RNA was quantified by SYBR green-based Quantitative PCR (qPCR) using the primers listed in Supplementary Table S3. The procedure is described in the Supplementary 'Materials and Methods' section.

RESULTS

MBNL and CELF proteins regulate the inclusion of *Clcn1* exon 7A

To examine whether the MBNL and CELF family proteins can regulate the splicing of *Clcn1*, we created a minigene covering exons 6 to 7 of the mouse *Clcn1* gene (Figure 1A). It is important to note that because the inclusion of exon 7A does not produce a premature termination codon in the context of our *Clcn1* minigene, the spliced products containing exon 7A are not substrates of NMD. Thus, the minigene would provide more faithful splicing patterns compared to the endogenous

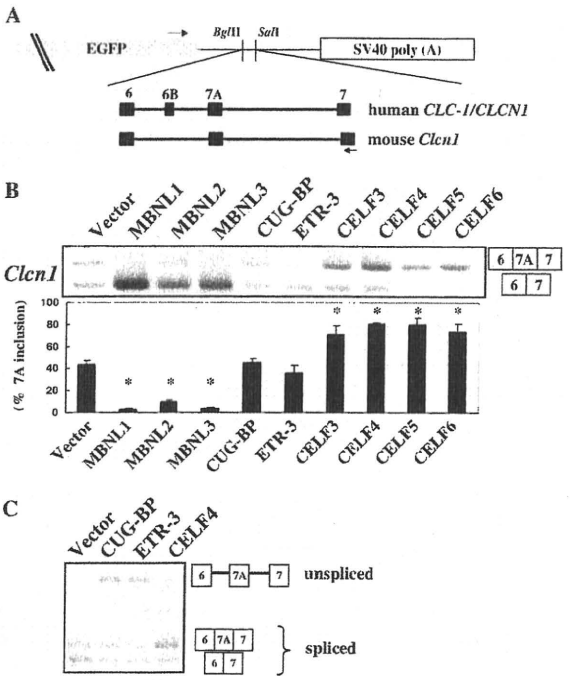

Figure 1. Splicing regulation of MBNL and CELF proteins. (**A**) Structure of chloride channel minigenes. Both human *CLCN1/CIC-1* and mouse *Clcn1* minigenes were subcloned between the BglII and SalI sites of pEGFP-C1. Black boxes represent exons of the minigenes. Arrows indicate the position of primers used in the splicing assays. Exon 6B is a human-specific exon and is absent in *Clcn1*. (**B**) Splicing regulation of *Clcn1* by MBNL and CELF proteins. Representative results of cellular splicing assays using the *Clcn1* minigene in COS-7 cells. The upper bands correspond to a splice product containing exon 7A, whereas lower bands correspond to a splice product lacking exon 7A. Bar chart shows quantified results of exon 7A inclusion (mean ± SD, $n = 3$). Statistical significance was analyzed by analysis of variance (ANOVA) and Dunnett's multiple comparisons. All MBNL proteins and CELF proteins except for CUG-BP and ETR-3 showed significant differences ($^*P < 0.0001$) compared to the empty vector. (**C**) CUG-BP and ETR-3 increased an unspliced product of the *Clcn1* minigene. Structures of PCR products are indicated.

Clcn1. We utilized non-muscle cell lines to minimize the effect of muscle-dependent backgrounds and focus on the direct effects of transgenes. When the *Clcn1* minigene was transfected into COS-7 cells, 45% of the spliced products contained exon 7A (Figure 1B). Next we expressed myc-tagged MBNL or CELF proteins with the *Clcn1* minigene and examined the patterns of *Clcn1* splicing. The expressions of MBNL and CELF proteins were confirmed by Western blot analysis using an anti-myc antibody (Supplementary Figure S6A). All three MBNL proteins strongly repressed exon 7A inclusion (Figure 1B). In contrast, CELF3, CELF4, CELF5 and CELF6 proteins significantly promoted the inclusion of 7A (Figure 1B). Remarkably, CUG-BP and ETR-3 did not alter the ratio of exon 7A inclusion (Figure 1B). These two proteins increased the unspliced product and reduced the spliced products with or without exon 7A (Figure 1C).

CELF4 and expanded CUG repeats act antagonistically against MBNL1 in the splicing regulation of *Clcn1*

We investigated whether CELF proteins can antagonize the effect of MBNL1. Among CELF proteins that promoted exon 7A inclusion in *Clcn1*, CELF4 was used for the following analyses because this protein is expressed in muscle (26,43). CELF4 and two other muscle-expressed CELF proteins, CUG-BP and ETR-3, were co-transfected with MBNL1, as well as the *Clcn1* minigene. As shown in Figure 2A, CELF4 increased the ratio of exon 7A in a dose-dependent manner (lanes 7 and 8), whereas CUG-BP or ETR-3 did not (lanes 3–6). When MBNL1 was titrated in the presence of CELF4, a decrease in exon 7A inclusion was observed, depending on the dosage of MBNL1 (Figure 2A, lanes 9–11). These results demonstrate that MBNL1 and CELF4, but not CUG-BP or ETR-3, can regulate *Clcn1* exon 7A splicing in an antagonistic manner.

We also examined DMPK constructs harboring either CTG18 (DM18) or interrupted CTG480 (DM480) in the 3′-UTR. When these constructs were expressed with the *Clcn1* minigene, DM480 increased exon 7A inclusion, whereas DM18 showed little effect (Figure 2B, left). These constructs were next expressed in the presence of MBNL1. As shown previously, MBNL1 alone strongly repressed the inclusion of exon 7A. DM480 reversed the repression of exon 7A inclusion by MBNL1, whereas DM18 did not (Figure 2B, right). Thus, expanded CUG repeats can antagonize the effect of MBNL1 on *Clcn1* splicing.

Finally, we performed knockdown experiments using vector-based RNAi to examine whether the regulation of *Clcn1* splicing is dependent on the dose of endogenous MBNL1 or CUG-BP. The RNAi vector used here expresses an artificial microRNA that is exactly complementary to a region of its target gene, leading to degradation of the target mRNA. For RNAi experiments, we utilized Neuro2a and HeLa cells rather than COS-7 because of the availability of cDNA sequence information essential for the design of microRNA. First, the overexpression of MBNL1 in Neuro2a cells strongly repressed the inclusion of exon 7A as in COS-7 cells (M, Figure 2C).

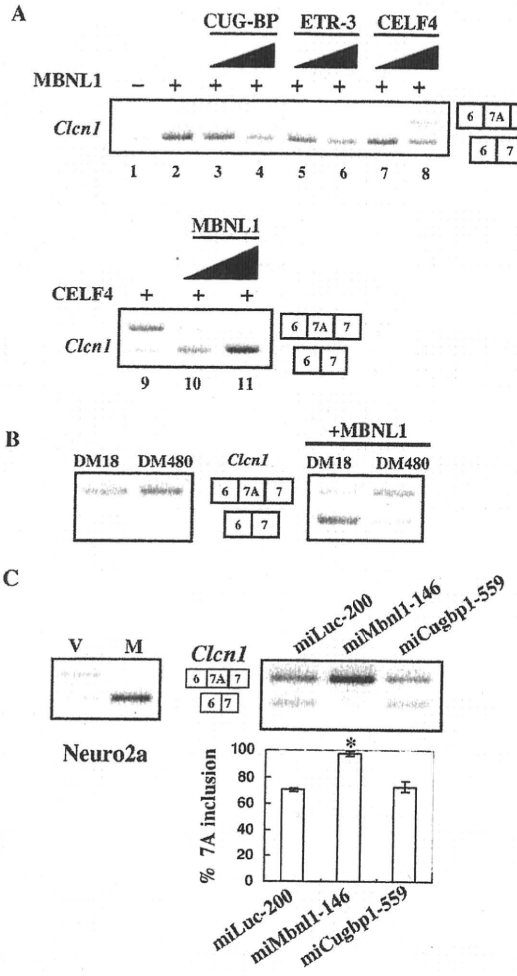

Figure 2. MBNL1 is antagonized by CELF4 and expanded CUG repeats. (**A**) Antagonistic effects of CELF4 against MBNL1 in the splicing regulation of *Clcn1*. *Clcn1* minigene was co-transfected with MBNL1 or MBNL1 plus CELF protein into COS-7 cells. Upper panel: lane 1, the pattern of empty vector-transfected cells; lanes 2–8, a constant amount of MBNL1-encoding vector and increasing amounts of one of the CELF-encoding vectors were co-transfected at ratios of 1:0 (lane 2), 1:1 (lanes 3, 5 and 7) and 1:3 (lanes 4, 6 and 8). Lower panel: similarly, a CELF4-encoding vector and MBNL1-encoding vector were transfected at ratios of 1:0 (lane 9), 1:1 (lane 10) and 1:3 (lane 11). The total amount of transfected plasmids was held constant by adjusting the amounts of empty vector. (**B**) Effects of an expanded CUG repeat on the chloride channel splicing. *Clcn1* minigene was transfected with an expression vector of DMPK harboring a normal (DM18) or expanded (DM480) CTG repeat (left panel). Either normal or expanded DMPK vector was co-transfected with MBNL1 and the *Clcn1* minigene, and the splicing patterns were analyzed (right panel). (**C**) Results of *Clcn1*-splicing assays in Neuro2a cells. MBNL1-mediated repression of exon 7A inclusion in Neuro2a cells (left). 'V' and 'M' indicate empty vector and MBNL1, respectively. Note that the basal inclusion of exon 7A in Neuro2a cells was higher than that in COS-7 cells. *Clcn1*-splicing regulation was affected by RNAi of Mbnl1 but not Cugbp1 in Neuro2a cells (right). Bar chart shows the quantified results of exon 7A inclusion (mean ±SD, $n = 4$). According to ANOVA and Dunnett's test, miMbnl1-146 induced a statistically significant increase of exon 7A compared to miLuc-200 ($^*P < 0.0001$), whereas miCugbp1-559 did not ($P = 0.36$).

An RNAi vector miMbnl1-146 was effective in reducing both Mbnl1 and Mbnl2 and increased the inclusion of *Clcn1* exon 7A compared to a control vector miLuc-200 (Figure 2C and Supplementary Figure S1). Another RNAi vector, miMbnl1-236, also increased exon 7A inclusion (data not shown). We also used HeLa cells for RNAi experiments and found that the knockdown of MBNL1 resulted in an increase in exon 7A inclusion (Supplementary Figure S2). Therefore, the splicing of *Clcn1* exon 7A is regulated by MBNL in a dose-dependent manner. In contrast, an effective RNAi vector targeting endogenous Cugbp1, miCugbp1-559, did not alter exon 7A splicing (Figure 2C and Supplementary Figure S1). We were not able to confirm the knockdown of endogenous Etr-3/Cugbp2 and Celf4 because of very low or undetectable endogenous expression in the cell lines used in this study (data not shown); thus, the RNAi of Etr-3 and Celf4 was not tested in the splicing analysis.

Involvement of the exon 7A sequence in the splicing regulation by MBNL1

To understand the mechanism by which MBNL1 represses the inclusion of *Clcn1* exon 7A, we tried to define the regions of the *Clcn1* minigene that are responsive to MBNL1. For this purpose, we examined a series of *Clcn1*deletion mutants lacking a region in either introns 6 or 7A (Δ1–Δ9, Figure 3A). As shown in Figure 3B, deletions in intron 6 altered the basal splicing pattern. For example, the Δ1 mutant exhibited increased inclusion of exon 7A, suggesting the presence of an element in the deleted region that represses exon 7A inclusion. However, this mutant and the other intron 6 mutants (Δ2–Δ5) were strongly repressed by MBNL1 (Figure 3B, upper panel). Similarly, the series of intron 7A mutants did not lose responsiveness to MBNL1 (Δ6–Δ9, Figure 3B). Thus, large portions of these intronic regions could be excluded from the region critical for responsiveness to MBNL1.

Next we examined another series of deletion mutants (Figure 4A, left). When minigene 6-7A was expressed with the empty vector, both spliced and unspliced products were observed (Figure 4A). Co-expression with MBNL1 resulted in the repression of splicing (Figure 4A). In the case of 7A-7, the expression of MBNL1 did not cause a significant change compared to that of the empty vector (Figure 4A). To examine whether the responsiveness of 6-7A was dependent on the sequence of exon 7A, we analyzed the 6/7 mutant, in which the sequence of exon 7A was virtually replaced with that of exon 7. Interestingly, this replacement completely abolished the responsiveness to MBNL1 (Figure 4A, right panels). The observed loss of response in the 6/7 mutant was not due to inefficient basal splicing of exon 7, because replacement of the first 12 nt of exon 7 in the 6/7 mutant with an exonic-splicing enhancer (ESE) markedly increased the spliced products but did not improve the responsiveness to MBNL1 (Supplementary Figure S3). Thus, exon 7A should contain at least part of the MBNL1-responsive region. In addition, intron 6 alone was insufficient for response to MBNL1. To further

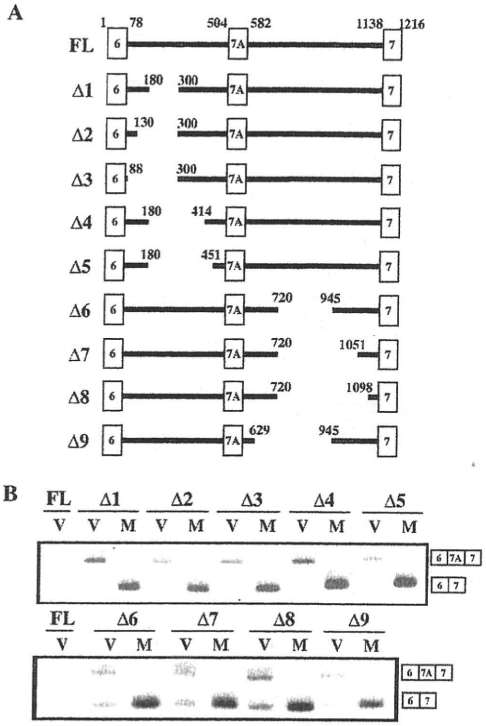

Figure 3. Intronic deletion analysis of *Clcn1*. (**A**) Structure of *Clcn1* deletion mutants. Deletion series of the *Clcn1* minigene were generated by PCR-mediated mutagenesis. FL corresponds to the full-length *Clcn1* minigene covering exons 6 7 and is the same construct as that used in Figures 1 and 2. The positions of nucleotides at the termini of exons and the junction of deleted regions are indicated. (**B**) Splicing analysis of *Clcn1* deletion mutants in COS-7 cells. Cellular-splicing assays were performed using deletion mutants Δ1–Δ9 and MBNL1. Lanes 'V' and 'M' indicate the co-expression of each deletion minigene with empty vector and MBNL1, respectively. Splicing patterns were detected by RT–PCR, as in Figure 1B.

define the responsive region, we tested another minigene, 6-7AΔ, a 3′ deletion of the 6-7A fragment containing only 30 nt from the 5′ part of exon 7A. This fragment was still responsive to MBNL1 (Figure 4B). Finally, we tested a minigene derived from 7A-7 that had exon 7 as well as the preceding 52 nt replaced by the corresponding region of exon 7A (Figure 4C). This minigene, 7A/(−52)7A, exhibited increased unspliced products upon MBNL1 expression, suggesting that the substituted fragment could mediate the response to MBNL1 in a heterologous context (Figure 4C). These results suggest the involvement of a 5′ portion of exon 7A (30 nt), as well as its upstream 52-nt region in the MBNL1-mediated-splicing regulation, at least in the context of the truncated minigenes.

