1760–1775 *Nucleic Acids Research, 2017, Vol. 45, No. 4* *Published online 28 November 2016*
doi: 10.1093/nar/gkw1158

Autoregulation of MBNL1 function by exon 1 exclusion from *MBNL1* transcript

Patryk Konieczny, Ewa Stepniak-Konieczna, Katarzyna Taylor, Łukasz J. Sznajder and Krzysztof Sobczak*

Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznan, Poland

Received August 02, 2015; Revised October 27, 2016; Editorial Decision November 03, 2016; Accepted November 07, 2016

ABSTRACT

Muscleblind-like proteins (MBNLs) are regulators of RNA metabolism. During tissue differentiation the level of MBNLs increases, while their functional insufficiency plays a crucial role in myotonic dystrophy (DM). Deep sequencing of RNA molecules cross-linked to immunoprecipitated protein particles (CLIP-seq) revealed that MBNL1 binds to *MBNL1* exon 1 (e1) encoding both the major part of 5′UTR and an amino-terminal region of MBNL1 protein. We tested several hypotheses regarding the possible autoregulatory function of MBNL1 binding to its own transcript. Our data indicate that MBNLs induce skipping of e1 from precursor *MBNL1* mRNA and that e1 exclusion may impact transcript association with polysomes and translation. Furthermore, e1-deficient protein isoform lacking the first two zinc fingers is highly unstable and its EGFP fusion protein has severely compromised splicing activity. We also show that *MBNL1* can be transcribed from three different promoters and that the transcription initiation site determines the mode of e1 regulation. Taken together, we demonstrate that MBNL proteins control steady-state levels of MBNL1 through an interaction with e1 in its precursor mRNA. Insights from our study open a new avenue in therapies against DM based on manipulation of the transcription initiation site and e1 splicing of *MBNL1* mRNA.

INTRODUCTION

The family of Muscleblind-like (MBNL) proteins controls RNA metabolism in several distinct ways. These include regulation of splicing of alternative exons and selection of a polyadenylation site in pre-mRNAs, influencing stability and differential localization of mRNAs based on their 3′UTR length, and processing of miRNAs (1–6). MBNLs recognize their target RNAs using four zinc fingers (ZnFs) arranged in two tandems linked via a flexible linker (7–9). ZnFs are located at the amino-terminus of MBNLs, while the following regions determine their cytoplasmic and/or nuclear localization (10–12), and dimerization (11,13).

MBNLs ZnFs recognize target transcripts through the common YGCY motif [where Y is either U or C; (4,14–17)], preferentially containing unpaired Y bases in a semistable RNA structure (18). The interaction seems to be also influenced by the motif frequency in the close regional proximity, as the majority of targets have either two or more binding sites for MBNLs [reviewed in (19)]. At least two theories based on experimental data hint on how MBNLs interact with RNA. One, grounded on the structure of di-domains, indicates that RNA is looped around an MBNL molecule, with each ZnF recognizing one GpC dinucleotide within the YGCY motif (9). Another theory presumes that MBNLs interact with a wide variety of RNA targets owing to the flexible linker that positions ZnF tandems appropriately for such interaction (7). MBNLs binding position relative to the alternative exon determines whether an alternative exon is positively or negatively regulated. Binding within the alternative exon and upstream intronic regions generally facilitates exon skipping, while binding to downstream intronic regions promotes exon inclusion (4,16).

Among three family members, *MBNL1* mRNA is most prominently expressed in majority of tissues (19). The MBNL1 protein level increases during differentiation, which facilitates transitions from embryonic-to-foetal and foetal-to-adult splicing patterns in mammals (20–22). Functions of other MBNLs are less well-established. However, MBNL2 plays a key role in splicing of pre-mRNAs in the brain (14) and compensates for the loss of MBNL1 (4,23). *MBNL3* mRNA is hardly detectable in adult tissues (8,19,24) and seems to regulate RNA metabolism in early differentiation stages as well as in regenerative processes (17,25–27).

The key cellular roles of MBNLs are underscored by their functional depletion, as it occurs in myotonic dystrophy [DM; (28)]. In both types of the disease, DM1 and DM2, all MBNL proteins are sequestered by pathogenic transcripts

*To whom correspondence should be addressed. Tel: +48 61 829 5958; Fax: +48 61 829 5949; Email: ksobczak@amu.edu.pl

that carry multiple C(C)UG motifs in cell nuclei (8,22,29). In DM1, this is a result of a CTG triplet repeat expansion in the 3′UTR of *DMPK* and in DM2, a CCTG tetramer multiplication in the first intron of *CNBP*. The sequestration of MBNLs leads to hundreds of splicing alterations, akin to the embryonic splicing pattern (1,4,5,14,15,30,31). Some of the misspliced mRNAs directly linked to the DM phenotype include *CLCN1*, *INSR*, *PKM*, *BIN1*, *CACNA1S* and *DMD*. Upon translation, these mRNAs may generate non-fully functional protein isoforms in the adult. Particularly, their expression in the adult skeletal muscle causes myotonia, insulin resistance and muscle weakness, respectively (32–39). Decreasing amounts of functional MBNL proteins result in increasing splicing aberrations, with some exons responding to relatively small changes in MBNL content (30,40). This indicates that in physiological conditions the level of MBNLs should be precisely controlled.

In the present study, we show that MBNLs bind to the 5′-most region of *MBNL1* exon 1 (e1) that encodes both the 5′-untranslated region (5′UTR) and an amino-terminal part of the protein. We hypothesized that by binding to e1 of *MBNL1* pre-mRNA, MBNLs might regulate e1 exclusion, which could lead to the production of a non-fully functional protein without two ZnFs. We also postulated that binding of MBNLs to e1 of mature *MBNL1* mRNA could exert an effect on the transcript subcellular localization and/or its translation. Based on our data, we conclude that all three paralogs of MBNLs do indeed regulate MBNL1 expression by inhibition of e1 inclusion. The e1-devoid transcript concentrates in a cytosolic fraction lacking polysomes and has affected translational activity, while the protein product displays decreased stability. We also reveal that the transcription initiation site plays a pivotal role in MBNL-dependent regulation of e1 inclusion. Altogether, these processes fine-tune cellular steady-state levels of MBNL1.

MATERIALS AND METHODS

CLIP-Seq

CLIP-seq was performed using mouse skeletal muscle tissues as previously reported (4,41,42).

Preparation of radiolabeled *Mbnl1* e1 transcripts

The template for transcription reaction of 190 nt-long *Mbnl1* e1 was obtained in two PCR steps. The first 447 bp product was amplified using *F1*/*R1* primers based on mouse genomic DNA. The 447 bp product subsequently constituted a template for the second PCR (190 bp product) with *F2*/*R1* primers. The 447 bp fragment was also used as a template to obtain Mut1–3 of 190 bp-long *Mbnl1* e1 through overlap extension. First, two short PCR products harboring substitutions were amplified using pairs of primers for Mut1: *F3*/*R1* and *F1*/*R3*; Mut2: *F4*/*R1* and *F1*/*R4*; and Mut3: *F5*/*R1* and *F1*/*R5*. After gel purification, both products constituted a template for a PCR using *F1*/*R1* primers. 447 bp mutants 1, 2 and 3 were subsequently subjected to PCRs with *F2*/*R1* primers amplifying 190 bp mutants of *Mbnl1* e1. Transcription reaction and radio-labeling with [γ-P^{32}] ATP was performed as previously described (43). The template for transcription reaction and internal labeling of 873 nt-long *Mbnl1* e1 was obtained using *F2*/*R2* primers. For [α-P^{32}] UTP labeling, 3.5 μl out of 10 μl PCR product was incubated with 3.125 μM of [α-P^{32}] UTP (6000 Ci/mmol), 40 U RNasin Plus RNase Inhibitor (Promega), 0.5 mM NTPs (Invitrogen), 100 U T7 RNA Polymerase (Ambion), 1x T7 transcription buffer (Ambion) for 30 min at 37°C. The purification of radio-labeled RNA was performed as previously described (43). Primer sequences are listed in Supplementary Table S1.

