEKE region (aa 524–598), the C_2H_2 ZnF region (aa 685–703), both the KEKE and C_2H_2 ZnF regions, or the after-ZnF region (aa 711–742) were deleted from pNB2/EGFP-C3 and these constructs were expressed in ARL cells. Green: EGFP. Some S1-1 NBs and nucleoli are indicated by arrows and arrowheads, respectively. (2) Analysis by ImageJ Plot Profile. The nuclear distributions of the mutants

were analyzed along the indicated lines, which traversed the extra-nucleolar nucleoplasm and S1-1 NBs. (3) Number of S1-1 NBs. NBs larger than 0.5 μm were counted by two individuals independently using enlarged images of cells (Fig. S2), and the numbers were averaged. **B.** Verification of the NB-targeting activity of the KEKE region and C_2H_2 ZnF. Sequences encoding the KEKE region (aa 529–590), the C_2H_2 ZnF (aa 676–711), and the control OCRE region (aa 475–533) were PCR-amplified, ligated to the FLAG-tag in p3xFLAG-Myc-CMV-24, and expressed in ARL cells. Cells were stained with an antibody against FLAG (green), and nuclei were counterstained with 4´,6-diamidino-2-phenylindole (blue). Scale bar: 10 μm.

Fig. 6 NBTS2/C_2H_2 ZnF is essential for regulation of AS by RBM10.

RBM10-knockout HepG2 cells were transduced with lentivirus carrying wild-type RBM10 or RBM10 in which the KEKKE sequence of NBTS1 was mutated to AAKAA or the CLLC sequence of NBTS2 was mutated to ALLA. After 3 days, AS of *fas* transcripts was examined for skipping versus inclusion of exon 6 (skipping/inclusion ratio). **A.** mRNA expression levels of RBM10 and its mutants. **B.** AS of *fas* pre-mRNA. Skipping/inclusion ratios of exon 6 were obtained as described by Inoue et al (2014). The results of two separate experiments were averaged, with deviations indicated by vertical lines at the top of the bars. (–) indicates RBM10-knockout HepG2 cells. WT, AAKAA, and ALLA indicate RBM10-knockout HepG2

cells transduced with wild-type RBM10 or RBM10 with mutation of AAKAA or ALLA, respectively. Two independent experiments in B yielded similar results, as indicated by the small deviations of the averaged exon 6-skipping/inclusion values. However, RBM10 mRNA expression levels in one experiment were roughly 50% lower than those in A (data not shown), suggesting that the expressed construct proteins in both experiments were at saturation levels in performing AS of *fas* pre-mRNA.

Fig. 7 RBM10 is not essential for the structural organization of S1-1 NBs.

A. HeLa cells were treated with RBM10-specific siRNA (1) or control AllStar siRNA (2). Immunoblotting of RBM10 and GAPDH is shown. The molecular sizes of the 930- and 852-aa isoforms in kDa are presented on the left. **B.** Nuclei of siRNA-treated HeLa cells were microinjected with Cy3-labeled *ftz* pre-mRNA and cells were incubated with 5 μg/mL Act D for 2 h. Fluorescence images show RBM10 labeled with anti-S1-1 antiserum (left panels) and Cy3-labeled *ftz* pre-mRNA (right panels). Arrows indicate representative enlarged S1-1 NBs.

Fig. 8 RBM10 is transiently sequestered in S1-1 NBs when transcription is suppressed.

A. (1) Effects of Act D on S1-1 NBs. ARL cells were incubated with Act D (5.0 μg/mL) for the indicated durations and stained with anti-S1-1 antiserum. Arrows indicate some S1-1

NBs. (2) Formation of enlarged S1-1 NBs upon suppression of transcription with various inhibitors. ARL cells were incubated with Act D, α-amanitin, flavopiridol, DRB, or LY294002 for 3 h. Enlarged S1-1 NBs in cells treated with α-amanitin and DRB reverted back to their initial state upon further incubation (chasing) in fresh media for 4 and 1 h, respectively. Cells were stained with anti-S1-1 antiserum. Bar: 10 μm. **B.** (1) Transcriptional activity in DRB-treated ARL cells under the conditions described in A (2). Western blotting of RBM10 in DRB-treated cells. (a) and (b) indicate the 930- and 852-aa isoforms of RBM10, respectively.