A region around the 3′ splice site of exon 7A mediates the splicing regulation of MBNL1

To further analyze the regions critical for exon 7A splicing, we generated heterologous minigenes in which the exon of interest was inserted in the context of constitutive exons of mouse tropomyosin 2 (*Tpm2*), a gene distinct from *Clcn1*. *Clcn1* fragments covering 414–720 or 451–720, or a reference fragment covering *Tpm2* exon 9 and its flanking intronic regions, were inserted into a *Tpm2* fragment covering exons 1 to 2 (Figure 5A). MBNL1 repressed *Clcn1* exon 7A inclusion of the heterologous minigenes, demonstrating that the inserted fragments of *Clcn1* were sufficient for response to MBNL1 (Figure 5B). In contrast, MBNL1 showed little effect on the inclusion of *Tpm2* exon 9 (Figure 5B). Next we made an additional chimeric minigene derived from the 451–720 minigene by replacing a region derived from intron 7A with a sequence derived from *Tpm2* intron 9 (int6-ex7A-int9; Figure 5A). Although int6-ex7A-int9 was still responsive to MBNL1, another minigene with a further replacement of exon 7A with *Tpm2* exon 9 (int6-ex9-int9) exhibited only a slight response to MBNL1 (Figure 5C). Thus, the exon 7A sequence was essential for regulation by MBNL1. Furthermore, a minigene with an intermediate replacement, 7A(30)-ex9(40), in which the first 30 nt of the exon and its upstream intronic region were derived from *Clcn1*, was responsive to MBNL1 (Figure 5C). These results suggest that a region covering a 3′ portion of intron 6, as well as the first 30 nt of exon 7, can mediate splicing regulation by MBNL1. This is consistent with the results of the truncated minigenes displayed in Figure 4.

We also tested several minigenes to identify the region of *Clcn1* responsive to CELF4. Interestingly, CELF4 did not increase the exon 7A inclusion of the 414–720 construct, which was responsive to MBNL1 (Figure 5B and D). However, minigenes covering 89–720 and 181–720, but not 414–1050, increased the inclusion of exon 7A upon CELF4 overexpression (Figure 5D and data not shown). Therefore, an element responsive to CELF4 was located in the region of 181–413, which was distinct from the regions responsive to MBNL1 identified previously. Lastly, we tested the Δ4 mutant that lacks the putative responsive region for CELF4 (Figure 3A). CELF4 did not promote exon 7A inclusion of Δ4 (Figure 5E). Thus, 181–413 was essential for the regulation of exon 7A by CELF4.

MBNL1 binds directly to the transcript of the *Clcn1* minigene

We next examined whether MBNL1 associates with transcripts of *Clcn1* in cells. In a ribonucleoprotein immunoprecipitation (RIP) analysis (42), intracellular RNA–protein complexes were reversibly fixed with paraformaldehyde and then co-immunoprecipitated. Fixation allowed for stringent washing of immunoprecipitates to exclude post-lysis association between RNA and proteins. We utilized a Neuro2a cell line stably expressing EGFP-fused MBNL1. This cell line exhibited a lower basal inclusion of exon 7A of endogenous *Clcn1* compared to normal Neuro2a cells (Supplementary Figure S5). EGFP-MBNL1 was precipitated by beads pretreated with anti-GFP antibody or rabbit IgG. RNA fragments derived from specific regions of endogenous *Clcn1* in the precipitates were detected by quantitative PCR. The Association of *Clcn1* intron 6 and EGFP-MBNL1 was detected only when immunoprecipitated by anti-GFP antibody (Figure 6A, left). Importantly, this

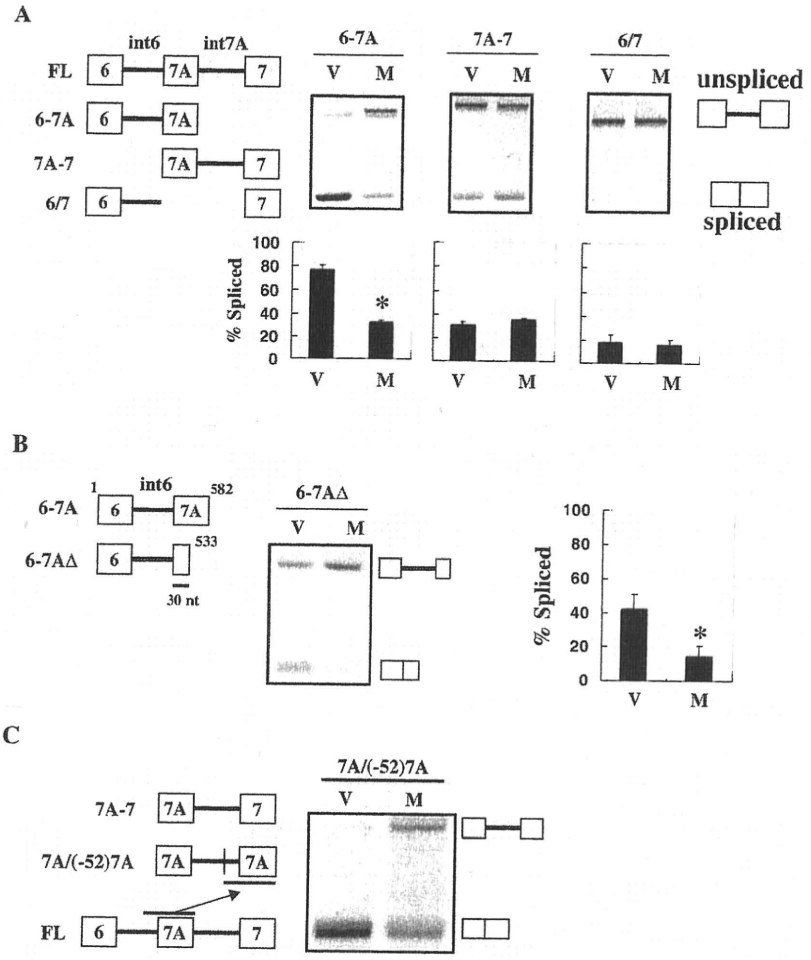

Figure 4. Involvement of exon 7A sequence in splicing regulation by MBNL1. (**A**) Response of truncated *Clcn1* minigenes to MBNL1 is dependent on exon 7A. Structures of 6-7A, 7A-7 and 6/7 deletion mutants are indicated in the upper left. Splicing analysis of these mutants shows both spliced and unspliced products as indicated. The bar chart shows the quantified ratio of spliced products. (**B**) Deletion in the 3′ region of exon 7A did not abolish the response to MBNL1. The structure of the 6-7AΔ mutant is shown (left). Splicing products of 6-7AΔ and quantification are shown as in (A). (**C**) Splicing regulation of the 7A/(−52)7A minigene. 7A/(−52)7A was made from 7A-7 by replacing exon 7 and its upstream 52 nt with the corresponding region of exon 7A (left). Splicing of 7A/(−52)7A was responsive to MBNL1 (right). Lanes 'V' and 'M' indicate co-transfection with the empty vector and MBNL1, respectively. Bars represent the ratio of spliced bands. Statistical analysis was performed by two-tailed *t*-test in comparison with the empty vector. $^{*}P < 0.05$ from three independent experiments.

association was specific because intron 6 was highly enriched compared to other *Clcn1* regions (Figure 6A, right).

In the results described previously, a region around the boundary of intron 6 and exon 7A was suggested to play an important role in the response to MBNL1 (Figures 4 and 5C). Previously, it was reported that MBNL1 prefers mismatch-containing double-stranded structures (22,44,45). It is interesting that a region around the 3′ splice site of exon 7A (corresponding to 473–518) was predicted to form a hairpin structure by Mfold (http://www.bioinfo.rpi.edu/applications/mfold; Figure 6B; 46). We examined whether MBNL1 can bind to this fragment, Clcn1(473–518), by gel shift analysis. GST-fused MBNL1, but not GST alone, bound to the *Clcn1* fragment in a dose-dependent manner (Figure 6C). GST-MBNL1 exhibited an apparently lower affinity for Clcn1(GAA), a mutant RNA in which the first 12 nt of exon 7A were substituted by a $(GAA)_4$ repeat (Figure 6C). Therefore, the sequence at the 5′ end of exon 7A contributes to the binding of MBNL1.

The 5′ end of exon 7A contains an ESE

Region 451–533 of *Clcn1* consistently accounted for the responsiveness to MBNL1 in multiple contexts (Figures 4B and 5C). We noticed that this region contained a YGCU(U/G)Y motif (Y: C or U; Figure 7A), which has been reported as a binding consensus sequence for MBNL1 (24). A mutant minigene containing a mutation of this motif (but otherwise the same as the full-length minigene) was examined in a splicing assay (Mut 1, Figure 7A). Mut1 was fully responsive to MBNL1 as well as to CELF4 (Figure 7B). Thus, this motif itself