Chemical and enzymatic analyses of *Mbnl1* e1 structure

The analyses were conducted as previously described (44), with slight modifications. Briefly, the ^{32}P-labeled 190 nt *Mbnl1* e1 was first subjected to denaturation and renaturation procedure in either buffer A containing 20 mM Tris–HCl pH 7.2, 80 mM NaCl, 2 mM $MgCl_2$ (ribonuclease T2) or in a commercial S1 buffer (Fermentas). Limited RNA digestion was initiated by mixing 5 μl of the RNA sample (∼0.5 pmol) with 5 μl of a probe solution containing S1 (final concentration 0.2, 0.3, 0.4 U/μl), T2 (final concentration 0.04, 0.05, 0.075 U/μl). An equal volume of a stop solution was added to stop all reactions. Cleavage products were separated in a 10% denaturing polyacrylamide gel along with the products of alkaline hydrolysis, lead ions digestion and T1 nuclease digestion. The alkaline hydrolysis ladder was generated by incubating labeled RNA in formamide at 99°C for 20 min. The T1 ribonuclease digestion was performed in 50 mM sodium citrate pH 4.3, 7 M urea and 0.1 or 1 U/μl of T1 at 55°C for 10 min. Lead ions ladder was generated in buffer A with 0.25 mM of Pb^{2+} at 55°C for 2 min.

Quantification of *Mbnl1* e1/protein interaction and its inhibition by AONs *in vitro*

Recombinant GST and His6-tagged MBNL1 have been previously described (43). Filter binding assay was performed in a 30 μl volume. To assess the MBNL1 affinity to RNA, 5′-labeled *Mbnl1* e1 190 nt-long WT and Mut1-3 transcripts (0.05 nM) were incubated with the indicated protein amounts (ranging from 0 to 200 nM) in buffer B containing 250 mM NaCl, 15 mM KCl, 50 mM, Tris–HCl pH 8, 0.05% Tween-20, 1 mM $MgCl_2$ at 37°C for 30 min. To estimate the inhibitory property of AONs, 0.05 nM of labeled transcript (*Mbnl1* e1 WT 190 nt or 873 nt) underwent a three step incubation with 10 μM of AONs, first at 90°C for 1 min, then on ice for 10 min and at 37°C for 25 min. Subsequently, the indicated concentrations of MBNL1 were added to each sample and incubated at 37°C for 25 min. 25 μl sample aliquots were loaded onto filter binding apparatus with stacked nitrocellulose (Protran BA 85, Whatman) and nylon (Hybond N+, Amersham) membranes prewetted in buffer B. The membranes were exposed to a phosphor screen and imaged on a FLA-1500 (FujiFilm). The signal was quantified with Multi Gauge software (FujiFilm). K_d of the RNA/MBNL1 complexes affected by different AONs was measured by Graph Pad using one site specific binding curve.

Plasmid preparation

EGFP_MBNL1 and *EGFPΔ_MBNL1* vectors

To obtain *EGFP_MBNL1_*41.4, Flag-tag was amplified using *F6*/*R6* primers on p*EGFP*_C1 vector template containing a coding sequence for *MBNL1* (–e5, +e7) (45). The purified PCR product was digested with BamHI and EcoRI and cloned in between BamHI and EcoRI sites of the p*EGFP*_C1_*MBNL1* (–e5, +e7) using T4 DNA ligase (Invitrogen). *EGFP_MBNL1_*43.4 was generated based on *EGFP_MBNL1_*41.4 and pEGFP_C1_*MBNL1* (+e5, –e7) (45) using EcoRI and NdeI. *EGFP_MBNL1_*41.2 was obtained from *EGFP_MBNL1_*41.4 with *F7*/*R7* primers. The amplified fragment was digested with HindIII and XhoI and cloned into HindIII and XhoI digested *EGFP_MBNL1_*41.4. *EGFP_MBNL1_*43.2 was generated from *EGFP_MBNL1_*41.2 and *EGFP_MBNL1_*43.4, following digestion with NdeI and HindIII and ligation of corresponding fragments. *EGFP* vector was generated by removing the coding sequence of *MBNL1* from *EGFP_MBNL1_*41.4 with BamHI and XhoI and ligation with annealed oligos coding for the Flag-tag *F8*/*R8*.

*EGFPΔ_MBNL1_*41.2 was obtained by digesting *EGFP_MBNL1_*41.2 with AgeI and BspEI and religation of the appropriate fragment. To generate *EGFPΔ_MBNL1_*41.4, *EGFP_MBNL1_*41.4 was first processed as *EGFP_MBNL1_*41.2 above. To add an ATG codon, a fragment was PCR-amplified with *F9*/*R9* primers using the newly generated vector as a template. The vector sequence was replaced with the amplified fragment with NheI and HindIII. *EGFPΔ_MBNL1_*43.4 was generated from *EGFPΔ_MBNL1_*41.4 and *EGFP_MBNL1_*43.4 using HindIII and EcoRI while *EGFPΔ_MBNL1_*43.2 from *EGFPΔ_MBNL1_*41.2 and *EGFPΔ_MBNL1_*43.4 using HindIII and NheI.

Luciferase constructs

To obtain *Mbnl1_*5′UTR2_*Luc2* Luc vector, pmirGLO (Promega) was digested with NheI and XbaI, ligated and then annealed oligos (*F10*/*R10*) where cloned into a HindIII site in front of *Luc2*. +e1_Luc and –e1_Luc constructs were obtained in a few sequential steps. First, *Mbnl1* fragments were amplified with *F11*/*R11* primers from C2C12 cDNA, and then the 5′UTR products were re-amplified with *F12*/*R12.1* or *F12*/*R12.2* primers and cloned into Luc vector with NheI and XhoI. Such obtained +e1_Luc and –e1_Luc constructs contained 132 nucleotides of e5′UTR2 and the full 5′UTR sequence of either e1 (875 nucleotides) or e2 (102 nucleotides) between the promoter and *Luc2*, respectively. The constructs retained short vector/multiple cloning site sequences downstream of the promoter and upstream of *Luc2*.

Luciferase *Mbnl1_*e1_5′UTR2_*Luc2* constructs were generated based on pGL4.51 (Promega) and pmirGLO vectors. *Mbnl1* fragment was PCR-amplified using *F13*/*R11* primers on a template of RNA obtained from mouse NIH/3T3 cells, and then a shorter fragment was generated using *F14*/*R2* primers. The sequence contained 65 nucleotides of e5′UTR2 and 863 nucleotides of e1. One ATG codon in e5′UTR2 was mutated [G to T; chr3: 60501374 (mm10)] using the forward primer to prevent expression of a shorter protein that could delay the translation of luciferase. Obtained *Mbnl1* fragment was digested with HindIII and cloned into HindIII digested pGL4.51 vector. To obtain Ctrl construct, a PCR fragment was amplified with *F15*/*R15* primers, digested with MluI and XhoI, and cloned into MluI and XhoI-digested pmirGLO. Δ1 and Δ2 were obtained by digestion of Ctrl with SmaI and XbaI. The ends were blunted with the Klenow fragment (Thermo Scientific) prior to the ligation. Deleted sequences in Δ1 and Δ2 constructs are listed in Supplementary Table S2. An empty vector lacking *Mbnl1* (Emp) was generated by removing the *Mbnl1* fragment from Ctrl vector with HindIII digestion.