Fig. 1

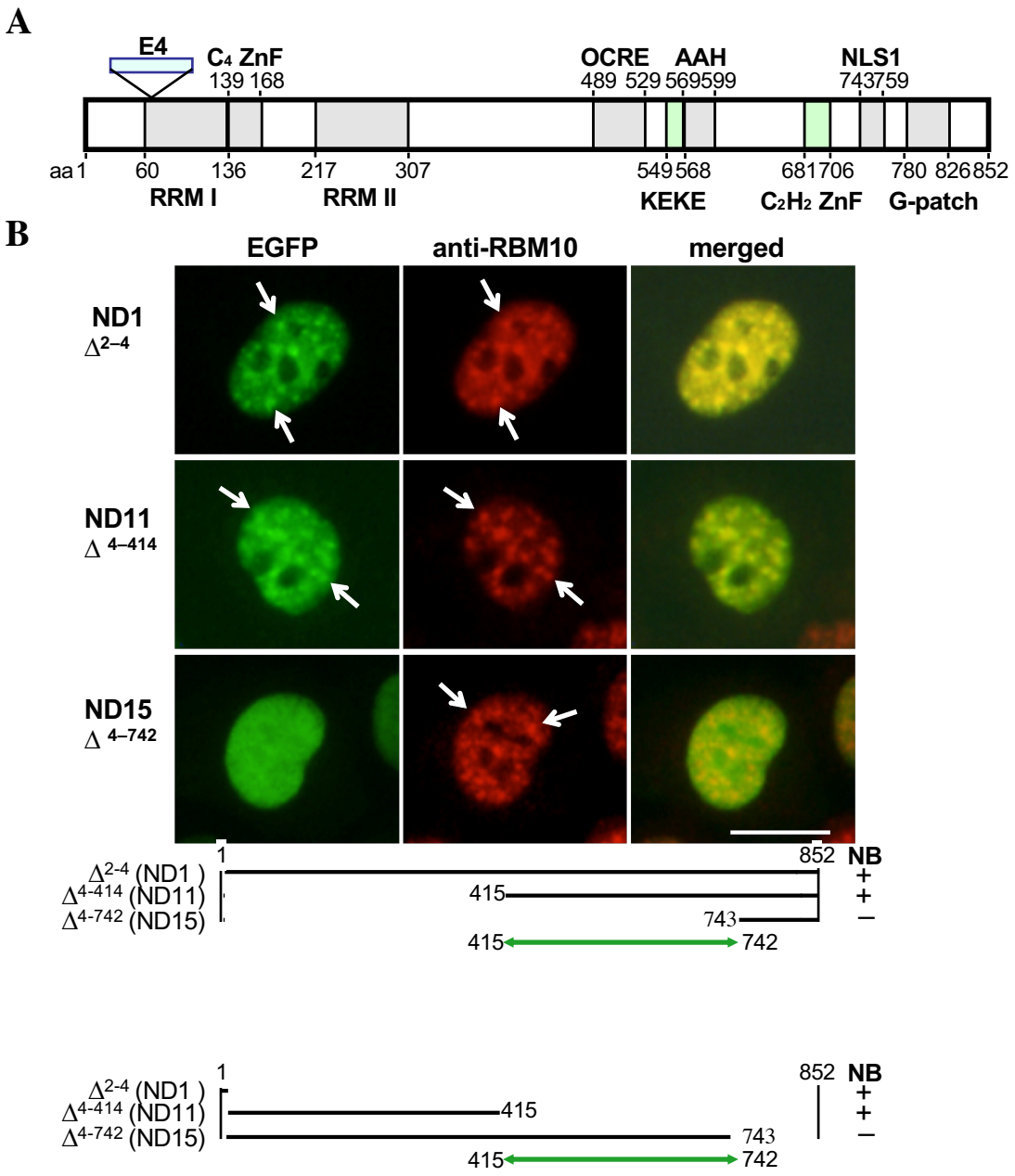

A
E4
C4 ZnF
139 168
OCRE
489 529
AAH
569599
NLS1
743759
aa 1
60
136
217
307
549 568
681 706
780 826 852
RRM I
RRM II
KEKE
C2H2 ZnF
G-patch
B
EGFP
anti-RBM10
merged
ND1
Δ2–4
ND11
Δ 4–414
ND15
Δ 4–742
1
852
NB
Δ2-4 (ND1)
Δ4-414 (ND11)
Δ4-742 (ND15)
415
743
415
742
+
+
–