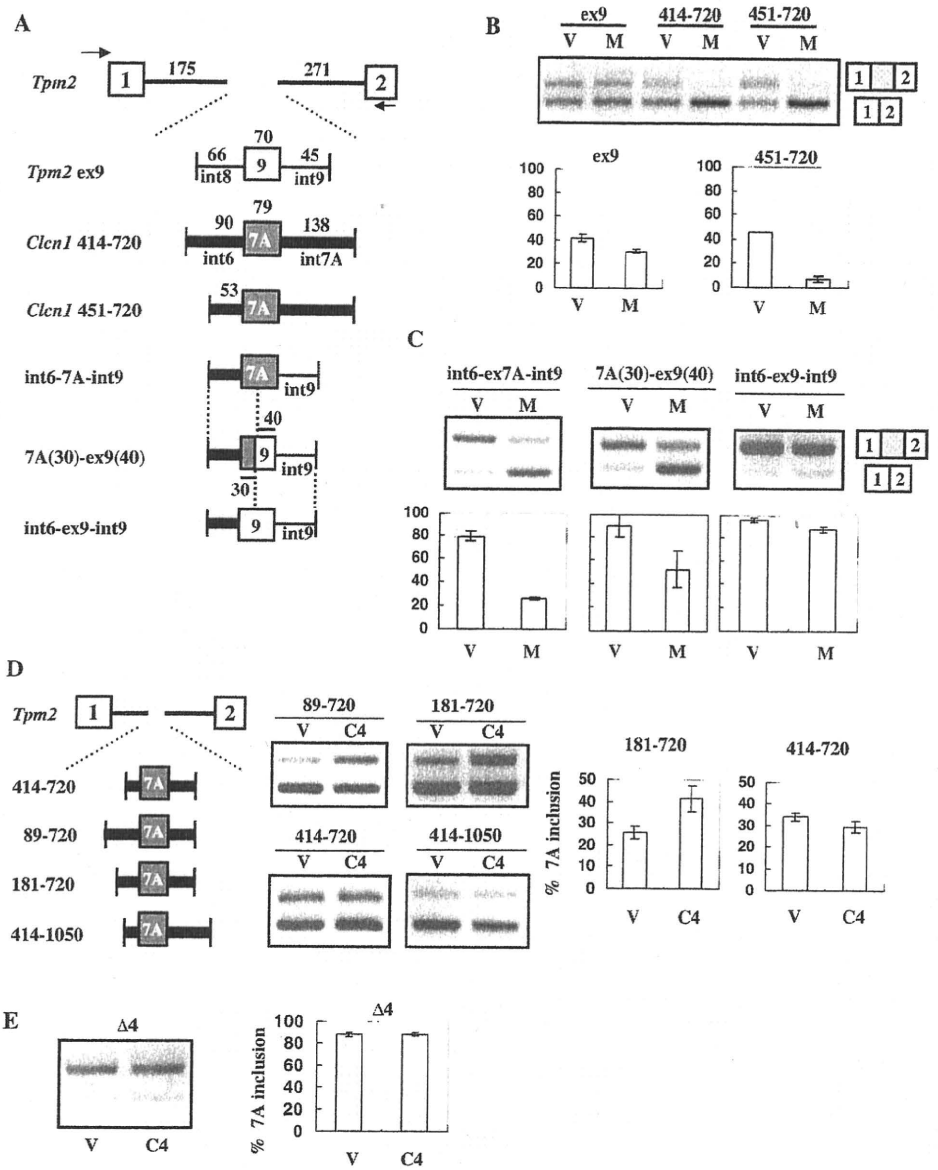

Figure 5. Exon 7A splicing regulation in heterologous minigenes. (**A**) Structure of the *Tpm2*-based heterologous minigene. Fragments of *Tpm2* covering exons 1 to 2 were inserted downstream of EGFP. Test exons together with their flanking regions were inserted into intron 1 of *Tpm2*. Intronic fragments derived from *Clcn1* are indicated by thick lines, whereas those derived from *Tpm2* (regions flanking exon 9) are indicated by thin lines. Exonic sequences of *Clcn1* exon 7A and *Tpm2* exon 9 are indicated by grey and white boxes, respectively. (**B**) Splicing assay results using *Tpm2*-based heterologous minigenes in COS-7 cells. Upper bands correspond to the spliced products containing an exon inserted between *Tpm2* exon 1 and 2. 'V' and 'M' indicate empty vector and MBNL1, respectively. Compared with *Tpm2* ex9, *Clcn1*-derived 414–720 and 451–720 minigenes exhibited evident responses to MBNL1. Bar chart shows quantified results of the splicing assay ($n = 3$). (**C**) Splicing regulation of heterologous minigenes containing a portion of *Clcn1* intron 6. Results of the splicing assay are shown as in B. The structures of minigenes are shown in A. (**D**) Determination of the *Clcn1* region responsible for CELF4-mediated exon 7A inclusion. *Tpm2*-based heterologous minigenes covering a *Clcn1* region indicated by the numbers were tested for their responsiveness to CELF4 (C4). The splicing assay was performed as in B, except that CELF4 was used in place of MBNL1. In the case of 181–720, CELF4 expression induced a significant increase compared to control ($P < 0.05$, $n = 3$, two-tailed t-test). (**E**) Splicing analysis of the Δ4 mutant minigene and CELF4. The structure of Δ4 is described in Figure 3A. Splicing analysis results are shown as in D. CELF4 did not significantly alter the splicing of Δ4 ($P = 0.97$, $n = 3$, two-tailed t-test).

was dispensable for the response to MBNL1 in exon 7A splicing.

The preceding analyses suggested that the 5′ end of exon 7A is involved in the responsiveness to MBNL1 (Figures 4B and 5C). To characterize the role of this region, we first examined a minigene lacking the first 15 nt of exon 7A (Mut 2, Figure 7A). Unexpectedly, this deletion resulted in the complete exclusion of exon 7A, even without MBNL1 overexpression (Figure 7B), indicating the presence of an ESE in this region that is

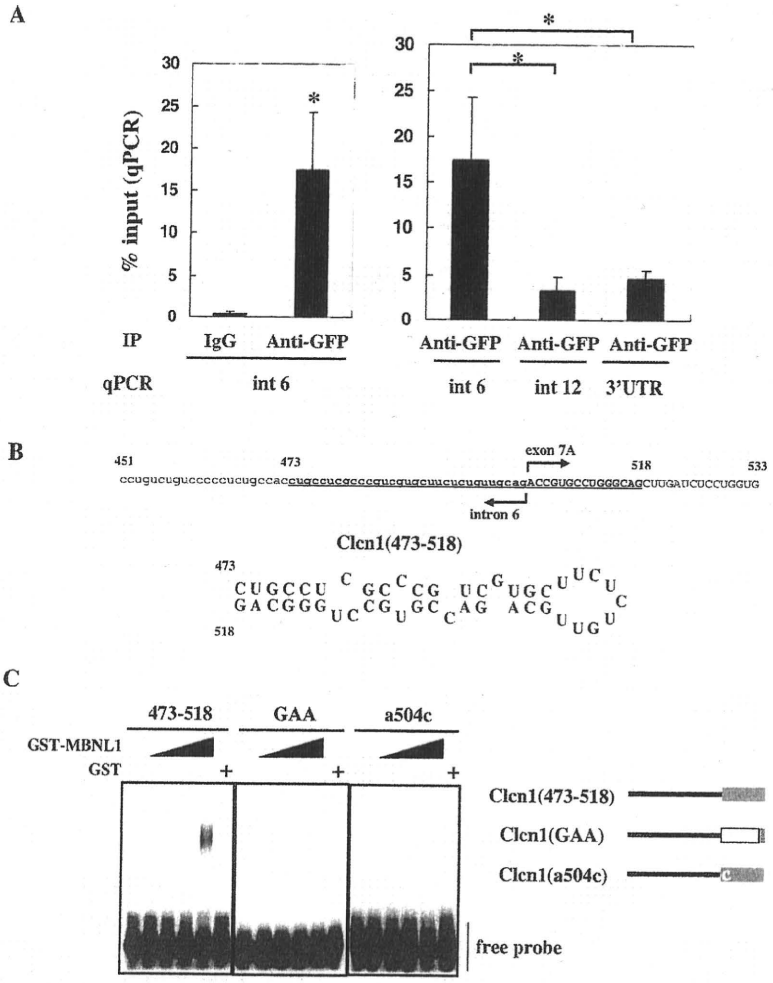

Figure 6. MBNL1 associates with *Clcn1* RNA. (**A**) The intracellular association between EGFP-MBNL1$_{40}$ and the transcript of endogenous *Clcn1* was analyzed by RIP. The cell lysate of a cell line stably expressing EGFP-MBNL1 was used for immunoprecipitation. The RNA fraction co-precipitated with anti-GFP antibody or IgG was reverse-transcribed. The amount of *Clcn1* (pre-)mRNA was quantified by real-time quantitative PCR using a primer set for the indicated regions of the *Clcn1* gene ($n = 5$). Bars represent the amount of *Clcn1* RNA co-precipitated with anti-GFP antibody or control IgG normalized by the amount of the *Clcn1* RNA in the input fraction (mean ± SD). Left: EGFP-MBNL1 was co-immunoprecipitated with RNA fragments containing *Clcn1* intron 6 by GFP antibody but not control IgG ($^*P < 0.01$, two-tailed *t*-test). Right: RIP analysis of multiple *Clcn1* regions. Three regions of *Clcn1* (intron 6, intron 12 and 3′-UTR) were amplified from immunoprecipitates of the GFP antibody. Intron 6 was significantly more enriched than the other regions ($^*P < 0.01$, ANOVA and Tukey's test). (**B**) A hairpin structure predicted in the putative MBNL1-responsive region of *Clcn1*. Nucleotides 473–518 of *Clcn1* are indicated in boldface (upper). Predicted secondary structure of the fragment Clcn1(473–518) (lower). (**C**) Binding between GST-MBNL1 and Clcn1(473–518) or its mutants was examined by gel shift analysis (left). ^{32}P-labeled probes were incubated with or without GST-MBNL1 (0.225, 0.45, 0.9, 1.8 μM) or GST (1.8 μM). The reaction mixture was separated by native PAGE and visualized by autoradiography. Structure of probes used in gel shift analysis (right). In the Clcn1(GAA) mutant, the first 12 nt of exon 7A are substituted by $(GAA)_4$ repeats. The Clcn1(a504c) mutant contains a point mutation of the first nucleotide of exon 7A but is otherwise the same as Clcn1(473 518).

essential for the basal inclusion of exon 7A. We further tested 6-nt deletions in either the 5′ or 3′ parts of the 15-nt region (Mut 3 and Mut 4, Figure 7C). Only Mut 3 exhibited complete exon 7A skipping like Mut 2 (Figure 7C). Direct comparison of the sequences of Mut 3 and Mut 4 showed that 5 nucleotides differed between these mutants (Figure 7D, upper). We then introduced substitutions of these nucleotides into the full-length minigene (Mut 5a and Mut 5b, Figure 7D). Both Mut 5a and 5b exhibited reduced basal inclusion of exon 7A, with a stronger effect from Mut 5a (Figure 7D). Finally, we tested two point mutants, a504c and t512g, to define the region of the ESE. Whereas the t512g mutation had no effect on the basal splicing of exon 7A, mutant a504c totally disrupted the inclusion of exon 7A (Figure 7D). Because Mut 5b (c511a and t512g), but not t512g alone, showed a reduction in exon 7A inclusion, mutation c511a should be critical for exon 7A inclusion. GST-MBNL1 bound to an RNA containing the a504c mutation with a lower affinity compared to Clcn1(473–518) (Figure 6C), suggesting that the strong repressive effect of a504c was due to the direct disruption of the ESE rather than a

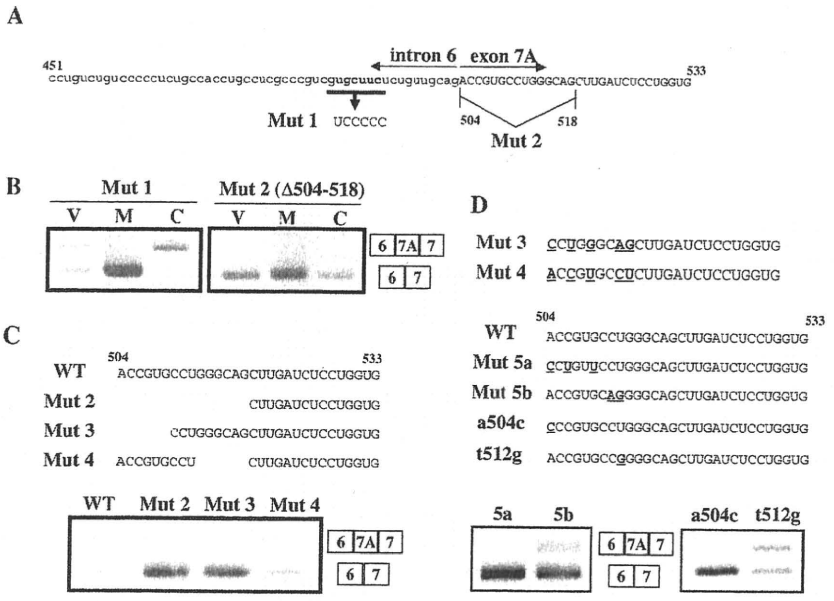

Figure 7. Analysis of the 5′ region of exon 7A in the *Clcn1* minigene. (**A**) Sequence of nt 451–533 of the *Clcn1* minigene that corresponds to a 3′ portion of intron 6 and a 5′ portion of exon 7A. Substituted nucleotides in Mut 1 are shown in boldface. The first 15 nt of exon 7A were deleted in Mut 2. (**B**) Splicing assays results for Mut 1 and Mut 2 minigenes in COS-7 cells. Mut 2 exhibited a complete loss of exon 7A inclusion. (**C**) Effects of deletions in the 5′ region of *Clcn1* exon 7A on basal-splicing efficiency. Deleted regions in Mut 2, Mut 3 and Mut 4 are indicated (upper). Splicing assay results of Mut 2/3/4/minigenes transfected into COS-7 cells are shown (lower). WT: wild type. (**D**) Point mutations at the 5′ end of exon 7A disrupted the basal inclusion of exon 7A. Alignment of part of Mut3 and Mut4 sequences (upper). Different nucleotides between Mut 3 and Mut 4 are indicated by boldface and underlined text. Point mutations that partly mimic the difference between Mut 3 and Mut 4 were introduced in the Mut 5a, Mut 5b, a504c and t512g mutants, as indicated by boldface and underlined text (middle). Splicing assay results of the mutant minigenes in COS-7 cells (lower).

hypersensitivity to MBNL1. Thus, we concluded that the 8 nt at the 5′ end of exon 7A, 504–511, contain an ESE essential for the basal inclusion of exon 7A.

DISCUSSION

CLCN1/*Clcn1* splicing is a key event in DM. Although the misregulation of splicing has been well established as a characteristic abnormality of DM, few misregulated genes have a clear causal relationship to symptoms of DM. *Clcn1* misregulation can account for myotonia in DM model mice (38,47,48). As demonstrated recently, the skipping of exon 7A induced by AON reversed the myotonic phenotype of DM model mice (38), making *CLCN1* splicing a promising target for therapeutic approaches. Understanding *Clcn1*/*CLCN1* splicing would aid in the design of rational strategies for correcting CLCN1 expression to perhaps prevent myotonia. In addition, the identification of molecular pathways causative of *Clcn1* missplicing is important for the evaluation or interpretation of DM model mice. Besides the aforementioned model mice, myotonic phenotypes are observed in several other mouse models (49–52). It is unclear whether the loss of MBNL function can explain all of these myotonic mice (51). In these contexts, a *Clcn1* minigene should be useful for identifying and characterizing the factors that modulate the splicing pattern of *Clcn1*. Indeed, we have identified multiple determinants for the regulation of exon 7A splicing using a *Clcn1* minigene and its derivatives.

MBNL proteins regulate *Clcn1* exon 7A

Previous genetic studies have indicated the involvement of MBNL proteins in the regulation of *Clcn1* splicing. The loss of either *Mbnl1* or *Mbnl2* can cause myotonia in mice, whereas overexpression of Mbnl1 rescues the myotonic phenotype in HSA^{LR} mice (13,23,37). However, these studies have not excluded the indirect involvement of MBNL proteins such that they regulate muscle differentiation or maintenance, which in turn alters the activity of other splicing factors to result in altered *Clcn1* splicing. Indeed, exon 7A of *Clcn1* is regulated through muscle development and is preferentially included in neonatal muscle but not in adult muscle (16,47). Here, we have demonstrated that the splicing regulation of *Clcn1* exon 7A by MBNL1 was observed in COS-7, HeLa and Neuro2A cell lines (Figures 1 and 2 and Supplementary Figure S2). Thus, the regulation of exon 7A can be determined directly by the expression level of MBNL proteins, even without the context of muscle cells. The inclusion of exon 7A was repressed by the overexpression of MBNL proteins but increased by their knockdown (Figures 1 and 2). In addition, MBNL1 associated with the pre-mRNA of *Clcn1* in the cell (Figure 6A), and direct binding between MBNL1 and an RNA fragment of *Clcn1* was verified *in vitro* (Figure 6C). These results are

consistent with the model that MBNL proteins directly regulate *CLCN1*/*Clcn1* and that the loss of MBNL function leads to *CLCN1*/*Clcn1* misregulation in DM.