Min_e1_WT and Min_e1_Mut3 minigenes

Min_e1_WT minigene was prepared by subcloning a ∼2.0 kb murine genomic *Mbnl1* fragment spanning e1 with parts of flanking introns into human *cTNT* minigene [(46), a kind gift from M. Disney; The Scripps Research Institute]. *Mbnl1* fragment was amplified from C2C12 genomic DNA using a two-step nested PCR with *F16*/*R16* and *F17*/*R17* primers. The second round PCR product was cloned into *cTNT* minigene backbone with BamHI and SalI, replacing alternative *cTNT* e5. Min_e1_Mut3 was prepared using a three-step PCR reaction. First, two PCR reactions were performed to introduce point mutations abolishing MBNL1 binding in e1 with *F17*/*R18* and *F18*/*R17* primers. Both PCR products were annealed and extended and then subjected to PCR 3 with *F17*/*R17* primers. The ∼2.0 kb PCR product was cloned into BamHI/SalI digested *cTNT* vector.

All clones and PCR bands were verified by sequencing. PCR fragments used for cloning were amplified with either Platinum Taq DNA Polymerase High Fidelity, Pfx50 DNA polymerase (Invitrogen), or Kappa HiFi polymerase (Kapa Biosystems). *F*/*R* sequences are listed in Supplementary Table S1.

DM patients

RNA from skeletal muscles of non-DM and DM1 patients used in all experiments, except for semi-quantitative multiplex PCR (transcription start site-specific mRNAs) and quantitative PCR, were kindly provided by Charles Thornton (University of Rochester Medical Center). This RNA panel was obtained from two females and one male diagnosed from skeletal muscles with more than 500 CTG repeats in the *DMPK* gene. The remaining DM1 patient samples were cDNA (also provided by Charles Thornton, University of Rochester Medical Center) obtained from three females and one male, diagnosed based on the number of CTG repeats in the *DMPK* gene in the peripheral blood (1520, 480, 430 and 270 repeats, respectively). Biopsies of skeletal muscles from non-DM and DM2 patients were a kind gift from Anna Kamińska and Anna Łusakowska (Medical University of Warsaw, Poland).

Cell culture

Human HeLa, monkey COS7, and mouse NIH/3T3 cell lines were grown at 37°C in a high glucose DMEM medium with L-Glutamine (Lonza) supplemented with 10% foetal bovine serum (Sigma) and 1% antibiotic/antimycotic (Sigma), in a humidified incubator containing 5% CO_2. Fibroblasts derived from DM1 patients expressing *DMPK* transcript with ∼1000 CUG repeats (line GM04033; Coriell Cell Repositories) and control fibroblasts obtained from non-DM patients (line GM07492; Coriell Cell Repositories) were grown in EMEM medium with L-glutamine (Lonza) supplemented with 10% foetal bovine serum, 1% antibiotic/antimycotic, and 1% non-essential amino acids solution (Sigma) in a humidified incubator containing 5% CO_2, at 37°C.

Transfection of plasmids and knock-down of *MBNL*s with siRNAs

Plasmids were delivered to HeLa cells with X-tremeGENE HP DNA Transfection Reagent (Roche) according to the manufacturer's instructions, and at a ratio of 2 μl X-tremeGENE HP DNA Transfection Reagent per 1 μg plasmid. Briefly, 2 μg of plasmid and 4 μl of X-tremeGENE HP DNA Transfection Reagent were added per single well of a 12-well plate, and the cells were harvested 36–48 h post-transfection. To test the splicing of *cTNT* e5, 200 ng of the minigene was delivered to each well. X-tremeGENE HP DNA Transfection Reagent (Mock)- and *EGFP*-treated HeLa cells were used as controls. For immunoblotting experiments, cells were grown in a 6-well plate and a double amount of plasmid and the transfection reagent were used. Annealed siRNA oligos were delivered to fibroblasts using Lipofectamine 2000 (Invitrogen) per manufacturer's instructions. Oligos targeting human *MBNL1* (47) and *MBNL2* (48) were obtained from Future Synthesis and RiboTask, respectively. The cells were harvested 72 h post siRNA delivery. Lipofectamine (mock)- and AllStars Negative Control siRNA (Qiagen; Ctrl)-treated fibroblasts were used as controls.

Splicing and expression analyses of precursor and mature mRNA

RNA was isolated using TRI reagent (Sigma) per manufacturer's instructions. cDNA was synthesized using either Superscript III Reverse Transcriptase kit (Invitrogen) or GoScript Reverse Transcription System (Promega) according to the manufacturer's protocols. Standard PCRs were performed using GoTaq DNA polymerase (Promega). The PCR products were amplified using 30–38 cycles, depending on the amount of RNA used for RT as well as primer dilution. PCR images were captured using G:Box EF2 (Syngene) and the splicing efficiency and mRNA expression were assessed using GeneTools image analysis software (Syngene). Real-time quantitative PCRs were performed in a 7900 HT Fast Real-Time PCR System (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems) according to the manufactures' instructions. Quantitative PCR reactions were performed using at least three distinct cDNAs per sample type, each in three technical replicates. Raw Ct data were analyzed in Microsoft Excel using $2^{-\Delta\Delta Ct}$ method. Results represent a relative expression of the analyzed target gene compared to an indicated reference gene, and normalized to averaged control/calibrator samples. Sequences of primers used for splicing and expression analyses are listed in Supplemental Experimental Procedures in Supplementary Tables S3 and S4.

Co-immunoprecipitation (Co-IP)

Four microgram of *Mbnl1*_e1_5′UTR2_*Luc2* Ctrl, Δ1 or Δ2 construct were co-delivered to HeLa cells grown in 10 cm plates along with 12 μg of *EGFP_MBNL1*_41.4 and X-tremeGENE HP DNA Transfection Reagent. After 36 h, the medium was removed, 1 ml of ice-cold PBS was added and the cells were cross-linked (150 mJ/cm^2 at 254 nm, UVP CL-1000) on ice to covalently bind RNA/protein complexes. The cells were then harvested with a cell lifter and transferred into a test tube. After centrifugation, the cell pellet was kept at –80°C until use. The cells were lysed in 0.5 ml lysis buffer (50 mM Tris–HCl, pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, benzamidine, PMSF). The lysate was subsequently passed ten times through a 25 G needle and spun at 15 000 x g for 15 min to clear it. 1 μl Turbo DNase (Ambion) and 12.5 μl RNasin Plus RNase Inhibitor (Promega) were added and the EGFP_MBNL1_41.4/RNA complexes were immunoprecipitated on Dynabeads Protein G (Invitrogen) coated prior to the incubation with 2 μg of FLAG-M2 antibody (Sigma, F1804) for 60 min at RT. The co-immunoprecipitation was performed on a rotator for 2 h at 4°C. The beads were then washed thrice with high-salt buffer (50 mM Tris–HCl, pH 7.4; 1 M NaCl; 1 mM EDTA; 1% NP-40; 0.1% SDS; 0.5% sodium deoxycholate) and RNA was isolated using the standard procedure with TRI reagent.

Luciferase assay

HeLa cells were grown on a Nunc F96 MicroWell Black Polystyrene Plate (137101, Thermo Scientific). Fifty nanogram of *Mbnl1*_e1_5′UTR2_*Luc2*/*Mbnl1*_5′UTR2_*Luc2* plasmid were co-delivered with 150/350 ng of *EGFP_MBNL1*/*EGFP or EGFP_MBNL1* construct with X-tremeGENE HP DNA Transfection Reagent. The cells were lysed and luminescence of firefly and *Renilla* luciferases was measured consecutively using Dual Luciferase Reporter Assay System (Promega), infinite F200 PRO, and i-control 1.8 SP1 microplate reader software (Tecan). Data sets obtained for Ctrl and Δ1 plasmids were normalized to the Emp plasmid data set.