Fig. 2

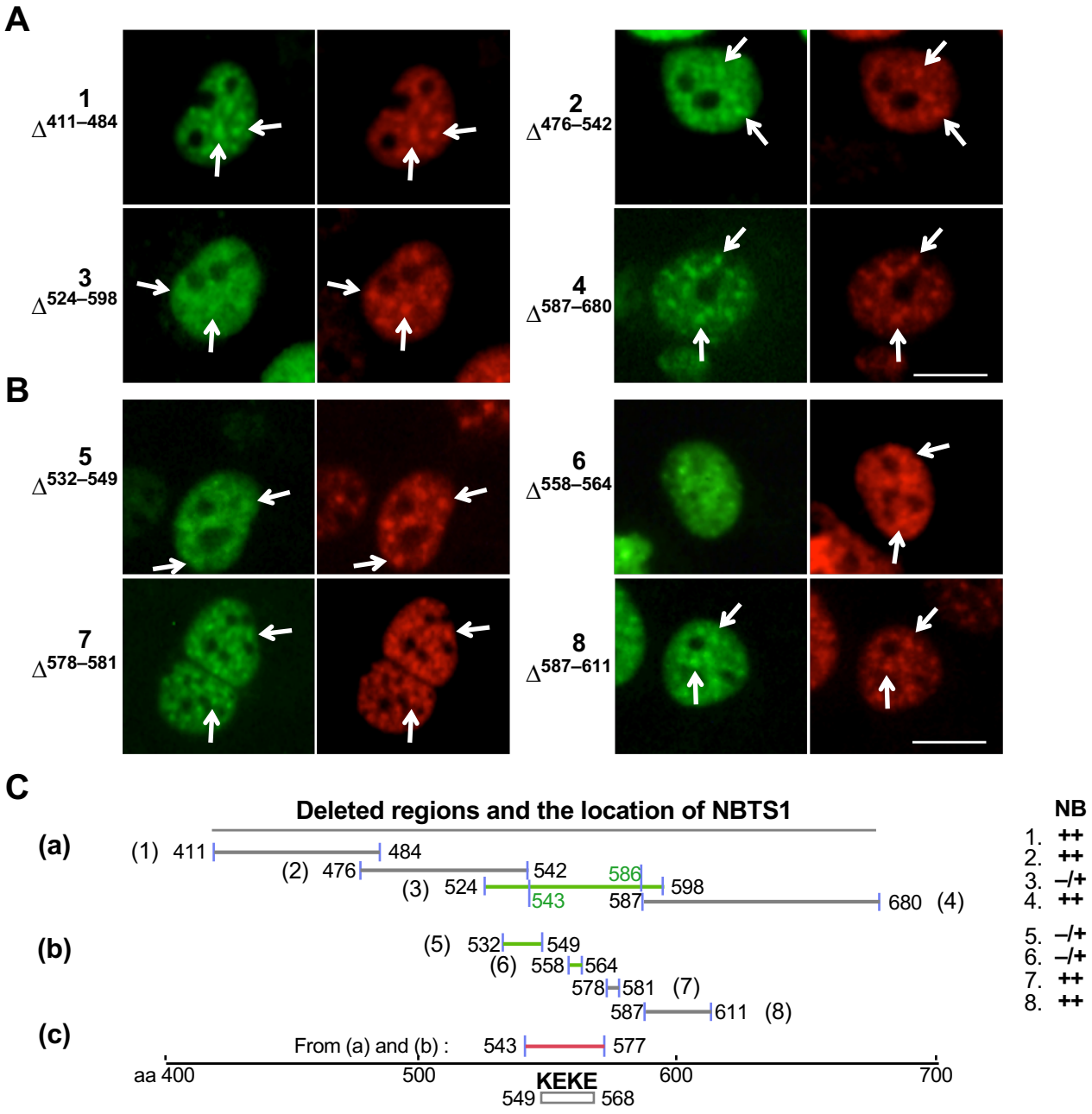

A
1
Δ411–484
2
Δ476–542
3
Δ524–598
4
Δ587–680
B
5
Δ532–549
6
Δ558–564
7
Δ578–581
8
Δ587–611
C
Deleted regions and the location of NBTS1
NB
(a)
(1) 411 484
(2) 476 542
(3) 524 586
543 587 598
680 (4)
1. ++
2. ++
3. –/+
4. ++
(b)
(5) 532 549
(6) 558 564
578 581 (7)
587 611 (8)
5. –/+
6. –/+
7. ++
8. ++
(c)
From (a) and (b) : 543 577
aa 400
500
KEKE
549 568
600
700