Differential effects of CELF proteins on *Clcn1* splicing

In contrast to MBNL proteins, CELF3/4/5/6 promoted increased inclusion of exon 7A of mouse *Clcn1* (Figure 1B). Among these CELF proteins, CELF4 is expressed in a wide variety of tissues, including muscle (26,43). Although mice deficient in *Celf4* have been reported to manifest a complex seizure phenotype (53), the physiological function of CELF4 is largely unclear. Although an elevation of CUG-BP and ETR-3 proteins was observed in DM1 patients, the other CELF proteins have not been well characterized. The expression level, intracellular localization, and activity of CELF4 (and CELF3/5/6) should be investigated in the context of DM. Although *Clcn1* is enriched in muscle, it is expressed in other tissues including brain even at a low level (data not shown). Because some CELF proteins are enriched in the brain (43), they might play a role in keeping *Clcn1* expression at a low level in the brain or possibly other tissues through a splicing-mediated regulation of expression.

In contrast, CUG-BP and ETR-3 did not directly change the ratio of exon 7A inclusion in our mouse *Clcn1* minigene (Figures 1B and 2C). This was unexpected because the inclusion of exon 7A has been reported to be increased in CUG-BP transgenic mice (33). It is possible that the regulation of exon 7A by CUG-BP is dependent on some specific cellular conditions and/or the genomic context of *Clcn1* that is missed in the minigene. Alternatively, the transgenic mouse might involve indirect or secondary effects of long-term CUG-BP overexpression. Nevertheless, our results might be compatible with a role of CUG-BP or ETR-3 in the misregulation of mouse *Clcn1*, as the overexpression of these proteins could reduce spliced *Clcn1* products and increase intron retention (Figure 1C). It should be noted, however, that both CUG-BP and ETR-3 increase neither the inclusion of exon 7A nor the unspliced product in the human *CLCN1* minigene in our preliminary results (Kino *et al.*, unpublished data). The determinants of the difference between human and mouse minigenes are currently under investigation. Our results also revealed a functional heterogeneity among CELF proteins. The *Clcn1* minigene would allow us to determine the cause of the difference in the effects of CELF proteins on exon 7A inclusion.

Multiple factors involved in the regulation of *Clcn1* exon 7A splicing

To clarify the regulatory mechanism of exon 7A splicing, we analyzed a series of *Clcn1* mutant minigenes. We found that the regulation of exon 7A by MBNL1, as well as their binding, was at least partly mediated by the sequence of exon 7A itself. Furthermore, the MBNL-responsive regions were mapped roughly to a boundary region of exon 7A and its flanking intronic sequence, or 451–533. In the previous examples, MBNL1 regulated its target exons through intronic sequences of *cTNT*, *TNNT3*, and *SERCA1* (24,45,54). In contrast, neither intron 6 nor 7A of *Clcn1* was sufficient for response to MBNL1 (Figures 4A and 5C). Our results suggest that not only intronic but also exonic sequences should be considered for analyses such as genome-wide identification of the MBNL splicing targets.

We found an ESE located at the 5′ end of exon 7A (504–511) that was essential for the basal inclusion of exon 7A (Figure 7). The nucleotide sequence of this ESE is completely conserved in human *CLCN1*. In general, ESEs are important motifs for both constitutive and alternative splicing and are bound by SR proteins that enhance the recognition of an exon and promote splicing (55,56). The ESEfinder algorithm (http://rulai.cshl.edu/tools/ESE/; 57) predicted that the candidate region overlaps with potential recognition sites for multiple SR proteins (data not shown). However, the prediction pattern went unchanged even when a critical mutation (a504c) was introduced. Interestingly, the ESE in the 5′ end of exon 7A overlaps with the region targeted by the AON effective against myotonia in mouse models (38). Hybridization of the AON may prevent the functioning of not only the polypyrimidine (PY) tract and 3′ splice site but also the ESE, leading to efficient exon 7A skipping.

Previously, we have suggested double-stranded RNA with mismatches as putative MBNL1-binding targets (22), which was supported by recent analyses directly demonstrating that stem-loop structures with a pyrimidine mismatch are binding sequences of MBNL1 and mediate its splicing regulation (44,45). In contrast, YGCU(U/G)Y motifs have been reported as binding sites for MBNL1 in *cTNT* and *SERCA1* (24,54). In the case of *cTNT*, however, the YGCU(U/G)Y motifs were later shown to be part of a stem-loop structure (44). Here, we found a stem-loop structure in the junctional region between intron 6 and exon 7A as a candidate region for MBNL1-binding. Indeed, this fragment could bind with recombinant MBNL1 (Figure 6C). Furthermore, substitutions that alter this hairpin region impaired the responsiveness to MBNL1 (Figures 4A, 5C, Supplementary Figure 3) and the binding of RNA to MBNL1 (Figure 6C). Interestingly, this hairpin region overlaps with splicing signals, such as the PY tract, 3′ splice site and the ESE described previously. In particular, the ESE overlaps with an RNA region essential for efficient binding to MBNL1 (Figures 6C and 7). Therefore, one possible explanation of how MBNL1 acts on exon 7A splicing is that MBNL1 binds to the hairpin structure and inhibits the access of splicing factors to these splicing signals, like the AON described previously. The Clcn1(GAA) mutant exhibited residual binding to MBNL1 (Figure 6C). This might be explained by the presence of a YGCU(U/G)Y motif in intron 6.

Lastly, the region of *Clcn1* that mediates the effect of CELF4 was located within 181–413 (Figure 5D and E), which was not necessary for the regulation by MBNL1 (Figure 3). Therefore, antagonistic regulation by MBNL1 and CELF4 is mediated by distinct regions of the minigene. This is similar to the case of *cTNT* exon 5, which is regulated by CUG-BP and MBNL1 through distinct

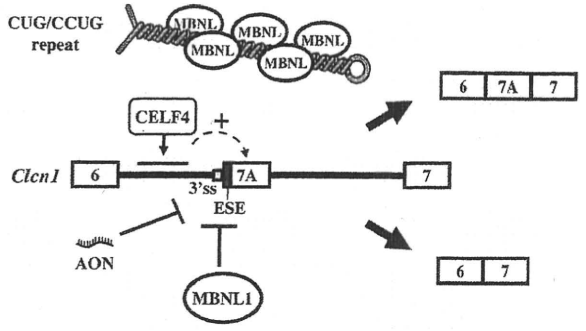

Figure 8. Model of mouse *Clcn1* splicing regulation by multiple factors. *Cis*- and *trans*-acting factors involved in the splicing regulation of *Clcn1* exon 7A are depicted. MBNL1 represses exon 7A inclusion through a region containing the 5′ region of exon 7A as well as its flanking intronic sequence. An exonic splicing enhancer (ESE) is located at the 5′ end of exon 7A. An antisense oligonucleotide (AON) previously reported (38), as well as MBNL1, might act in part through inhibiting this ESE. The facilitation of exon 7A inclusion by CELF4 is mediated by a region located in intron 6. Expanded CUG/CCUG repeat RNA may deplete MBNL proteins, resulting in the facilitation of exon 7A inclusion.

binding motifs (24). Because the Δ4 mutant exhibited a higher basal inclusion of exon 7A (Figure 5E), the deleted region (181–413) may contain an element that represses exon 7A inclusion. Perhaps CELF4 might antagonize this repressive element, which might explain how the effect of CELF4 was lost in the case of the Δ4 mutant.

Our current model for *Clcn1* regulation is depicted in Figure 8. Alternative splicing of exon 7A can be regulated by MBNL and by certain CELF proteins, such as CELF4. Expanded CUG/CCUG repeats may deplete MBNL and decrease the probability of binding between MBNL and the pre-mRNA of *Clcn1*. The ESE located at the 5′ end of exon 7A and the flanking splicing signals might be common targets of MBNL and the previous AON. Taken together, our results reveal novel regulatory properties of MBNL and CELF proteins and provide a molecular basis for the mechanism of splicing regulation of *Clcn1*, which might underlie rational therapeutic strategies targeting myotonia in DM.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Yoshihide Hayashizaki for FANTOM3 clones and Dr Gen Matsumoto for the R-miR vector. They are also grateful to the staffs of the Research Resource Center of RIKEN Brain Science Institute for DNA sequencing and cell sorting.

FUNDING

Ministry of Health, Labor and Welfare; Ministry of Education, Culture, Sports, Science, and Technology of Japan; Human Frontier Science Program. JSPS Research Fellowship for Young Scientists and the RIKEN Special Postdoctoral Researchers Program to Y.K. Funding for open access charge: Ministry of Health, Labor and Welfare, Japan.

Conflict of interest statement. None declared.

REFERENCES

1. Harper,P.S. (2001) *Myotonic Dystrophy*, 3rd edn. WB Saunders, London.
2. Brook,J.D., McCurrach,M.E., Harley,H.G., Buckler,A.J., Church,D., Aburatani,H., Hunter,K., Stanton,V.P., Thirion,J.P., Hudson,T. *et al.* (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell*, **68**, 799–808.
3. Fu,Y.H., Friedman,D.L., Richards,S., Pearlman,J.A., Gibbs,R.A., Pizzuti,A., Ashizawa,T., Perryman,M.B., Scarlato,G., Fenwick,R.G. *et al.* (1993) Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science*, **260**, 235–238.
4. Mahadevan,M., Tsilfidis,C., Sabourin,L., Shutler,G., Amemiya,C., Jansen,G., Neville,C., Narang,M., Barcelo,J., O'Hoy,K. *et al.* (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science*, **255**, 1253–1255.
5. Mankodi,A., Logigian,E., Callahan,L., McClain,C., White,R., Henderson,D., Krym,M. and Thornton,C.A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*, **289**, 1769–1773.
6. Liquori,C.L., Ricker,K., Moseley,M.L., Jacobsen,J.F., Kress,W., Naylor,S.L., Day,J.W. and Ranum,L.P. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science*, **293**, 864–867.
7. Taneja,K.L., McCurrach,M., Schalling,M., Housman,D. and Singer,R.H. (1995) Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.*, **128**, 995–1002.
8. Davis,B.M., McCurrach,M.E., Taneja,K.L., Singer,R.H. and Housman,D.E. (1997) Expansion of a CUG trinucleotide repeat in the 3′ untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc. Natl Acad. Sci. USA*, **94**, 7388–7393.
9. Philips,A.V., Timchenko,L.T. and Cooper,T.A. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science*, **280**, 737–741.
10. Savkur,R.S., Philips,A.V. and Cooper,T.A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat. Genet.*, **29**, 40–47.
11. Charlet-B,N., Savkur,R.S., Singh,G., Philips,A.V., Grice,E.A. and Cooper,T.A. (2002) Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cell*, **10**, 45–53.
12. Mankodi,A., Takahashi,M.P., Jiang,H., Beck,C.L., Bowers,W.J., Moxley,R.T., Cannon,S.C. and Thornton,C.A. (2002) Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cell*, **10**, 35–44.
13. Kanadia,R.N., Johnstone,K.A., Mankodi,A., Lungu,C., Thornton,C.A., Esson,D., Timmers,A.M., Hauswirth,W.W. and Swanson,M.S. (2003) A muscleblind knockout model for myotonic dystrophy. *Science*, **302**, 1978–1980.
14. Kimura,T., Nakamori,M., Lueck,J.D., Pouliquin,P., Aoike,F., Fujimura,H., Dirksen,R.T., Takahashi,M.P., Dulhunty,A.F. and Sakoda,S. (2005) Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca2+-ATPase in myotonic dystrophy type 1. *Hum. Mol. Genet.*, **14**, 2189–2200.
15. Savkur,R.S., Philips,A.V., Cooper,T.A., Dalton,J.C., Moseley,M.L., Ranum,L.P. and Day,J.W. (2004) Insulin receptor splicing alteration in myotonic dystrophy type 2. *Am. J. Hum. Genet.*, **74**, 1309–1313.