Subcellular fractionation

Subcellular fractionation was performed using a standard procedure (detailed protocol available on the Abcam website). Briefly, HeLa or NIH/3T3 cells from two 10 cm plates were lysed in 1 ml of subcellular fractionation buffer [250 mM sucrose, 20 mM Hepes (pH 7.4), 10 mM KCl, 1.5 mM

$MgCl_2$, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, SigmaFAST Protease inhibitor Cocktail (Sigma)] and passed 10 times through a 25 G needle. After 20 min incubation on ice, the nuclear pellet was harvested by centrifugation at 700 x g for 5 min at 4°C. The supernatant was subsequently centrifuged at 10 000 x g for 10 min. The supernatant was transferred to a new tube and centrifuged again at 100 000 x g for 1 h at 4°C to pellet the membrane fraction. RNA was isolated from the total cell extract and the nuclear, membrane and cytosolic (the remaining supernatant) fractions using TRI reagent.

Immunoblotting

Cells were washed in PBS, lifted using a cell scraper and transferred into a test tube in ice-cold PBS. Following centrifugation at 5000 x g for 5 min at 4°C, the cell pellets were frozen at –20°C until further use. The cells were homogenized in 40 μl lysis buffer (see the IP protocol) with a micropestle, sonicated for 5 cycles (30 s on/ 90 s off) using Bioruptor Plus (Diagenode), and spun at 15 000 x g for 15 min at 4°C to clear the lysate. The lysate was then mixed with the standard sample buffer (4×), heated for 5 min at 95°C, and loaded on an 8–10% polyacrylamide gel. After electrophoresis (Mini-PROTEAN Tetra System, Bio-Rad), proteins were transferred onto the nitrocellulose membrane (Protran BA-85, Whatman) for 1 h at 100 V at 4°C, which was subsequently blocked in 5% skim milk in 0.1% PBS–Tween for 1 h at RT. Rabbit anti-human MBNL1 primary antibody (a kind gift from C. Thornton; University of Rochester Medical Center) was diluted in the blocking solution and incubated with the membrane O/N at 4°C. Mouse anti-human GAPDH antibody (sc-47724; Santa Cruz) was diluted in PBS-T and incubated with the membrane for 1 h at RT. The blots were washed 3 × 5 min each in 0.1% PBS-Tween and incubated with a secondary goat anti-rabbit (A9169, Sigma) or anti-mouse (12–349, Millipore) IgG peroxidase conjugate for 45 min at RT. After the subsequent washing steps, the signal was developed using Pierce ECL Plus Western Blotting Substrate or SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific). The images were captured in G:Box Chemi-XR5 (Syngene) and the densitometrical analysis was performed using GeneTools image analysis software (Syngene).

Microscopy

Images were taken with Axio Observer.Z1 microscope equipped with AxioCam MRm camera, filter set 09, A-Plan 10×/0.25 Ph1 objective (Zeiss), and AxioVs40 module.

RESULTS

MBNL1 binds the 5′-most region of exon 1 in *MBNL1* pre-mRNA

To reveal novel Mbnl1 targets, we sequenced RNA molecules obtained from UV cross-linking and immunoprecipitation procedure (CLIP-Seq) that involved mouse skeletal muscles and a polyclonal antibody against Mbnl1 (42,45). One of the most abundant CLIP-Seq clusters was identified in the 5′-most fragment of *Mbnl1* exon 1 (e1, Figure 1A). We observed no other CLIP-Seq clusters in e1 or the adjacent introns, which suggested that Mbnl1 could autoregulate its function by binding to this specific region contained within the precursor and mature mRNA. Furthermore, MBNL1 could induce formation and binding of a circular RNA (circRNA) composed of e1 (49). Considering these data, we hypothesized that these interactions could regulate nuclear maturation, stability, cellular localization or translational activity of the *MBNL1* transcript. Based on the sequence analysis of mouse and human e1, we singled out nine conserved YGCY MBNL-specific binding motifs contained in the CLIP-Seq cluster as potential interaction sites (marked green in Figure 1A and Supplementary Figure S1A).

We used a set of *in vitro* filter binding assay experiments to validate MBNL1 binding to the selected YGCY motifs within the CLIP-Seq cluster of *Mbnl1* e1. The 190 nt-long RNA molecules contained either non-mutated or point-mutated YGCY motifs (Figure 1B and Supplementary Figure S1B). These experiments revealed high binding affinity of recombinant MBNL1 to the wild-type (WT) e1 fragment (K_d = 3 nM) when compared to mutant RNAs (Mut1–Mut3), in which different number of YGCY sequence motifs were mutated. Particularly, Mut3 with point mutations in five consecutive YGCY sites had K_d over 35 times higher than the corresponding WT e1 fragment (Figure 1B). These data were further corroborated using a set of antisense oligonucleotides (AON1–AON9) that bound to and blocked various potential MBNL1 interaction sites in e1 (Supplementary Figures S1C and S2). AON4 that blocked the same YGCY motifs, shown in Figure 1B to markedly affect MBNL1 binding when mutated, resulted in the lowest MBNL1 affinity to the WT e1 region. CLIP-Seq data revealed also that three out of the five consecutive YGCY motifs (located in the region facilitating MBNL1 binding) contained characteristic for MBNL1 targets cross-linking induced mutations (CIMs), i.e. substitutions of C to T [Figure 1C, (4,50)].

Next, we analysed structural features of the WT e1 fragment, to which MBNL1 bound with high affinity. Limited digestion of 5′-end-labeled RNA with T2 RNase and S1 nuclease revealed that WT e1 contains three semistable hairpins (Figure 1C, h1–h3). One of them, h3, characterized by presence of pyrimidine-rich internal loops and bulges, contains YGCY motifs shown to facilitate MBNL1 binding. All G residues from the five consecutive GC steps are paired and separated by 1–5, mainly unpaired, pyrimidines. The cluster of three other YGCY motifs, which upon mutation did not significantly reduce MBNL1 binding affinity (Mut1) locate in h2, a more stable hairpin structure.

All MBNL paralogs induce skipping of e1 in *MBNL1* pre-mRNA

To test putative MBNL-dependent regulation of e1 alternative splicing, we first estimated relative amounts of *MBNL1* mRNAs initiated from three different transcription start sites (T1–T3, Figure 2A). Their prediction was based on trimethylation of histone H3 at lysine 4 (H3K4Me3) in nine cell lines, including human skeletal muscle myoblasts [Sup-