Fig. 3

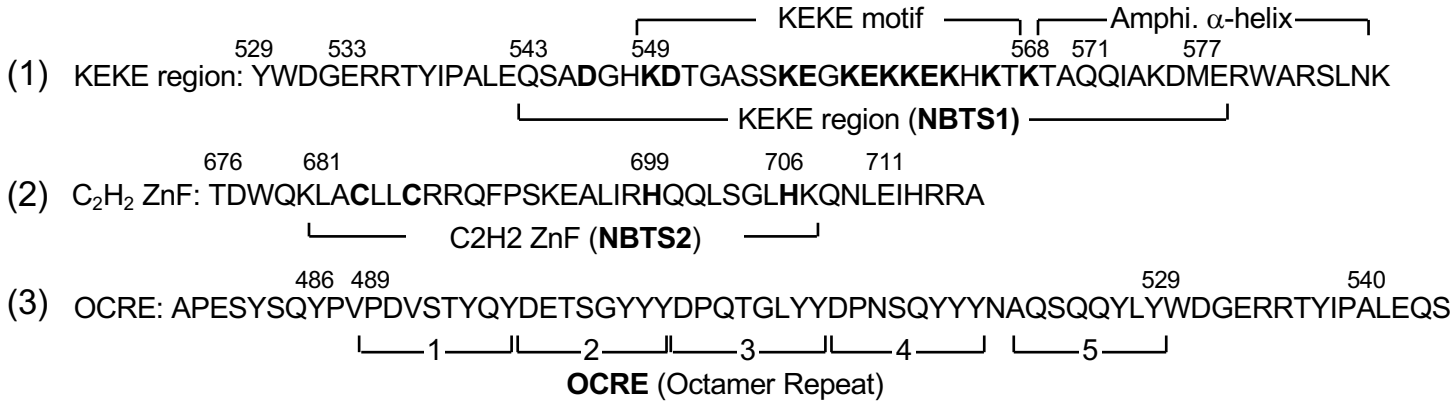

KEKE motif
Amphi. α-helix
529
533
543
549
568
571
577
(1) KEKE region: YWDGERRTYIPALEQSADGHKDTGASSKEGKEKKEKHKTKTAQQIAKDMERWARSLNK
KEKE region (NBTS1)
676
681
699
706
711
(2) C2H2 ZnF: TDWQKLACLLCRRQFPSKEALIRHQQLSGLHKQNLEIHRRA
C2H2 ZnF (NBTS2)
486
489
529
540
(3) OCRE: APESYSQYPVPDVSTYQYDETSGYYYDPQTGLYYDPNSQYYYNAQSQQYLYWDGERRTYIPALEQS
1
2
3
4
5
OCRE (Octamer Repeat)