16. Lin,X., Miller,J.W., Mankodi,A., Kanadia,R.N., Yuan,Y., Moxley,R.T., Swanson,M.S. and Thornton,C.A. (2006) Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.*, **15**, 2087–2097.
17. Ranum,L.P. and Cooper,T.A. (2006) RNA-mediated neuromuscular disorders. *Annu. Rev. Neurosci.*, **29**, 259–277.
18. Begemann,G., Paricio,N., Artero,R., Kiss,I., Perez-Alonso,M. and Mlodzik,M. (1997) muscleblind, a gene required for photoreceptor differentiation in Drosophila, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development*, **124**, 4321–4331.
19. Artero,R., Prokop,A., Paricio,N., Begemann,G., Pueyo,I., Mlodzik,M., Perez-Alonso,M. and Baylies,M.K. (1998) The muscleblind gene participates in the organization of Z-bands and epidermal attachments of Drosophila muscles and is regulated by Dmef2. *Dev. Biol.*, **195**, 131–143.
20. Fardaei,M., Rogers,M.T., Thorpe,H.M., Larkin,K., Hamshere,M.G., Harper,P.S. and Brook,J.D. (2002) Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum. Mol. Genet.*, **11**, 805–814.
21. Miller,J.W., Urbinati,C.R., Teng-Umnuay,P., Stenberg,M.G., Byrne,B.J., Thornton,C.A. and Swanson,M.S. (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.*, **19**, 4439–4448.
22. Kino,Y., Mori,D., Oma,Y., Takeshita,Y., Sasagawa,N. and Ishiura,S. (2004) Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.*, **13**, 495–507.
23. Hao,M., Akrami,K., Wei,K., De Diego,C., Che,N., Ku,J.H., Tidball,J., Graves,M.C., Shieh,P.B. and Chen,F. (2008) Muscleblind-like 2 (Mbnl2) -deficient mice as a model for myotonic dystrophy. *Dev. Dyn.*, **237**, 403–410.
24. Ho,T.H., Charlet-B,N., Poulos,M.G., Singh,G., Swanson,M.S. and Cooper,T.A. (2004) Muscleblind proteins regulate alternative splicing. *EMBO J.*, **23**, 3103–3112.
25. Dansithong,W., Paul,S., Comai,L. and Reddy,S. (2005) MBNL1 is the primary determinant of focus formation and aberrant insulin receptor splicing in DM1. *J. Biol. Chem.*, **280**, 5773–5780.
26. Ladd,A.N., Charlet,N. and Cooper,T.A. (2001) The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol. Cell Biol.*, **21**, 1285–1296.
27. Iakova,P., Wang,G.L., Timchenko,L., Michalak,M., Pereira-Smith,O.M., Smith,J.R. and Timchenko,N.A. (2004) Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate. *EMBO J.*, **23**, 406–417.
28. Anant,S., Henderson,J.O., Mukhopadhyay,D., Navaratnam,N., Kennedy,S., Min,J. and Davidson,N.O. (2001) Novel role for RNA-binding protein CUGBP2 in mammalian RNA editing. CUGBP2 modulates C to U editing of apolipoprotein B mRNA by interacting with apobec-1 and ACF, the apobec-1 complementation factor. *J. Biol. Chem.*, **276**, 47338–47351.
29. Mukhopadhyay,D., Houchen,C.W., Kennedy,S., Dieckgraefe,B.K. and Anant,S. (2003) Coupled mRNA stabilization and translational silencing of cyclooxygenase-2 by a novel RNA binding protein, CUGBP2. *Mol. Cell*, **11**, 113–126.
30. Timchenko,N.A., Cai,Z.J., Welm,A.L., Reddy,S., Ashizawa,T. and Timchenko,L.T. (2001) RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J. Biol. Chem.*, **276**, 7820–7826.
31. Kuyumcu-Martinez,N.M., Wang,G.S. and Cooper,T.A. (2007) Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Mol. Cell*, **28**, 68–78.
32. Timchenko,N.A., Patel,R., Iakova,P., Cai,Z.J., Quan,L. and Timchenko,L.T. (2004) Overexpression of CUG triplet repeat-binding protein, CUGBP1, in mice inhibits myogenesis. *J. Biol. Chem.*, **279**, 13129–13139.
33. Ho,T.H., Bundman,D., Armstrong,D.L. and Cooper,T.A. (2005) Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy. *Hum. Mol. Genet.*, **14**, 1539–1547.
34. Kleopa,K.A. and Barchi,R.L. (2002) Genetic disorders of neuromuscular ion channels. *Muscle Nerve*, **26**, 299–325.
35. Pusch,M. (2002) Myotonia caused by mutations in the muscle chloride channel gene CLCN1. *Hum. Mutat.*, **19**, 423–434.
36. Berg,J., Jiang,H., Thornton,C.A. and Cannon,S.C. (2004) Truncated ClC-1 mRNA in myotonic dystrophy exerts a dominant-negative effect on the Cl current. *Neurology*, **63**, 2371–2375.
37. Kanadia,R.N., Shin,J., Yuan,Y., Beattie,S.G., Wheeler,T.M., Thornton,C.A. and Swanson,M.S. (2006) Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc. Natl Acad. Sci. USA*, **103**, 11748–11753.
38. Wheeler,T.M., Lueck,J.D., Swanson,M.S., Dirksen,R.T. and Thornton,C.A. (2007) Correction of ClC-1 splicing eliminates chloride channelopathy and myotonia in mouse models of myotonic dystrophy. *J. Clin. Invest.*, **117**, 3952–3957.
39. Orengo,J.P., Chambon,P., Metzger,D., Mosier,D.R., Snipes,G.J. and Cooper,T.A. (2008) Expanded CTG repeats within the DMPK 3′ UTR causes severe skeletal muscle wasting in an inducible mouse model for myotonic dystrophy. *Proc. Natl Acad. Sci. USA*, **105**, 2646–2651.
40. Takahashi,N., Sasagawa,N., Usuki,F., Kino,Y., Kawahara,H., Sorimachi,H., Maeda,T., Suzuki,K. and Ishiura,S. (2001) Coexpression of the CUG-binding protein reduces DM protein kinase expression in COS cells. *J. Biochem.*, **130**, 581–587.
41. Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559–1563.
42. Niranjanakumari,S., Lasda,E., Brazas,R. and Garcia-Blanco,M.A. (2002) Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods*, **26**, 182–190.
43. Ladd,A.N., Nguyen,N.H., Malhotra,K. and Cooper,T.A. (2004) CELF6, a member of the CELF family of RNA-binding proteins, regulates muscle-specific splicing enhancer-dependent alternative splicing. *J. Biol. Chem.*, **279**, 17756–17764.
44. Warf,M.B. and Berglund,J.A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA*, **13**, 2238–2251.
45. Yuan,Y., Compton,S.A., Sobczak,K., Stenberg,M.G., Thornton,C.A., Griffith,J.D. and Swanson,M.S. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.*, **35**, 5474–5486.
46. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
47. Lueck,J.D., Lungu,C., Mankodi,A., Osborne,R.J., Welle,S.L., Dirksen,R.T. and Thornton,C.A. (2007) Chloride channelopathy in myotonic dystrophy resulting from loss of posttranscriptional regulation for CLCN1. *Am. J. Physiol. Cell Physiol.*, **292**, C1291–C1297.
48. Lueck,J.D., Mankodi,A., Swanson,M.S., Thornton,C.A. and Dirksen,R.T. (2007) Muscle chloride channel dysfunction in two mouse models of myotonic dystrophy. *J. Gen. Physiol.*, **129**, 79–94.
49. Seznec,H., Agbulut,O., Sergeant,N., Savouret,C., Ghestem,A., Tabti,N., Willer,J.C., Ourth,L., Duros,C., Brisson,E. *et al.* (2001) Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. *Hum. Mol. Genet.*, **10**, 2717–2726.
50. O'Cochlain,D.F., Perez-Terzic,C., Reyes,S., Kane,G.C., Behfar,A., Hodgson,D.M., Strommen,J.A., Liu,X.K., van den Broek,W., Wansink,D.G. *et al.* (2004) Transgenic overexpression of human DMPK accumulates into hypertrophic cardiomyopathy, myotonic myopathy and hypotension traits of myotonic dystrophy. *Hum. Mol. Genet.*, **13**, 2505–2518.
51. Mahadevan,M.S., Yadava,R.S., Yu,Q., Balijepalli,S., Frenzel-McCardell,C.D., Bourne,T.D. and Phillips,L.H. (2006) Reversible model of RNA toxicity and cardiac conduction defects in myotonic dystrophy. *Nat. Genet.*, **38**, 1066–1070.
52. Chen,W., Wang,Y., Abe,Y., Cheney,L., Udd,B. and Li,Y.P. (2007) Haploinsuffciency for Znf9 in Znf9 + /- mice is associated with multiorgan abnormalities resembling myotonic dystrophy. *J. Mol. Biol.*, **368**, 8–17.
53. Yang,Y., Mahaffey,C.L., Bérubé,N., Maddatu,T.P., Cox,G.A. and Frankel,W.N. (2007) Complex seizure disorder caused by Brunol4 deficiency in mice. *PLoS Genet.*, **3**, e124.

54. Hino,S., Kondo,S., Sekiya,H., Saito,A., Kanemoto,S., Murakami,T., Chihara,K., Aoki,Y., Nakamori,M., Takahashi,M.P. *et al.* (2007) Molecular mechanisms responsible for aberrant splicing of SERCA1 in myotonic dystrophy type 1. *Hum. Mol. Genet.*, **16**, 2834–2843.
55. Berget,S.M. (1995) Exon recognition in vertebrate splicing. *J. Biol. Chem.*, **270**, 2411–2314.
56. Blencowe,B.J. (2000) Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.*, **25**, 106–110.
57. Cartegni,L., Wang,J., Zhu,Z., Zhang,M.Q. and Krainer,A.R. (2003) ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res.*, **31**, 3568–3571.

Polyalanine tracts directly induce the release of cytochrome *c*, independently of the mitochondrial permeability transition pore, leading to apoptosis

Kazuya Toriumi[1], Yoko Oma[1], Ai Mimoto[1], Eugene Futai[1], Noboru Sasagawa[1], Boris Turk[2] and Shoichi Ishiura[1]*

[1]*Department of Life Sciences, Graduate School of Arts and Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan*
[2]*Department of Biochemistry and Molecular and Structural Biology, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

In recent years, several novel types of disorder caused by the expansion of triplet repeats in specific genes have been characterized; in the "polyalanine diseases", these expanded repeats result in proteins with aberrantly elongated polyalanine tracts. In this study, we fused expanded polyalanine tracts to yellow fluorescent protein to examine their physical interaction with mitochondria. Tracts containing more than 23 alanine repeats were found to physically associate with mitochondria, strongly suggesting that an interaction between polyalanine tracts and mitochondria is a contributing factor in the pathology of polyalanine diseases. Furthermore, in *in vitro* experiments, polyalanine tracts induced release of cytochrome *c* from mitochondria and caspase-3 activation, independently of the mitochondrial permeability transition pore. These results suggest that oligomerized polyalanine tracts might induce the rupture of the mitochondrial membrane, the subsequent release of cytochrome *c*, and apoptosis. This novel mechanism for polyalanine tract cytotoxicity might be common to the pathogenesis of all polyalanine diseases. Further investigation of this mechanism might aid the development of therapies for these diseases.

Introduction

The expansion of trinucleotide repeats encoding polyalanine tracts has been recently shown to cause nine human diseases (Amiel *et al.* 2004; Brown & Brown 2004). The presence of protein aggregates in the affected tissues is a characteristic feature of these polyalanine diseases (Albrecht *et al.* 2004; Nasrallah *et al.* 2004), and the length of the polyalanine repeat correlates with the severity of the phenotype. For example, in oculopharyngeal muscular dystrophy (OPMD), an adult-onset disorder characterized by progressive eyelid drooping, swallowing difficulties, and proximal limb weakness, there is an abnormal expansion of the N terminal polyalanine tract of the PABPN1 protein, producing an autosomal dominant phenotype (Brais *et al.* 1998). In patients with OPMD, the polyalanine tract contains 12–17 alanine repeats, whereas in normal individuals, it contains only 10 repeats. In affected skeletal muscle, intranuclear aggregations are observed, and cell death is induced through an unknown mechanism.

In a previous study (Toriumi *et al.* 2008), we examined the effect of expanded polyalanine tracts fused with yellow fluorescent protein (YFP) on cultured cell lines. Cells expressing the YFP-fused expanded polyalanine tracts exhibited cytotoxicity and aggregate formation like that seen in polyalanine diseases. We speculated that the polyalanine tract itself possesses properties that bring about cell death by a mechanism common to the polyalanine diseases. We also found that polyalanine tracts are associated with mitochondria and that expression of polyalanine tracts decreases the mitochondrial membrane (MM) potential, a step considered key in the initial apoptotic process (Crompton 1999). In response to an apoptotic signal, mitochondria decrease their membrane potential, which increases the permeability of their outer membranes and allows the release of various apoptogenic factors that normally reside in the intermembrane space of these organelles. Among these factors is cytochrome *c*, an important component of the mitochondrial respiratory

Communicated by: Masayuki Yamamoto (Tohoku University)
**Correspondence*: cishiura@mail.ecc.u-tokyo.ac.jp

DOI: 10.1111/j.1365-2443.2009.01307.x

chain. Once released into the cytosol, cytochrome *c* activates caspase-9, which in turn activates caspase-3 and caspase-7. These activated caspases kill the cell by proteolysis, leading to biochemical and morphological features characteristic of apoptosis (Li *et al.* 1997). The release of cytochrome *c* has been widely believed to be mediated by the opening of mitochondrial permeability transition (PT) pores; these pores include the voltage-dependent anion channel (VDAC), the adenine nucleotide translocator (ANT), and cyclophilin D (Tsujimoto & Shimizu 2007). Opening of the PT pore is strictly regulated by proteins of the Bcl-2 family. Bax, one of these proteins, physically interacts with VDAC in the outer MM, leading to opening of the PT pore and induction of cytochrome *c* release. Conversely, the anti-apoptotic factors Bcl-2 and Bcl-xL induce PT pore closing and the blocking of cytochrome *c* release through their association with VDAC (Shimizu *et al.* 1999, 2000a).

In the present study, we show that binding of polyalanine tracts to mitochondria induces the release of cytochrome *c* into the cytoplasm, independently of the PT pore and the subsequent activation of caspase-3. Our results suggest a novel mechanism by which polyalanine tracts bring about cytotoxicity through the apoptotic process. The investigation of this mechanism may provide insights into the pathogenesis of, and possible therapeutic strategies for, polyalanine diseases.