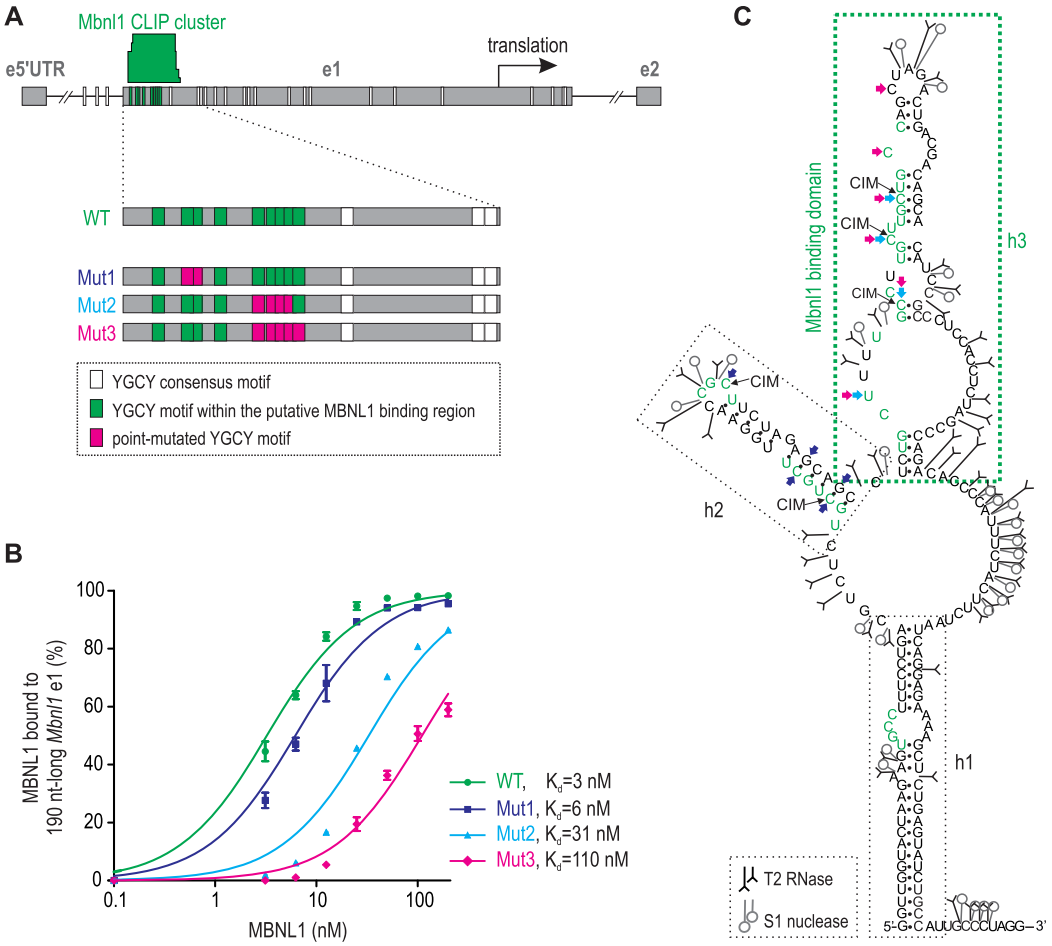

Figure 1. Mbnl1 protein interacts with the 5′-most part of e1 of its own mRNA. (**A**) Mbnl1 CLIP cluster in e1 of mouse *Mbnl1* and schematic representation of *Mbnl1* RNA fragments used in **B**. (**B**) A filter binding assay showing an *in vitro* interaction of MBNL1 with the selected 190 nt-long region of e1 (WT) and reduced interaction with the corresponding fragment carrying point mutations in YGCY motifs (Mut1–Mut3). Mean values from two experiments ± standard deviation (SD) are shown on the graph. (**C**) The proposed structure of the e1 fragment based on the RNA digestion with two enzymatic structural probes, T2 RNase and S1 nuclease. The nuclease digestion sites are marked as indicated in the legend. Mutation sites of Mut1-Mut3 are marked with color-coded arrows as in **A** and **B**. CLIP-seq CIMs are indicated with black arrows.

plementary Figure S3A, (51,52)], and mapped transcription start sites and their usage in human and mouse cells and tissues [Supplementary Figure S3A and B, (53,54)]. Each of the three transcription start sites is associated with a different exon constituting a part of *MBNL1* 5′UTRs (e5′UTR1–e5′UTR3) that could be potentially spliced to e1 (Supplementary Figure S4A). Interestingly, the length of introns between e5′UTRs and e1 ranges from approximately 50 kb to only 0.2 kb (Figure 2A). Semi-quantitative RT-PCR analyses with primers specific to e5′UTRs revealed predominant amounts of T2 mRNA in mouse C2C12 myoblasts and human foetal and adult skeletal and cardiac muscle tissues (Figure 2B and C), with T2 and T3 transcripts increasing in the course of myoblast differentiation three and five times, respectively (Figure 2B). In other tissues, expression of e5′UTR2 and e5′UTR3 was more balanced while contents of T1 mRNA was marginal in all samples, except the liver, spleen and thymus, where the three tested transcriptional start sites were similarly operational (Figure 2C).

Based on the level of *MBNL1* transcripts initiated from different transcription start sites, we tested if e1 is alternatively spliced in skeletal muscles from DM patients and the HSA^{LR} mouse, a DM model deficient in functional MBNLs due to their sequestration (55), using a forward primer specific to e5′UTR2. In both, human and mouse tissues with functional MBNL protein depletion, we detected a substantial shift towards *MBNL1* mRNA with e1 inclusion (Figure 3A). Importantly, a group of tested DM1 patients showing severe pathology had a much higher percentage of e1 inclusion than moderately affected group of DM2 patients (Figure 3A). As our data (Figure 2B) and several other studies indicated that *MBNL1* mRNA and protein levels rise during differentiation (2,19,21,22), we compared inclusion of e1 in various human foetal and adult tissues (Supplementary Figure S4B). Our analyses show a tendency towards reduced inclusion of e1 during differentiation, especially in tissues in which *MBNL1* expression is prominently induced during differentiation, such as heart and skeletal muscle (19).

Next, we delivered into HeLa cells either *EGFP* or one of the two *EGFP_MBNL1* expression constructs differing in inclusion of exon 5 [e5; 41.4 (–e5) and 43.4 (+e5)] that de-