Fig. 4

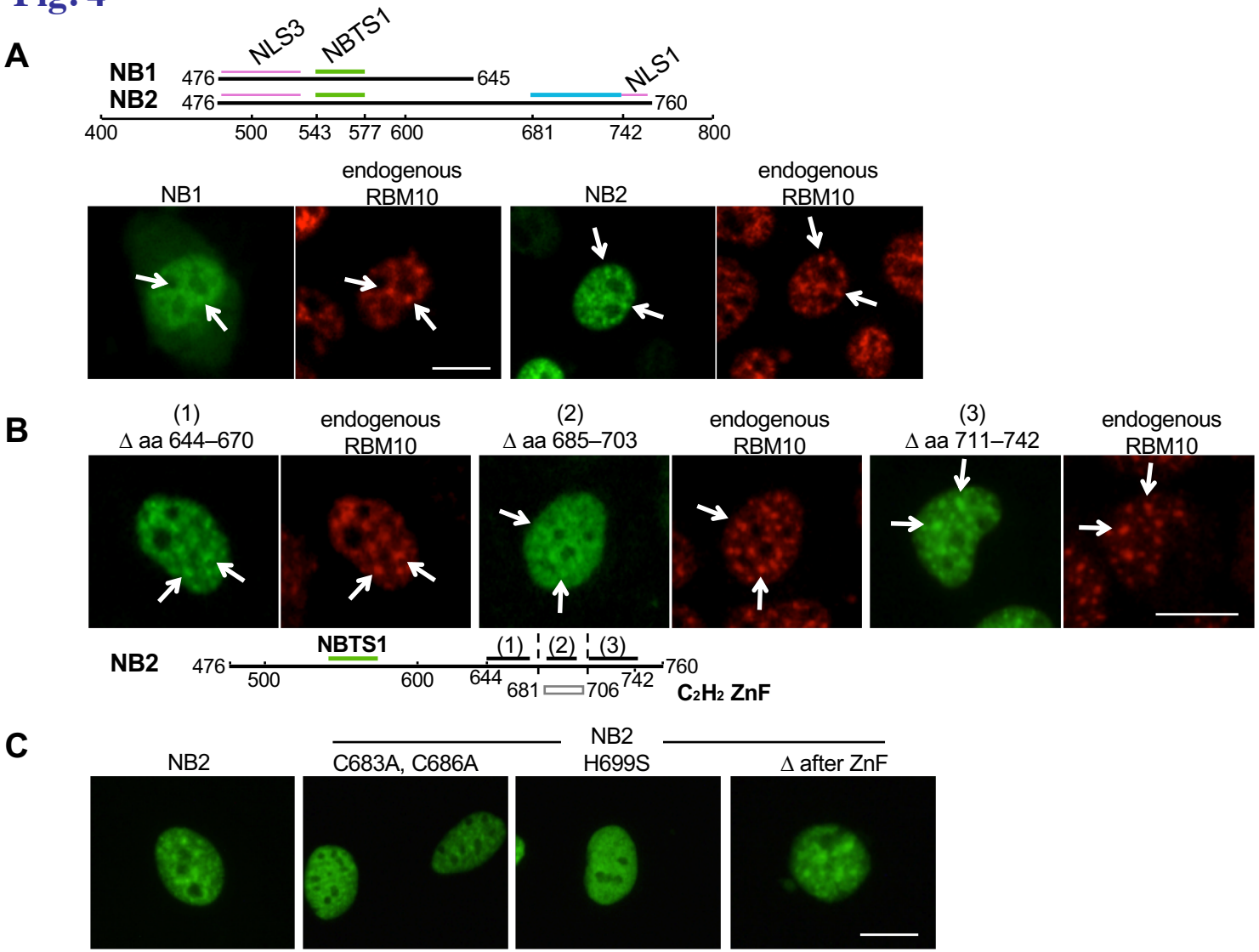

A
NLS3
NBTS1
NLS1
NB1
NB2
476
645
760
400
500
543
577
600
681
742
800
NB1
endogenous RBM10
NB2
endogenous RBM10
B
(1)
Δ aa 644–670
endogenous RBM10
(2)
Δ aa 685–703
endogenous RBM10
(3)
Δ aa 711–742
endogenous RBM10
NB2
NBTS1
(1)
(2)
(3)
476
500
600
644
681
706
742
760
C2H2 ZnF
C
NB2
NB2
C683A, C686A
H699S
Δ after ZnF