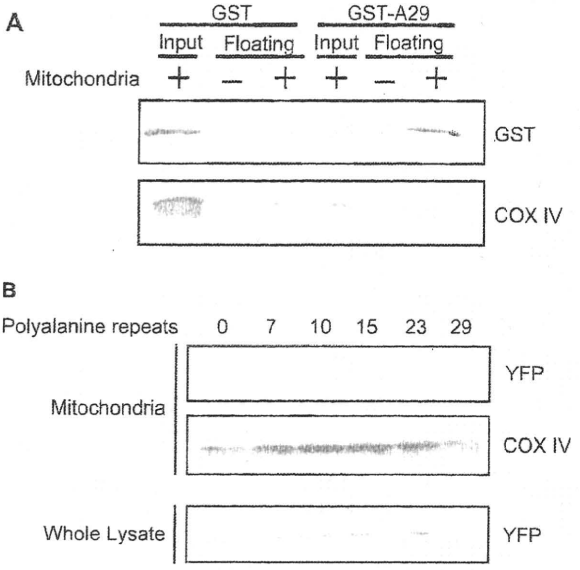

Figure 1 Long polyalanine tracts directly associate with mitochondria. (A) The *in vitro* mitochondrial-binding assay. Isolated mouse liver mitochondria were incubated with GST or GST-Ala29 and collected by floating using discontinuous sucrose gradient centrifugation. Input represents the mixture of mitochondria with GST or GST-Ala29 before floating. (B) YFP-polyalanine is physically associated with mitochondria *in vivo*. Immunoblot analysis of mitochondria isolated from COS-7 cells expressing YFP or a YFP-polyalanine fusion suggests an association between long polyalanine tracts and mitochondria. The mitochondrial content was quantified with anti-COX IV antibodies.

Results

Long polyalanine tracts physically associate with mitochondria

In a previous study, we detected polyalanine tracts in a mitochondrial fraction, suggesting that polyalanine tracts associate with mitochondria (Toriumi *et al.* 2008). To confirm this association *in vitro*, we performed a mitochondrial-binding assay using isolated mouse liver mitochondria. The mitochondria were incubated with GST or GST-Ala29 purified from *Escherichia coli* and then collected by discontinuous sucrose gradient centrifugation. Immunoblot analysis with an anti-GST antibody indicated that GST-Ala29, but not GST, was tightly associated with the recovered mitochondrial pellet (Fig. 1A).

Next, to determine the length of the polyalanine tract necessary for this observed association with mitochondria, we isolated mitochondria from COS-7 cells expressing YFP or YFP fused with polyalanine tracts of various lengths. Immunoblot results indicated that tracts with more than 23 alanine repeats strongly bound to the mitochondria (Fig. 1B), implying that a tract with 20 alanine repeats, which is lengthy enough to initiate polyalanine disease, is the threshold length necessary for association with mitochondria.

Expression of a long polyalanine tract induces cytochrome *c* release from mitochondria

In our previous study, we showed that the expression of polyalanine tracts decreases the MM potential, suggesting that an initial step of apoptosis is mediated by the mitochondria. We hypothesized that by directly binding to the mitochondria, polyalanine tracts induce the release of apoptogenic mitochondrial factors leading to an activation of an apoptotic cascade. Here, we evaluated this hypothesis by quantifying the amount of cytochrome *c* in the cytosol of COS-7 cells expressing YFP, YFP-Ala29, or YFP-Ala70 (Fig. 2). Staurosporine, an inducer of apoptosis, was used as a positive control. We found that the amount of cytosolic cytochrome *c* increased as the length of the expressed polyalanine tracts increased, suggesting that apoptosis induced by expression of polyalanine tracts is mediated by the mitochondria.

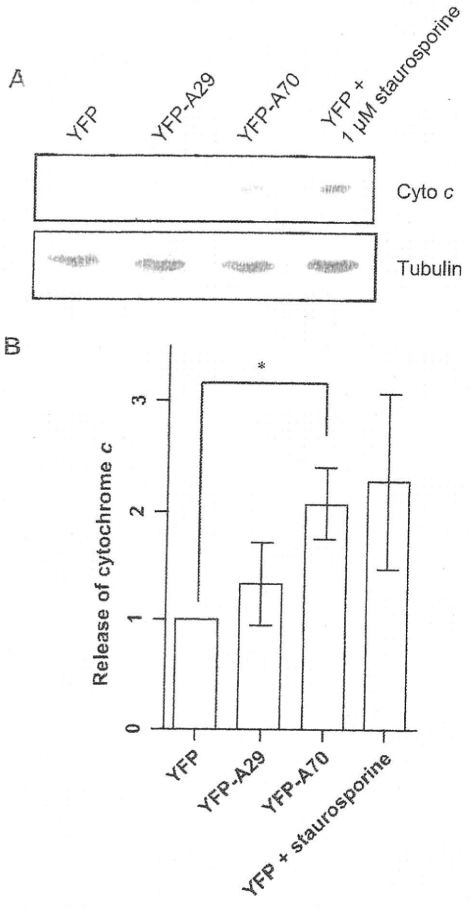

Figure 2 Expression of a long polyalanine tract induces mitochondrial release of cytochrome *c* into the cytoplasm. (A) Release of cytochrome *c* from the mitochondria of COS-7 cells expressing YFP, YFP-Ala29, or YFP-Ala70 into the cytoplasm was detected by immunoblot analysis using anti-cytochrome *c* antibody. The apoptosis inducer staurosporine (1 μM) was used as a positive control. (B) Band intensities from four independent experiments are shown as means ± SEM. Data were analyzed using ANOVA and *post hoc* Tukey's tests. $*P < 0.05$, compared with cells expressing YFP.

Polyalanine tracts directly induce release of cytochrome *c*

Next, we examined the more direct effects of abnormally expanded polyalanine tracts on the release of cytochrome *c* using an *in vitro* cytochrome *c*-release assay. Mitochondria isolated from mouse liver were incubated in the absence or presence of GST or GST-Ala29 (2 μM), and the organelles were centrifuged. The supernatant fractions were then analyzed for their cytochrome *c* content by immunoblotting (Fig. 3A), which demonstrated that incubation of the mitochondria with GST-Ala29 resulted in a significant release of cytochrome *c* to the cytoplasm. $CaCl_2$ was used as a positive control because Ca^{2+} is a known permeabilizer and PT pore opener.

To determine the concentration of GST-Ala29 required to trigger release of cytochrome *c*, GST or GST-Ala29 at various concentrations were added to isolated mitochondria, and the release of cytochrome *c* was determined. In these experiments, GST-Ala29 at concentrations of at least 1.5 μM brought about the release of cytochrome *c* (Fig. 3B), whereas GST at concentrations of up to 8.0 μM did not (data not shown). Additionally, when GST-Ala29 was present at 4.0 or 8.0 μM, COX IV was detected in the supernatant fraction after centrifugation. Since COX IV normally localizes to the inner MM, this result suggests that, at high concentrations, the polyalanine tract ruptured the MM, leading to the release of cytochrome *c*.

Polyalanine tracts induce cytochrome *c* release, independently of the PT pore

Although regulatory release of cytochrome *c* has been reported to occur through the PT pore, the GST-Ala29-induced release of cytochrome *c* seen in our study was probably not a result of regulatory opening of the PT pore; rather, it probably leaked through a rupture in the outer MM. To explore this possibility, we performed the *in vitro* cytochrome *c*-release assay in the presence of an inhibitor of PT pore opening. As the inhibitor, we used the conserved N-terminal homology domain (BH4) of Bcl-xL (amino acids 4–23), which inhibits VDAC activity in liposomes and isolated mitochondria (Shimizu *et al.* 2000b), thereby inhibiting cytochrome *c* release and loss of the MM potential. Although addition of the PT pore inhibitor prevented the release of cytochrome *c* normally caused by calcium overload, it did not prevent the GST-Ala29-induced release of cytochrome *c* (Fig. 4). This result is consistent with the proposal that polyalanine tracts disrupt the MM, allowing release of cytochrome *c* into the cytoplasm.

Expression of polyalanine tracts activates caspase-3

Release of cytochrome *c* into the cytoplasm has been reported to activate a sequential apoptotic cascade. To confirm that the polyalanine-induced release of cytochrome *c* leads to apoptosis, we examined caspase-3 activity, which is known to rise at the point in the cascade immediately downstream of cytochrome *c* release. Forty-eight hours after transfection of COS-7 cells with plasmids expressing YFP, YFP-Ala29, or YFP-Ala70, we assessed caspase-3 activity by measuring cleavage of Ac-Asp-Glu-Val-Asp-MCA, a specific peptide substrate of caspase-3. The caspase-3 activity in cells expressing

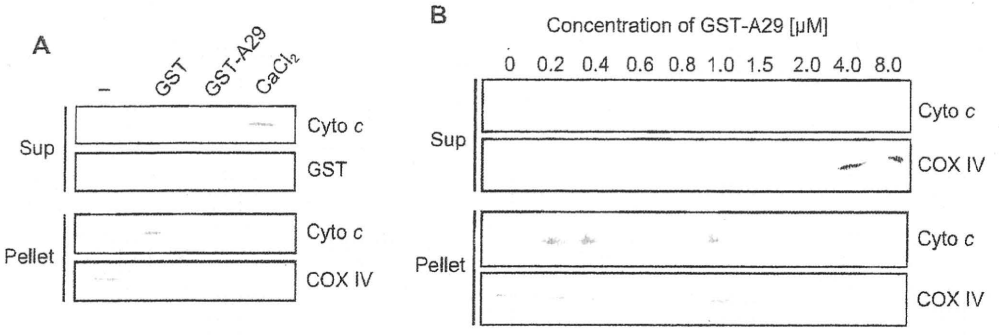

Figure 3 Polyalanine tracts directly induce the release of cytochrome *c in vitro*. Isolated mouse liver mitochondria were incubated in the absence or presence of GST or GST-Ala29 and then centrifuged at 1600 × *g* for 15 min. Cytochrome *c* in the resulting supernatant fractions was detected by immunoblotting. (A) Mitochondria were incubated in the absence or presence of 2 μM GST or GST-Ala29. $CaCl_2$ (500 μM) was used as a positive control. (B) Mitochondria were incubated in the presence of varying concentrations of GST-Ala29.

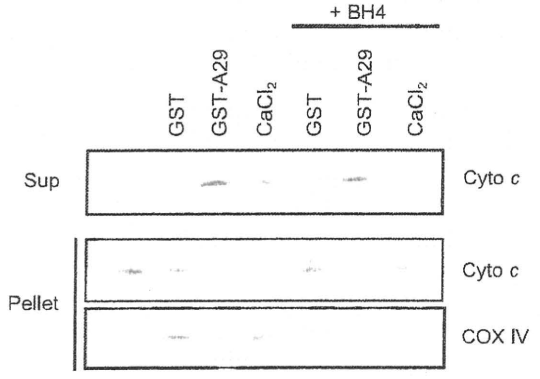

Figure 4 Polyalanine tracts induce cytochrome *c* release independent of the PT pore. The *in vitro* cytochrome *c*-release assay was performed in the presence of BH4 peptides, which inhibit induction of the PT pore. Mitochondria were incubated with BH4 peptides (20 μg/mL) for 15 min prior to the addition of polyalanine tracts. $CaCl_2$ (500 μM) was used as a positive control.

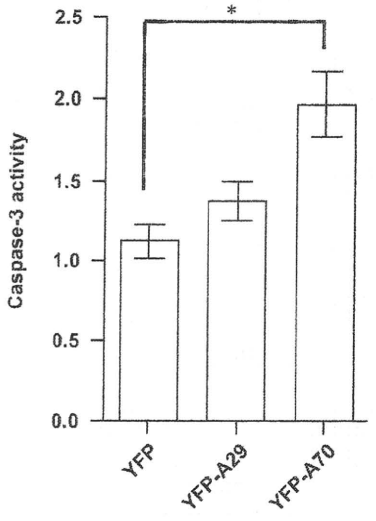

Figure 5 Expression of polyalanine tracts activates caspase-3. Ac-Asp-Glu-Val-Asp-MCA, a specific caspase-3 peptide substrate, was used to measure caspase-3 activity 48 h after transfection of COS-7 cells with plasmids expressing YFP or a YFP-polyalanine fusion protein. Data shown are means ± SEM. $*P < 0.05$, compared with cells expressing YFP. Data were analyzed by ANOVA and post-hoc Tukey's tests.

YFP-polyalanine fusion proteins increased in a polyalanine-length-dependent manner and was greater than that in cells expressing YFP alone (Fig. 5), suggesting that the apoptosis induced by expression of polyalanine tracts is mediated by the mitochondria.

Discussion

In this study, we found that polyalanine tracts containing more than 29 alanine repeats physically interact with mitochondria (Fig. 1). Since the onset of almost all polyalanine diseases results from an expansion of polyalanine tracts to more than 20 alanine repeats (Amiel *et al.* 2004; Brown & Brown 2004), these data strongly suggest that the interaction of polyalanine tracts with mitochondria is related to the pathology of polyalanine diseases. In fact, in affected skeletal muscle of OPMD patients, the mitochondria have an abnormal morphology (Gambelli *et al.* 2004). In this regard, we are aware of the relationship between optic atrophy induced by mitochondrial deficiency and eyelid drooping, which is a major phenotype of OPMD, one of the polyalanine diseases.

In a previous study of cell lysates from COS-7 cells transfected with YFP–polyalanine constructs (Oma *et al.* 2007), immunoblotting showed that the constructs with more than 23 alanine repeats were retained in the stacking gel after native PAGE, suggesting that they formed

oligomers. After SDS-PAGE, however, only the construct with 70 repeats was retained in the stacking gel; those constructs with 23 to 35 alanine repeats were not retained. These results suggest that moderately long polyalanine tracts (23–35 repeats) might form SDS-sensitive oligomers in the cell. In short, our study shows that the oligomerization of polyalanine tracts is sufficient to induce their cytotoxicity.

Since we were able to detect an association between polyalanine tracts and mitochondria *in vitro* in the absence of a cytosolic factor for mitochondrial import, such as MSF or Hsp70 (Hachiya *et al.* 1994; Komiya *et al.* 1996), we conclude that the polyalanine tracts were probably bound directly to the outer MM. Furthermore, we found that the amount of polyalanine tract detected in association with the mitochondria decreased gradually with a high-pH wash (data not shown), consistent with their binding to the outer MM.