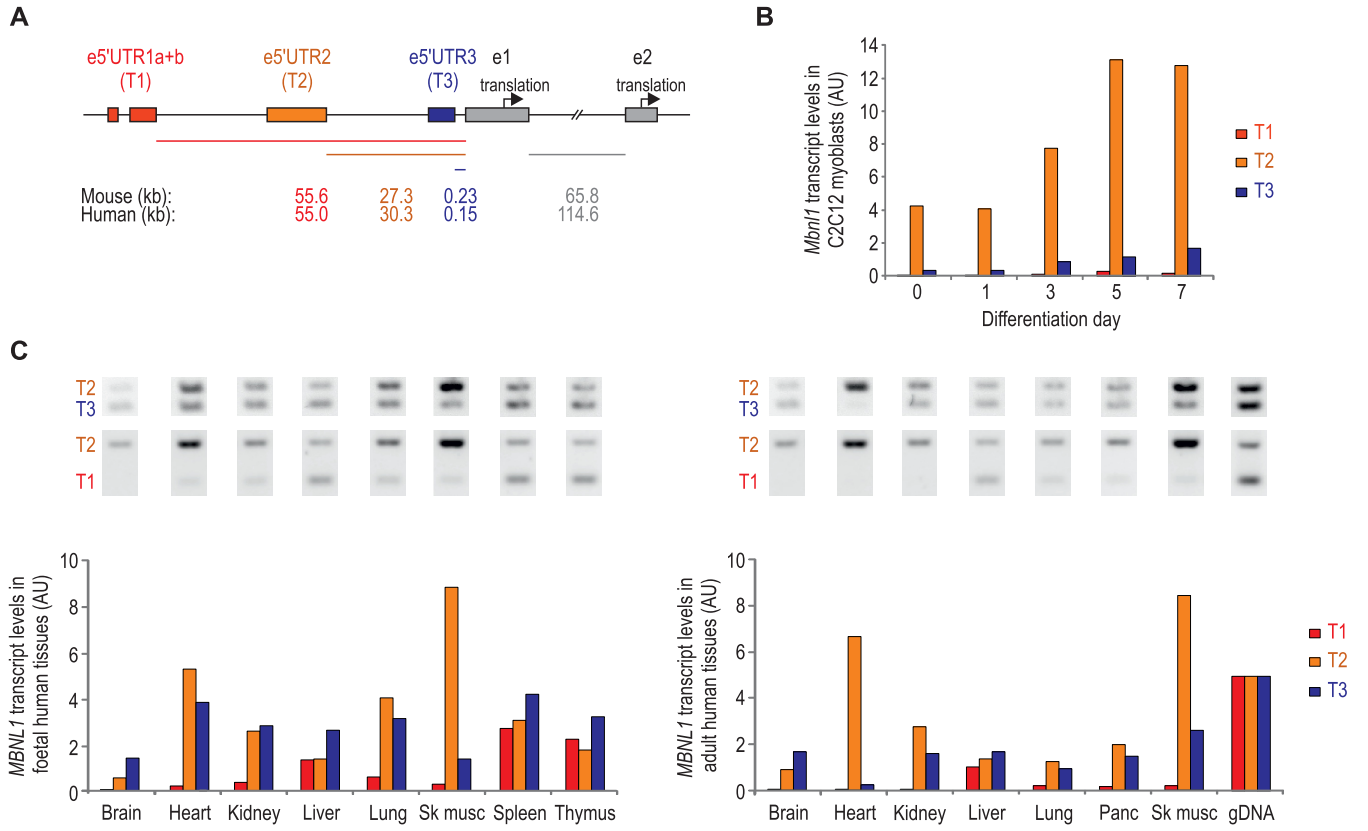

Figure 2. *MBNL1* expression is predominantly driven from T2 in cardiac and skeletal muscles. (**A**) Schematic representation of a genomic DNA fragment spanning transcription start sites (T1–T3) and their associated 5′UTR exons (e5′UTR1–e5′UTR3) of *MBNL1* that could be potentially spliced to either e1 or e2. Intron lengths between e5′UTRs and e1, as well as e1 and e2 in either the mouse or human pre-mRNAs, are indicated. (**B** and **C**) Semi-quantitative multiplex PCRs showing *Mbnl1* T1–T3 levels in C2C12 myoblasts following induction of differentiation (**B**) and in foetal (left panel) and adult (right panel) human tissue cDNA panels from Clontech (**C**). Specific signals for each of the e5′UTR were related to signals obtained from amplification of genomic DNA isolated from either mouse C2C12 myoblasts (**B**) or human HeLa cells (**C**). Signals from **B** were additionally related to *Gdp2* mRNA signals obtained from a semi-quantitative multiplex RT-PCR. Values were obtained based on quantification of bands intensities from one of two gels. Sk musc, skeletal muscle; Panc, pancreas; gDNA, genomic DNA. AU indicates arbitrary units.

termines exclusive nuclear localization of MBNL1 (10,11). Upon forced expression of either of the EGFP_MBNL1 proteins, we observed a marked increase in *MBNL1* mRNA isoform lacking e1 and, to a lower extent, also e3 (Supplementary Figure S4C). We also observed this effect in *MBNL2* pre-mRNA, however, in this case the MBNL1-induced skipping of *MBNL2* e1 was markedly less efficient (Supplementary Figure S4C). Previous CLIP-Seq analyses indicated that all three MBNL proteins interact with RNA through the YGCY motif (4,14,17), which prompted us to compare *MBNL1* e1 splicing upon forced expression of MBNL1, 2 and 3. Importantly, all tested MBNL proteins induced skipping of e1 to a similar extent (Figure 3B). To further confirm the autoregulatory splicing of e1 in *MBNL1* pre-mRNA, we knocked-down *MBNL1*, *MBNL2* or both in human fibroblasts derived from non-DM and DM1 patients (Figure 3C and Supplementary Figure S4D). MBNL knock-down induced a splicing pattern opposite to the one observed after MBNL overexpression, i.e. increased inclusion of e1 in mature *MBNL1* transcripts.

To verify whether *MBNL1* splicing of e1 is directly induced by MBNL1 binding to e1, we generated *MBNL1* e1 minigene (min_e1_WT) based on the previously described human *cTNT* plasmid (46), in which alternative *cTNT* e5 was replaced by a genomic fragment carrying *Mbnl1* e1 with adjacent intronic regions, giving rise to upstream and downstream introns of 0.9 and 0.83 kb, respectively. Surprisingly, following delivery of min_e1_WT to HeLa cells, we observed that e1 can be either included into mRNA or processed into a central intron and shorter peripheral exons, one of which contained a 5′ region of e1 (e1_5′) with the entire putative MBNL1 binding site (Figures 1A and 3D). The observed alteration in splicing of e1 presumably results from the fact that e1 is relatively long (1049 bp) and that in *MBNL1* pre-mRNA it is surrounded by exceptionally long introns (Figure 2A). Nonetheless, forced expression of MBNL1 induced a marked exclusion of e1_5′, reflecting MBNL-dependent skipping of e1 in its natural context (Figure 3D and E). Furthermore, mutation of the five consecutive YGCY motifs in e1_5′ (min_e1_Mut3; see also Figure 1A and Supplementary Figure S1B) abolished skipping of this exon upon delivery of MBNL1, confirming the autoregulatory-splicing role of MBNL1 by binding to e1.

Next, we analyzed whether e1 splicing is sensitive to MBNL1 protein when transcription of *MBNL1* is initiated from T3, the second most active transcription start site in