Fig. 5

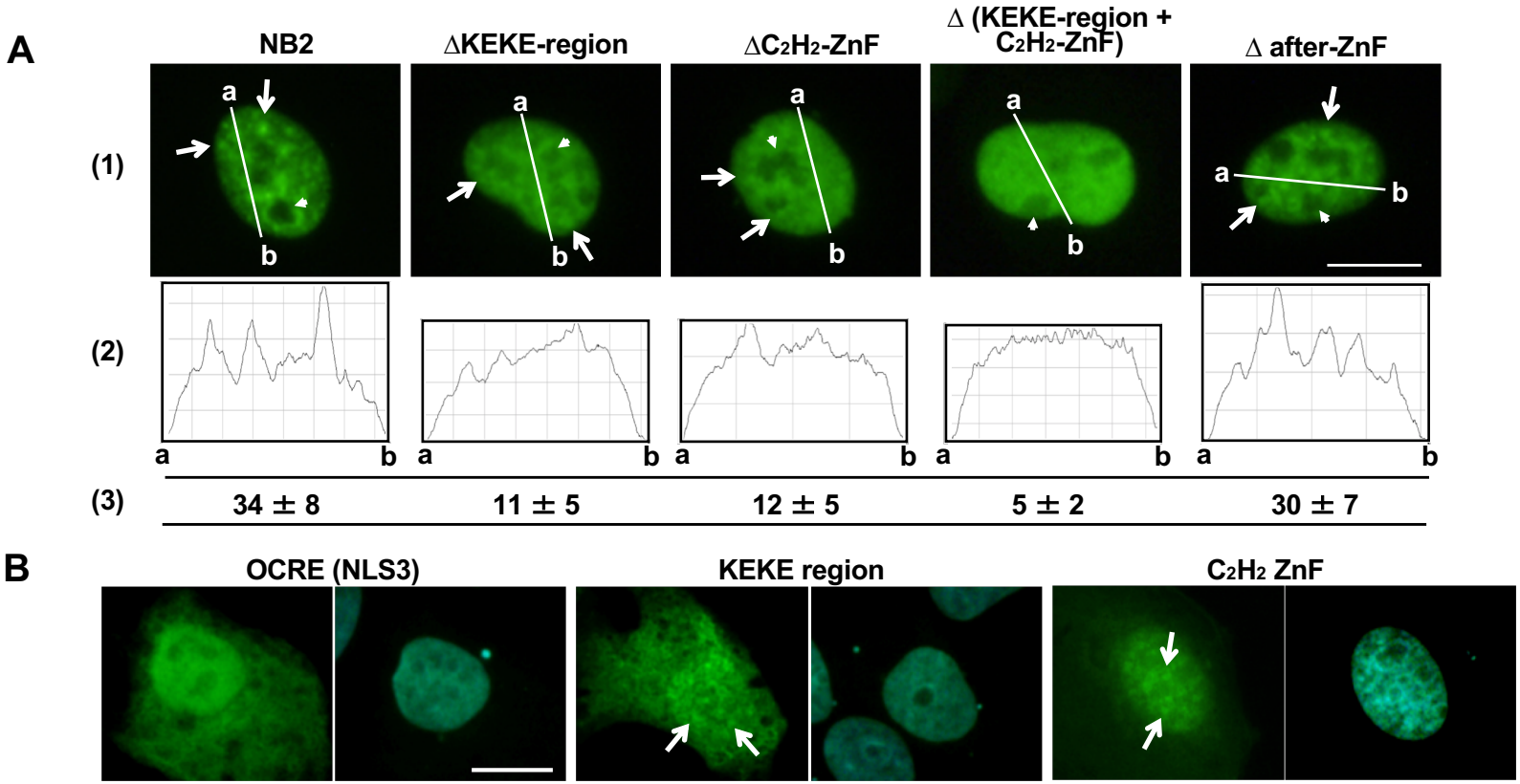

A
NB2
ΔKEKE-region
ΔC2H2-ZnF
Δ (KEKE-region + C2H2-ZnF)
Δ after-ZnF
(1)
(2)
(3)
a
b
34 ± 8
11 ± 5
12 ± 5
5 ± 2
30 ± 7
B
OCRE (NLS3)
KEKE region
C2H2 ZnF