Next, using an *in vitro* cytochrome *c*-release assay, we found that polyalanine tracts at concentrations greater than 1.5 μM directly induced the release of cytochrome *c*. This release was not blocked by BH4 peptides, suggesting that the release was not mediated by the PT pore (Figs 3 and 4). The addition of polyalanine tracts at 4.0 or 8.0 μM also promoted an efflux of COX IV, which is normally localized to the inner MM. These results suggest that polyalanine tracts cause the MM to rupture, leading to the release of cytochrome *c*. This possibility is also supported by the observation that the amount of cytochrome *c* released from mitochondria was independent of the polyalanine tract concentration in an "all-or-none" fashion, and the release occurred in the absence of other apoptotic signals, such as Ca^{2+}. In the affected cells, oligomerization causes a local increase in the polyalanine tract concentration. When these tracts attack mitochondria, they may in turn induce MM rupture, leading to cytochrome *c* release.

We demonstrated the release of cytochrome *c* and the activation of caspase-3 in COS-7 cells expressing YFP-A70 (Figs 2 and 5). These findings were consistent with a report that expression of PABPN1 containing a polyalanine tract long enough to cause OPMD induced the release of cytochrome *c*, followed by apoptosis (Davies *et al.* 2008). In addition, polyalanine tracts have been reported to form a β-sheet and then to activate an apoptotic pathway *via* caspase-8 (Giri *et al.* 2003). Thus, the cytotoxicity induced by polyalanine tracts might be related to apoptotic cell death. However, YFP-A29 didn't induce the release of cytochrome *c* and the activation of caspase-3 (Figs 2 and 5) although it can induce its release in *in vitro* experiment (Fig. 3). We think the inconsistency might be produced by the expression level of YFP-A29 in COS-7 cells. Since it is unlikely that the concentration of YFP-Ala29 in cells is the same as used in *in vitro* experiment, YFP-A29 couldn't induce the release of cyctochrome *c* from mitochondria and subsequent activation of caspase-3. However, it is possible that the local concentration increased due to the oligomerizing property of YFP-A29, so that the increasing tendencies were observed in YFP-A29 expressing cells in Figs 2 and 5. We think that increase in local concentration of polyalanine stretch is important to attack mitochondria. If we consider this idea as a pharmacological target, we can suggest some potentially effective therapeutics against polyalanine diseases. In fact, some reagent, inhibitors of oligomerization such as trehalose and geldanamycine, have been reported to be effective against polyalanine disease model (Albrecht *et al.* 2004; Davies *et al.* 2006).

In conclusion, we propose a novel mechanism by which the polyalanine tracts cause apoptosis (Fig. 6). This mechanism for polyalanine tract cytotoxicity might be common to the pathogenesis of all polyalanine diseases. These findings

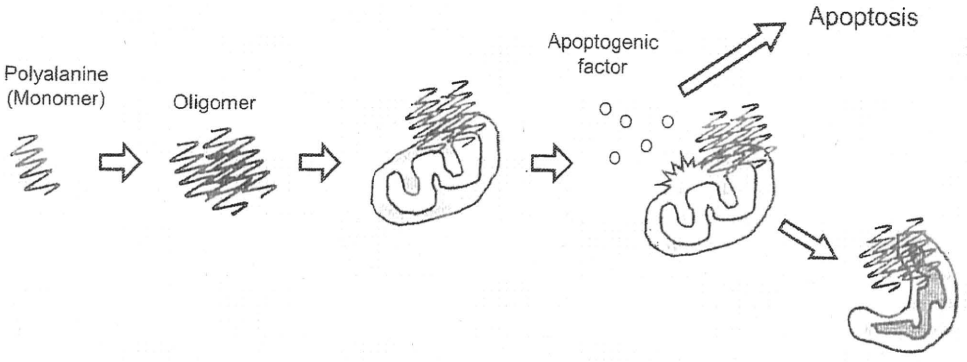

Figure 6 A model scheme for polyalanine tract cytotoxicity induced by mitochondrial dysfunction. Expanded polyalanine tracts oligomerize and then associate with the outer MM. This association induces rupture of the MM, leading to the release of apoptogenic factors such as cytochrome *c* and then to apoptosis. Loss of MM integrity might result in mitochondrial dysfunction. Some drugs known to prevent this cascade might be useful in treatment of polyalanine diseases.

in this report point to the mitochondria as good therapeutic targets for polyalanine diseases. We hope that these findings will benefit patients suffering from these diseases.

Experimental procedures

Construction of polyalanine-repeat sequences

Polyalanine repeat sequences [(Ala-)*n*] were synthesized by annealing of double-stranded oligonucleotides. These sequences were inserted into the *Eco*RV site of the pBlueScript KS(−) vector (TaKaRa), and the resulting constructs were digested with *Bam*HI and *Apa*I. The excised repeat sequences were then inserted into the *Bgl*II and *Apa*I sites of the pEYFP-C1 vector (Clontech) to allow the expression of the alanine repeats fused to the C-terminus of YFP. The resulting constructs were named YFP-Ala(*n*) according to the number of alanine repeats (*n*).

For *in vitro* assays of cytochrome *c* release, we prepared a construct encoding an $(Ala)_{29}$ tract fused to the glutathione-S-transferase (GST) protein. This construct, GST-Ala29, was prepared by excision of the $(Ala)_{29}$ sequence from YFP-Ala29 by digestion with *Xho*I and *Bam*HI and insertion of the fragment into the *Xho*I–*Bam*HI site of pGEX 4T-2 (Amersham Bioscience), a vector for expression of GST-fusion proteins in *Escherichia coli*.

Isolation of mitochondrial and cytosolic fractions

Transfected COS-7 cells were washed twice with ice-cold phosphate-buffered saline and scraped into mitochondrial buffer [20 mM Tris–HCl, pH 7.2, 250 mM sucrose, 2 mM EGTA, 40 mM KCl, 1 mg/mL bovine serum albumin (BSA)] containing 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1 : 1000 protease inhibitor cocktail (Sigma). The cells were fractured on ice using a 27-gauge needle. Unlysed cells, large debris, and nuclei were removed by centrifugation at low speed (50 × *g*, 10 min, 4 °C). The supernatant fraction was centrifuged at 1000 × *g* (10 min, 4 °C), and the resulting supernatant fraction was removed and centrifuged at 20 000 × *g* (10 min, 4 °C) to obtain the cytosolic fraction.

To purify mitochondria, the 1000 × *g* pellet was resuspended in mitochondrial buffer containing protease inhibitors, layered over a discontinuous 1.0/1.5 M sucrose gradient, and separated ("floated") by centrifugation at 100 000 × *g* in a swinging-bucket rotor (2 h, 4 °C). The hazy ring corresponding to the mitochondrial fraction was recovered carefully from the interface between the two sucrose solutions, diluted 1:2 in a buffer consisting of 20 mM Tris–HCl, pH 7.2, 2 mM EGTA, 40 mM KCl, and 1 mg/mL BSA, and centrifuged at 20 000 × *g* (20 min). The resulting pellet was resuspended in mitochondrial buffer containing protease inhibitors and washed twice by centrifugation (6800 × *g* for 5 min followed by 3800 × *g* for 5 min). The identity and purity of the mitochondrial fraction were determined by Western blot analysis using various cellular markers. Mitochondria were identified with antibodies against the nucleusencoded cytochrome *c* oxidase subunit IV (COX IV; BD Biosciences); the cytosolic fraction was identified with antibodies against α-tubulin (Sigma).

Isolation of mitochondria from mouse liver

Mice were fasted overnight, and their liver mitochondria were isolated by the method of Hogeboom (Hogeboom 1955) using a medium containing 0.25 M sucrose, 10 mM Tris–HCl, pH 7.4, and 0.1 mM EDTA. EDTA was omitted in the final wash, and the mitochondrial preparation was suspended in 0.25 M sucrose containing 10 mM Tris–HCl, pH 7.4.

Preparation of GST-fusion proteins

Escherichia coli cells were transformed with pGEX 4T-2 or the GST-Ala29-encoding plasmid construct and grown to an OD600 of 1.0. Expression was then induced by the addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and the cells were grown at 27 °C for 1 h. After centrifugation of the cells (1600 × *g*, 15 min), the pellet was placed on ice and resuspended in lysis buffer (50 mM Tris–HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA) containing 1 mM PMSF and 1:1000 protease inhibitor cocktail. The cells were fractured by sonication, and Triton X-100 was added to a final concentration of 1%. The resulting lysates were chilled on ice for 30 min and clarified by centrifugation (16 000 × *g*, 15 min). The supernatant fractions were incubated for 2 h at 4 °C with glutathione-Sepharose beads (Amersham Bioscience) that had been washed three times and resuspended in lysis buffer containing 1% Triton X-100.

Mitochondrial binding assay

Purified GST or GST-Ala29 (0.8 μM) were added to isolated mouse liver mitochondria (1 μg/μL) in binding buffer (20 mM HEPES, pH 6.8, 250 mM sucrose, 150 mM potassium acetate, 5 mM magnesium acetate), and the mixture was incubated at room temperature for 10 min. After centrifugation (1600 × *g*, 10 min), the pellet containing mitochondria was washed with binding buffer and resuspended in 50 μL of binding buffer. The mitochondrial suspension was added to 350 μL of 2.5 M sucrose buffer (20 mM HEPES, pH 6.8, 2.5 M sucrose, 150 mM potassium acetate, 5 mM magnesium acetate), and then layered under a discontinuous 1.0/1.5 M sucrose gradient. The mitochondria were floated by centrifugation at 100 000 × *g* in a swinging-bucket rotor (2 h, 4 °C). The hazy ring between the two sucrose solutions was recovered, diluted 1:2 in binding buffer, and centrifuged at 16 000 × *g* for 20 min. The resulting pellet was resuspended in binding buffer and washed by centrifugation at 16 000 × *g* for 10 min. GST-Ala29 protein associated with mitochondria was detected by Western blot analysis using anti-GST antibody. Mitochondria were identified with antibodies against the nucleus-encoded COX IV protein.

In vitro cytochrome *c*-release assay

To determine the effect of polyalanine tracts on the release of cytochrome *c*, isolated mouse liver mitochondria (0.6 μg/μL) were incubated in the presence or absence of GST or GST-Ala29 (various concentrations) for 30 min at 30 °C in a KCl-based medium

(125 mM KCl, 15 mM HEPES, pH 7.4, 0.5 mM EGTA, 4 mM $MgCl_2$, 5 mM Na_2HPO_4) containing glutamate and malate (1 mM each) as respiratory substrates. At the end of the incubation period, the mitochondrial suspensions were centrifuged at 1600 × *g* for 15 min. The supernatant fractions were removed and mixed with 5× SDS sample buffer, and the mitochondrial pellets were resuspended in 1× SDS sample buffer. Samples were boiled for 5 min and electrophoresed on a 12.5% SDS–polyacrylamide gel. The separated proteins were transferred to nitrocellulose and immunoblotted with antibodies specific for cytochrome *c* or COX IV.

For assays of cytochrome *c* release in the presence of PT pore inhibitor, the mitochondria were incubated with BH4 peptides (20 μg/mL, Calbiochem), which inhibit induction of the PT pore, for 15 min at 30 °C before polyalanine tracts were added.

Measurement of caspase-3 activity

COS-7 cells were transiently transfected with plasmids encoding YFP, YFP-Ala29, or YFP-Ala70, and the cells were harvested 48 h after transfection. The harvested cells were dissolved in extraction buffer (50 mM Tris–HCl, pH 7.5, 10 mM 2-mercaptoethanol, 1 mM EDTA) and subjected to three freeze-thaw rounds consisting of 60 s in liquid nitrogen followed by 90 s in a 30 °C water bath. The samples were centrifuged at 10 000 × *g* for 5 min, and the total protein (7.5 μg) in the supernatant fraction was dissolved in 200 μL of assay buffer (25 mM Tris–HCl, pH 7.5, 10 mM 2-mercaptoethanol, 1 mM EDTA). The fluorescent caspase-3 substrate Ac-Asp-Glu-Val-Asp-MCA (Peptide Institute) was added to a final concentration of 5 μM, and the mixtures were incubated at 37 °C for 30 min. The reactions were stopped by the addition of 100 μL of 10% SDS and 1 ml of 0.1 M sodium acetate, and the fluorescence was measured with a JASCO FP-777 fluorescence spectrometer with excitation at 380 nm and emission at 460 nm.

Acknowledgements

This work was supported in part by the Human Frontier Science Program and by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Albrecht, A.N., Kornak, U., Boddrich, A., Suring, K., Robinson, P.N., Stiege, A.C., Lurz, R., Stricker, S., Wanker, E.E. & Mundlos, S. (2004) A molecular pathogenesis for transcription factor associated poly-alanine tract expansions. *Hum. Mol. Genet.* **13**, 2351–2359.

Amiel, J., Trochet, D., Clement-Ziza, M., Munnich, A. & Lyonnet, S. (2004) Polyalanine expansions in human. *Hum. Mol. Genet.* **13**, R235–R243.

Brais, B., Bouchard, J.P., Xie, Y.G. *et al.* (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat. Genet.* **18**, 164–167.

Brown, L.Y. & Brown, S.A. (2004) Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends Genet.* **20**, 51–58.

Crompton, M. (1999) The mitochondrial permeability transition pore and its role in cell death. *Biochem. J.* **341** (Pt 2), 233–249.

Davies, J.E., Sarkar, S. & Rubinsztein, D.C. (2006) Trehalose reduces aggregate formation and delays pathology in a transgenic mouse model of oculopharyngeal muscular dystrophy. *Hum. Mol. Genet.* **15**, 23–31.