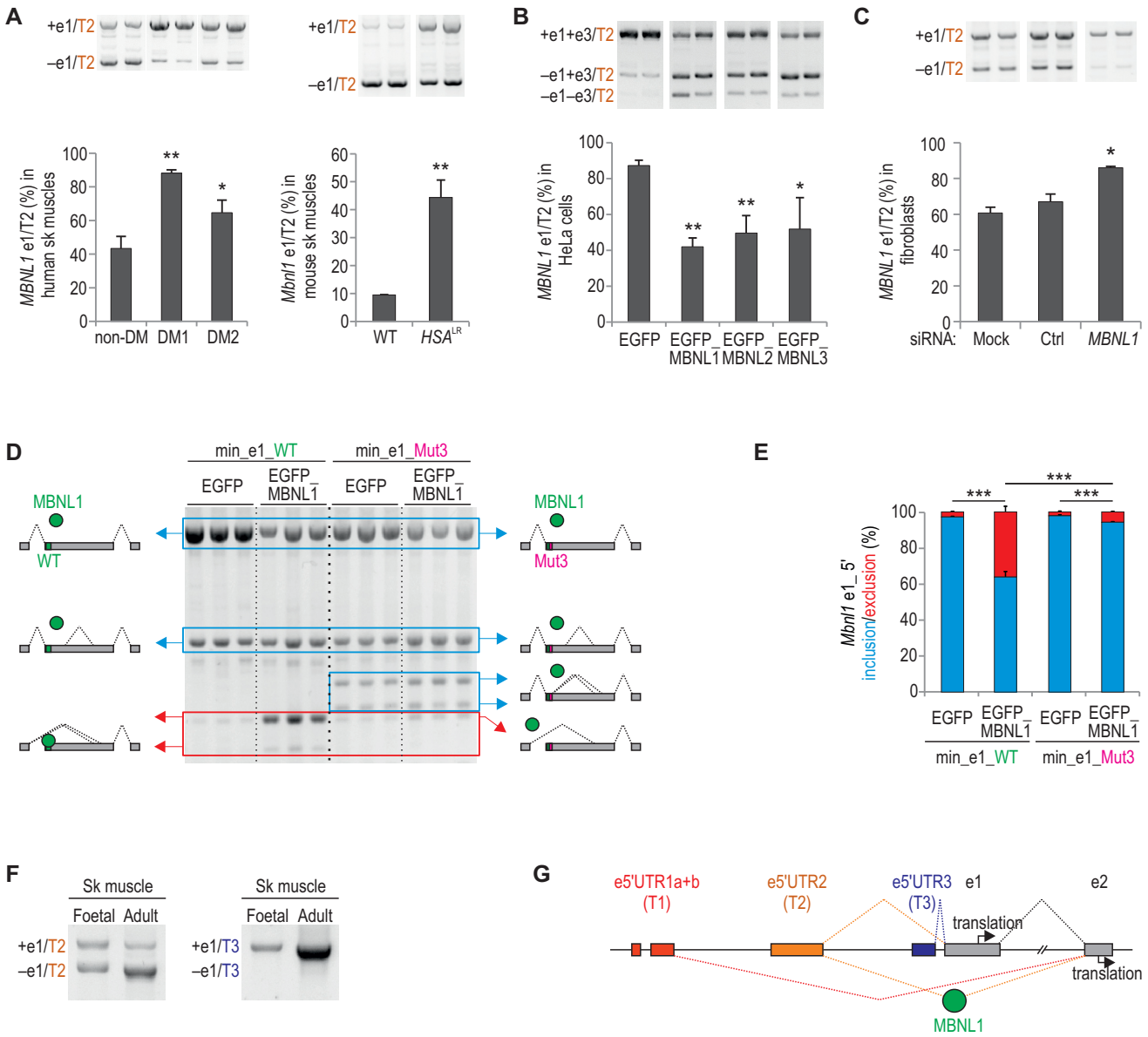

Figure 3. Splicing of *MBNL1* e1 is autoregulated by MBNL proteins in pre-mRNA transcribed from T2. (**A**) RT-PCR analyses of *MBNL1* e1 splicing in skeletal muscles of non-DM ($n = 3$), DM1 ($n = 3$) and DM2 ($n = 5$) patients (left panel), and WT ($n = 3$) and HSA^{LR} ($n = 3$) mice (right panel). (**B**) E1 splicing in *MBNL1* pre-mRNA following forced expression of EGFP_MBNL1, 2 and 3 (–e5, –e7) in HeLa cells. (**C**) Increased inclusion of e1 in *MBNL1* mRNA upon administration of siRNA against *MBNL1* in fibroblasts. 25 nM of either control AllStars (Ctrl) or *MBNL1*-specific siRNA was administered. $n = 3$ per each treatment group. (**D**) RT-PCR analyses of *Mbnl1* e1 splicing in HeLa cells following co-delivery of wild-type (min_e1_WT) or mutant (min_e1_Mut3) minigenes containing *Mbnl1* e1 with either a plasmid encoding EGFP or EGFP_MBNL1 (41.4). Note abolished skipping of the most-5′ part of e1 (e1_5′) in min_e1_Mut3 upon forced expression of EGFP_MBNL1 (red frame). (**E**) Quantification of the bands from the correspondingly color-coded frames in **D**, showing proportion of transcripts with included (blue) and excluded (red) e1_5' region. (**F**) RT-PCRs of human skeletal muscle samples (Clontech) with primers specific to either e5′UTR2 or e5′UTR3 and 3′UTR of *MBNL1* showing that these transcripts are full-length. (**G**) Schematic representation of splicing outcomes of *MBNL1* e1 depending on the transcription initiation start site. Bar graphs show mean values + SD. Statistical significance was evaluated by two-tailed Student's *t*-test; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

muscle tissues, and T1, shown to be substantially active only in liver, spleen and thymus (Figure 2C). Surprisingly, in transcripts starting from T1 we detected no *MBNL1* with included e1, while only e1 containing mRNA was observed when transcription was initiated from T3 (Supplementary Figure S4E). Finally, we tested whether transcripts originating from T2 and T3 promoters are full-length mRNAs. In the case of mRNA with e1 (driven from T2 and T3) and without e1 (T2 only), we detected full-length products, indicative of their potential to give rise to functional MBNL1 protein isoforms (Figure 3F and Supplementary Figure S4F). Taken together, these results indicate that only transcripts derived from T2 undergo autoregulatory e1 alternative splicing changes (Figure 3G).

5′UTR with e1 determines higher translational activity of MBNL1

To understand how e1 splicing might affect cellular localization of endogenous *MBNL1* transcripts, we fractionated HeLa and NIH/3T3 cells and estimated their distribution (Figure 4A). Although both cell line models differ in the ratio of *MBNL1* mRNA with and without e1, they displayed a very similar trend following fractionation, i.e. concentration of isoforms without e1 in the cytosolic fraction devoid of polysomes and enrichment of e1-containing mRNA in the membrane fraction containing polysomes with actively translated mRNAs (56). We hypothesised that the specific localization of the isoforms could determine altered translational activity of MBNL1. To address this question we generated *Mbnl1*_5′UTR2_*Luc2* vectors comprised of endogenous 5′UTR fragments, with and without e1, fused to luciferase *Luc2* (Figure 4B; +e1_Luc and –e1_Luc). Vector without 5′UTR was used as a control (Figure 4C; Luc). Interestingly, a marked reduction in luciferase production was observed upon addition of either of the 5′UTR constructs (Figure 4C). Most importantly, direct comparison of *Mbnl1* 5′UTRs revealed a 50% drop in translational activity of 5′UTR lacking e1 when compared to the isoform containing e1 (Figure 4C), suggesting that e1 skipping in *MBNL1* pre-mRNA has a negative effect on translation of MBNL1 protein, possibly due to inefficient loading of the mRNA lacking e1 on polysomes.

Next, we tested whether MBNL1 binding to its mature transcript could have an additional effect on translation. To confirm MBNL1 interaction with e1-containing mRNA, we co-transfected HeLa cells with *EGFP_MBNL1* and *Mbnl1*_e1_5′UTR2_*Luc2* vectors (Figure 4D) harbouring either an *Mbnl1* e1 fragment (Ctrl) or *Mbnl1* e1 deletion constructs lacking the sequence spanning the MBNL-binding site (Δ1 and Δ2), and co-immunoprecipitated EGFP_MBNL1 protein with target RNAs (Figure 4E). Deletion of the MBNL1 binding site in e1 (Figure 4D, marked green) markedly prevented EGFP_MBNL1 binding to e1 and thus impaired co-immunoprecipitation of Δ1 and Δ2 mRNAs, confirming MBNL1 binding to the mature transcript containing e1. Subsequently, our data suggested slightly altered subcellular localization of *Mbnl1*_e1_5′UTR_*Luc2* Ctrl and Δ1 mRNAs (Figure 4F), reminiscent of the pattern observed following fractionation of HeLa and NIH/3T3 cells and detection of MBNL1 endogenous isoforms with and without e1 (Figure 4A). Nonetheless, we did not notice any direct effect of MBNL1 binding to e1-containing transcripts on translation (Figure 4C and G). These results imply that MBNL1 translation is influenced by e1 exclusion from pre-mRNA and determination of the 5′UTR sequence by MBNL1, however, factors other than MBNL1 itself seem to play the major role in directing mRNA with e1 to polysomes.

Increased content of *MBNL1* isoforms containing e1 in muscles lacking functional MBNLs

We then investigated whether the total level of *MBNL1* mRNA differs depending on the presence or absence of functional MBNLs using semi-quantitative multiplex RT-PCR and quantitative RT-PCR (Figure 5 and Supplementary Figure S5). Our data indicated elevated total amounts of *MBNL1* mRNA in skeletal muscles of both, DM patients (Figure 5A, Supplementary Figure S5A) and HSA^{LR} mice (Figure 5B), in which +e1 mRNA isoforms predominate (Figure 3A). A reverse trend, although not statistically significant, was observed upon delivery of *EGFP_MBNL1* constructs to HeLa cells (Supplementary Figure S5B). In addition, amplification of transcription site-specific *MBNL1* mRNAs revealed the highest content of T2 and T3 transcripts in muscles from DM1 patients and HSA^{LR} mice (Figure 5), from which e1-containing isoforms originate (Supplementary Figure S4F). Elevated total amount of *MBNL1* observed upon functional loss of MBNLs could be explained by decreased *MBNL1* e1 circRNA production (49). While we detected relatively small amounts of *MBNL1* e1 circRNA in human and mouse muscles, neither DM nor HSA^{LR} samples showed any significant change in its quantity (Supplementary Figures S6A-C), despite its substantial increase during differentiation in C2C12 myoblasts (Supplementary Figure S6D). Taken together, we conclude that the overall increase of *MBNL1* mRNA containing e1 in the absence of functional MBNLs is a consequence of increased transcription and/or increased stability of mRNA.