Fig. 6

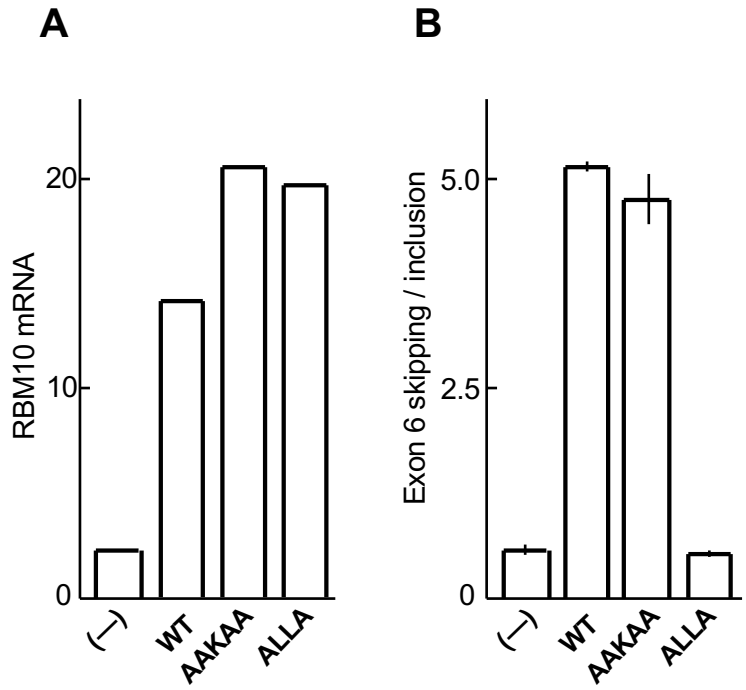

A
B
RBM10 mRNA
20
10
0
Exon 6 skipping / inclusion
5.0
2.5
0
(—)
WT
AAKAA
ALLA

Fig. 7

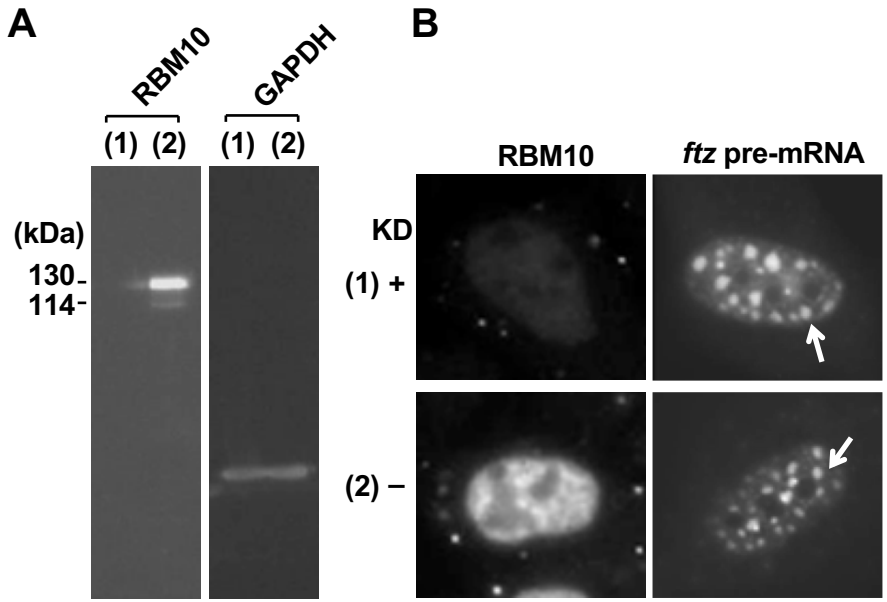
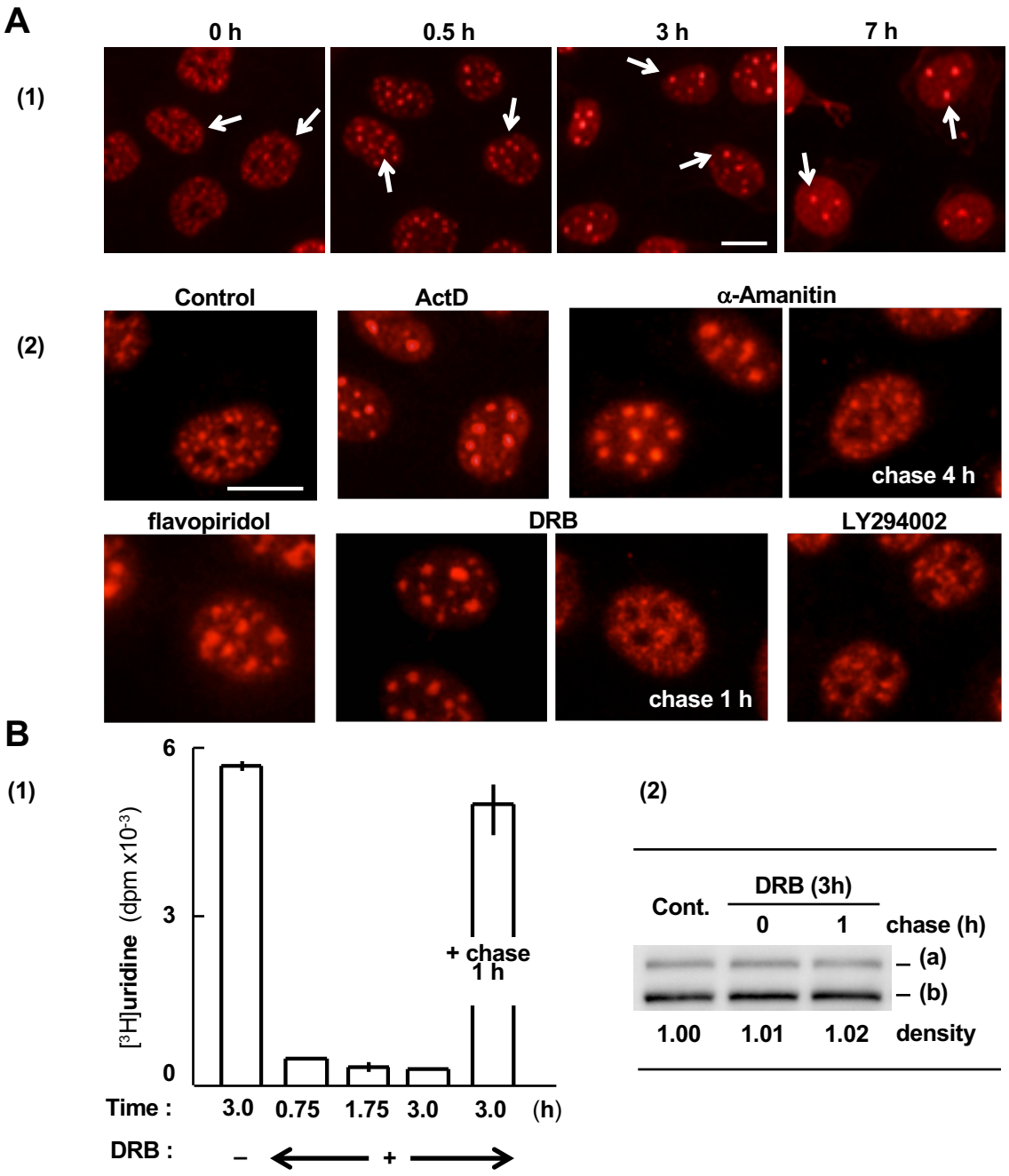


Fig. 8


A
(1)
0 h
0.5 h
3 h
7 h
(2)
Control
ActD
α-Amanitin
chase 4 h
flavopiridol
DRB
chase 1 h
LY294002
B
(1)
6
3
0
[3H]uridine (dpm x10-3)
+ chase 1 h
Time : 3.0 0.75 1.75 3.0 3.0 (h)
DRB : – + –
(2)
Cont.
DRB (3h)
0
1
chase (h)
– (a)
– (b)
1.00 1.01 1.02 density