Davies, J.E., Sarkar, S. & Rubinsztein, D.C. (2008) Wild-type PABPN1 is anti-apoptotic and reduces toxicity of the oculopharyngeal muscular dystrophy mutation. *Hum. Mol. Genet.* **17**, 1097–1108.

Gambelli, S., Malandrini, A., Ginanneschi, F., Berti, G., Cardaioli, E., De Stefano, R., Franci, M., Salvadori, C., Mari, F., Bruttini, M., Rossi, A., Federico, A. & Renieri, A. (2004) Mitochondrial abnormalities in genetically assessed oculopharyngeal muscular dystrophy. *Eur. Neurol.* **51**, 144–147.

Giri, K., Ghosh, U., Bhattacharyya, N.P. & Basak, S. (2003) Caspase 8 mediated apoptotic cell death induced by b-sheet forming polyalanine peptides. *FEBS Lett.* **555**, 380–384.

Hachiya, N., Komiya, T., Alam, R., Iwahashi, J., Sakaguchi, M., Omura, T. & Mihara, K. (1994) MSF, a novel cytoplasmic chaperone which functions in precursor targeting to mitochondria. *EMBO J.* **13**, 5146–5154.

Hogeboom, G.H. (1955) *Methods in Enzymology*. New York: Academic Press.

Komiya, T., Sakaguchi, M. & Mihara, K. (1996) Cytoplasmic chaperones determine the targeting pathway of precursor proteins to mitochondria. *EMBO J.* **15**, 399–407.

Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S.M., Ahmad, M., Alnemri, E.S. & Wang, X. (1997) Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* **91**, 479–489.

Nasrallah, I.M., Minarcik, J.C. & Golden, J.A. (2004) A polyalanine tract expansion in Arx forms intranuclear inclusions and results in increased cell death. *J. Cell Biol.* **167**, 411–416.

Oma, Y., Kino, Y., Toriumi, K., Sasagawa, N. & Ishiura, S. (2007) Interactions between homopolymeric amino acids (HPAAs). *Protein Sci.* **16**, 2195–2204.

Shimizu, S., Ide, T., Yanagida, T. & Tsujimoto, Y. (2000a) Electrophysiological study of a novel large pore formed by Bax and the voltage-dependent anion channel that is permeable to cytochrome c. *J. Biol. Chem.* **275**, 12321–12325.

Shimizu, S., Konishi, A., Kodama, T. & Tsujimoto, Y. (2000b) BH4 domain of antiapoptotic Bcl-2 family members closes voltage-dependent anion channel and inhibits apoptotic mitochondrial changes and cell death. *Proc. Natl. Acad. Sci. USA* **97**, 3100–3105.

Shimizu, S., Narita, M. & Tsujimoto, Y. (1999) Bcl-2 family proteins regulate the release of apoptogenic cytochrome c by the mitochondrial channel VDAC. *Nature* **399**, 483–487.

Toriumi, K., Oma, Y., Kino, Y., Futai, E., Sasagawa, N. & Ishiura, S. (2008) Expression of polyalanine stretches induces mitochondrial dysfunction. *J. Neurosci. Res.* **86**, 1529–1537.

Tsujimoto, Y. & Shimizu, S. (2007) Role of the mitochondrial membrane permeability transition in cell death. *Apoptosis* **12**, 835–840.

Received: 1 January 2009
Accepted: 22 March 2009

Journal of Neuroscience Research 87:2290–2296 (2009)

Biochemical Analysis of Oligomerization of Expanded Polyalanine Repeat Proteins

Jun Nojima,[1] Yoko Oma,[1] Eugene Futai,[1] Noboru Sasagawa,[1] Reiko Kuroda,[1] Boris Turk,[2] and Shoichi Ishiura[1]*

[1]Department of Life Sciences, Graduate School of Arts and Sciences, University of Tokyo, Tokyo, Japan
[2]Department of Biochemistry and Molecular and Structural Biology, Jozef Stefan Institute, Ljubljana, Slovenia

Many human proteins contain amino acid repeats that can form homopolymeric amino acid (HPAA) tracts. HPAA tract proteins that contain polyalanine sequences promote diseases, including oculopharyngeal muscular dystrophy. The pathological properties of these proteins develop when the repeats match or exceed ~20 residues. We analyzed the oligomerization of yellow fluorescent protein (YFP) and GST fusion proteins containing >20 alanine repeats by using sucrose density gradient centrifugation. YFP and GST fusion proteins having 23 polyalanine residues sedimented readily in sucrose density gradients, suggesting instability and oligomerization of proteins with an excess of 20 alanine repeats. Moreover, GST fusion proteins were resistant to trypsin digestion after oligomerization. Oligomerized artificial proteins with long polyalanine repeats may be suitable models for studying polyalanine-related diseases.

Key words: polyalanine diseases; oligomerization; sucrose density centrifugation; trypsin digestion

Many human proteins contain consecutive amino acids known as homopolymeric amino acid (HPAA) tracts. HPAA tracts are found in transcription factors and DNA binding proteins (Alba and Guigo, 2004). Some diseases are known to be caused by proteins containing HPAA expanded repeats of glutamine (Gln) and alanine (Ala). Specifically, many polyalanine diseases are predominant hereditary disorders, and nine distinct diseases have been identified (Albrecht and Mundlos, 2005). Eight genes responsible for these nine polyalanine-related diseases are transcription factors, whereas poly-(A) binding protein nucleus 1 (PABPN1), which is the gene responsible for oculopharyngeal muscular dystrophy (OPMD), is not a transcription factor. PABPN1 controls the length of the poly-A tail in mRNA and, with its expanded polyalanine tracts, forms aggregates within the nucleus of skeletal muscle fibers in OPMD patients and in cultured cell and mouse models of OPMD (Becher et al., 2000; Shanmugam et al., 2000). In other cell culture models of polyalanine diseases, aggregates are formed within the nuclei and cytoplasm (Albrecht et al., 2004), and it is thought that transcription is affected by the expanded alanine repeats in the abnormal protein. Furthermore, intracellular aggregates and nuclear fragmentation were observed in cells expressing green fluorescent protein (GFP) fused to a polyalanine repeat peptide (Rankin et al., 2000; Oma et al., 2004).

Polyalanine peptides containing 15–20 residues form β-sheets in vitro (Shinchuk et al., 2005), especially in strongly alkaline medium (pH > 10). Two polyalanine peptides containing either 11 or 17 residues remarkably enhanced β-sheet content and formed fibrils (Giri et al., 2007). Aggregation of proteins with expanded polyglutamine repeats is accompanied by transition of α-helices to β-sheets, suggesting that aggregation of proteins in polyalanine diseases has a common molecular mechanism. However, the fundamental mechanism of aggregation resulting from expanded polyalanine repeats is unknown.

We previously showed that transfection of COS-7 cells with yellow fluorescent protein (YFP)-polyalanine constructs (Ala23 to Ala35) results in YFP localization in cytosol (Oma et al., 2007). In native gels, purified GST-fusion proteins, especially GST-Ala29 to GST-Ala35 products, were retained in the upper part of the running gel, whereas GST-Ala7 to GST-Ala23 products migrated into the running gel (Oma et al., 2007). The minimum pathological length of polyalanine proteins is assumed to be ~20 residues, but conformational changes caused by expanded polyalanine lengths differ in fusion proteins. Therefore, we investigated oligomer formation of fusion proteins containing greater than 20 polyalanine repeats using a sucrose density gradient (Dehay and Bertolotti, 2006). Moreover, because aggregation in the nucleus and cytoplasm may be caused not only by dysfunction of protein degradation systems, such as the ubiquitin-

Contract grant sponsor: Human Frontier Science Program.

*Correspondence to: Shoichi Ishiura, Department of Life Sciences, Graduate School of Arts and Sciences, University of Tokyo, Tokyo 113-8421, Japan. E-mail: cishiura@mail.ecc.u-tokyo.ac.jp

Received 16 November 2008; Revised 6 January 2009; Accepted 16 January 2009

Published online 19 March 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jnr.22052

proteasome system (Abu-Baker et al., 2003), but also by resistance of the oligomerized proteins to degradation, we investigated resistance of the oligomerized GST fusion proteins to trypsin digestion.

MATERIALS AND METHODS

Constructs

Polyalanine repeat sequences were synthesized by annealing double-strand oligonucleotides that were inserted into the EcoRV site of a pBlueScript KS(–) vector (TaKaRa, Shiga, Japan). The construct was digested with EcoRI, blunted with T4 polymerase (TaKaRa) to adjust the frame for polyalanine, and ligated with T4 ligase (TaKaRa). The repeat was inserted into the BglII and ApaI sites of a pEYFP-C1 vector (Clontech, Logan, UT) to allow expression of the alanine repeats fused to YFP. GST-Ala7, -Ala15, -Ala23, -Ala30, and -Ala35 constructs were prepared; alanine repeat sequences were obtained by digestion of the YFP-Ala7, -Ala15, -Ala23, and -Ala35 constructs with XhoI and BamHI and were inserted into the same sites of pGEX 4T-2 (Amersham Bioscience, Arlington Heights, IL), a vector expressing GST in *Escherichia coli*.

Preparation of GST-Fusion Proteins

A pGEX vector containing each polyalanine insertion was transformed into the BL21 (DE3) strain. An overnight culture of each transformant in Luria broth (LB) medium was diluted and agitated at 37°C for 1 hr (or until the optical density reached 0.4–0.6), and then 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added. During induction by IPTG, the culture was agitated at 25°C for 3 hr. The bacterial cells were collected by centrifugation; suspended in 1/20 LB medium volume of sonication buffer containing 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), 0.2 mM phenylmethylsulfonyl fluoride (PMSF), and 1/1,000 volume of protease inhibitor cocktail (Sigma, St. Louis, MO); and sonicated on ice 10 times for 10 sec each. Triton X-100 was added to a final concentration of 1%, and the suspension was incubated on ice for 30 min. The lysate was subjected to affinity purification with glutathione Sepharose 4B (Amersham Bioscience) according to the manufacturer's protocol.

Native PAGE Assay

GST fusion proteins were eluted from Sepharose beads with buffer containing 50 mM Tris-HCl (pH 7.5), 30 mM glutathione, and 50 mM NaCl. After elution, GST fusion proteins were dialyzed against PBS. Protein concentration was determined by the Lowry method using the DC protein assay kit (Bio-Rad, Hercules, CA). Oligomerization of GST-Ala23 in PBS occurred during incubation at 37°C for 24 hr. Native PAGE was performed without SDS or reducing agents. The protein was stained with Coomassie brilliant blue.

Sucrose Density Gradient Sedimentation

COS-7 cells were transiently transfected with the YFP-Ala7, -Ala15, -Ala23, and -Ala35 plasmids. After transfection for 48 hr, the cells were harvested in PBS containing 1 mM DTT, 0.2 mM PMSF, and 1/1,000 volume of protease inhibitor cocktail and then sonicated on ice for 15 sec, and the lysate was centrifuged. The supernatant fraction containing the cleared lysate was loaded intermittently (600 μl per application) onto a 9.0-ml sucrose density gradient (0.1–1.5 M) and centrifuged for 17 hr at 250,000*g* at 4°C using a P40ST rotor (Hitachi) and a CP70MX (Hitachi) ultracentrifuge. Thirteen fractions of ~800 μl were collected in each gradient from top (fraction 1) to bottom (fraction 13), and, after acetone concentration, they were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), followed by immunoblotting for YFP. YFP was quantified by calculating it as a percentage of the total amount of protein loaded on the gradient. As markers for the density centrifugation, bovine serum albumin (66 kDa), β-amylase (200 kDa), and ferritin (440 kDa) were used, and the proteins were stained with Coomassie brilliant blue. After elution from the Sepharose beads, dialysis in PBS, and incubation at 37°C for 24 hr, GST-Ala7, -Ala15, -Ala23, and -Ala35 were ultracentrifuged. Immunoblots were performed with a GST antibody.

Degradation of GST-Fusion Proteins With Trypsin

The GST-fusion proteins were incubated with trypsin (Sigma) at 37°C in buffer containing 2 mM $CaCl_2$ at an enzyme-to-substrate ratio of 1:40 (w/w). Digestions were terminated by the addition of SDS sample buffer. The degree of proteolysis was determined by SDS-PAGE, followed by immunoblotting with an anti-GST antibody. In this assay, GST-Ala23 was incubated at 37°C for 1 week. Without acetone concentration, GST-Ala35 was fractionated by centrifugation for 7 hr at 252,000*g* at 4°C with a TLS55 rotor (Beckman Coulter) and an Optima TLX (Beckman Coulter). GST-Ala35 was loaded intermittently (120 μl per application) onto a 1.8-ml sucrose density gradient (0.1–1.5 M). Thirteen fractions of ~160 μl each were collected in each gradient from top (fraction 1) to bottom (fraction 13). Fractions 7–13 of GST-Ala35 were treated with trypsin.

RESULTS

Polyalanine Expansion Causes Oligomerization of YFP Fusion Protein in COS-7 Cells

We previously reported that YFP-fused elongated polyalanine stretches (>23 residues) were localized in cytosol (Oma et al., 2007). This localization may be caused by oligomerization of the YFP-fused elongated polyalanine stretches. To analyze this oligomerization further, total cell lysates of COS-7 cells expressing YFP-A7, -A15, -A23, and -A35 constructs were fractionated on sucrose gradients (Fig. 1). YFP-A7 and -A15 were recovered from light fractions 1–5; in contrast, the ratio of fractions 1–4 to the total was 38%, and that of YFP-A35 was 22%. Therefore, we conclude that YFP-fused proteins containing >20 polyalanine residues facilitate oligomerization in COS-7 cells.

GST-Polyalanine Fractionation

Previously, we showed length-dependent differences between polyalanine-containing proteins on native-