MBNL1 proteins lacking e1 coding sequence show decreased stability

Exclusion of e1 shifts the start codon from e1 to the second half of e2, which is predicted to result in expression of isoforms lacking two of the four ZnFs. To investigate functional properties of this truncation, we first forced expressed MBNLs missing e1 on –e5 or +e5, and EGFP background (Figure 6A and B, EGFP_MBNL1_41.2/43.2) and compared their cellular localization and stability to their full-length counterparts (EGFP_MBNL1_41.4/43.4). In transiently transfected HeLa cells we detected no change in the distribution pattern of truncated proteins when compared to isoforms containing all four ZnFs. EGFP_MBNL1_41.4 and EGFP_MBNL1_41.2 localized to both, the cytoplasm and nucleus, while proteins containing e5 were confined to the nucleus (Figure 6C). However, the stability of MBNL1 proteins containing only two ZnFs was significantly decreased (Supplementary Figure S7). As EGFP could influence stability of MBNL1 proteins, we additionally generated MBNL1 constructs lacking this tag (*EGFPΔ_MBNL1*; Figure 6D) and transiently expressed them in COS7 cells. Importantly, while MBNL1s containing all four ZnFs could be readily detected, no isoforms lacking the amino-acid sequence encoded by e1 were observed (Figure 6E), despite relatively high content of their mRNAs (Figure 6F). Based on these results, we concluded that in addition to reduced translational activity (Figure 4C), MBNL1 isoforms containing only two ZnFs have also decreased stability.

Our model predicts that upon expression of MBNLs there is a shift towards isoforms without e1, which should result in downregulation of endogenous MBNLs. To test this hypothesis, we forced expressed EGFPΔ_MBNL1s in HeLa and COS7 cells and immunoblotted for MBNL1 (Figure 6G and H). Importantly, we observed significantly

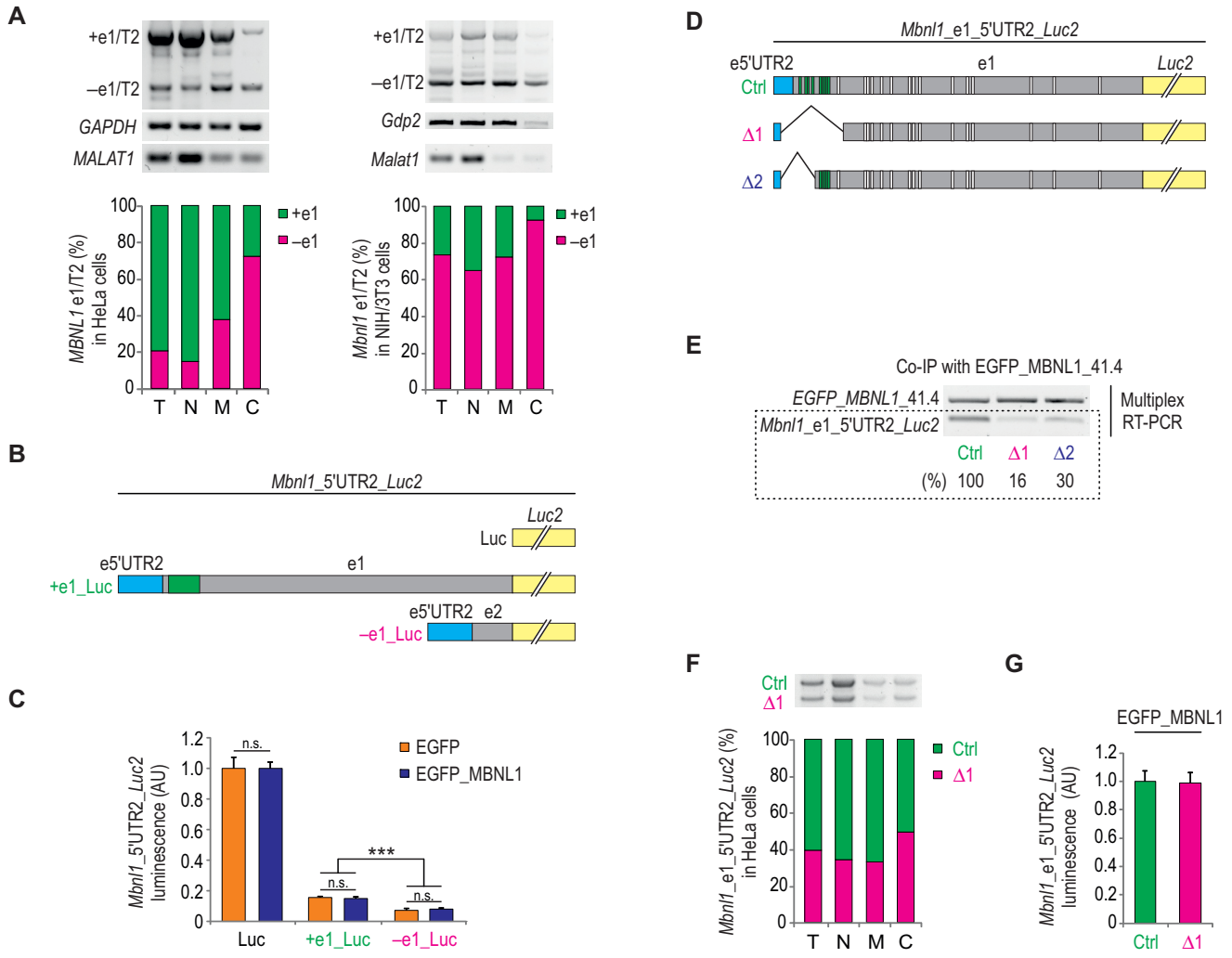

Figure 4. Reduced translation from *MBNL1* 5′UTR lacking e1. (**A**) Subcellular fractionation of HeLa (left panel) or NIH/3T3 cells (right panel) followed by RT-PCR analyses for *MBNL1* e1, *GAPDH*/*Gdp2* and *MALAT1*. *MALAT1* was used as a positive control for the nuclear fraction. (**B**) A schematic representation of *Mbnl1*_ 5′UTR2_*Luc2* (Luc, +e1_Luc, –e1_Luc) luciferase constructs used in **C**. The vectors contain *Mbnl1* e1 5′UTRs based either on e1 (+e1_Luc) or e2 (–e1_Luc) fused to luciferase *Luc2* or *Luc2* alone (Luc). (**C**) Luciferase assay showing translational activities of Luc, +e1_Luc and –e1_Luc mRNAs following plasmid co-delivery with *EGFP* or *EGFP_MBNL1* (41.4) constructs. AU indicates arbitrary units. $n = 3$ per each treatment group. Bar graphs show mean values + SD. (**D**) A schematic representation of *Mbnl1*_e1_5′UTR2_*Luc2* (Ctrl, Δ1, Δ2) used in **E-G.** Ctrl contains a 3′ part of e5′UTR2 and most of the non-coding region of e1 of *Mbnl1*. Δ1 and Δ2 lack a variable number of YGCY motifs at the 5′-most part of e1. (**E**) Co-immunoprecipitation (Co-IP) of EGFP_MBNL1_41.4 and RNA complexes following UV-cross linking of proteins with target RNAs. (**F**) Subcellular fractionation following co-transfection of Ctrl and Δ1 plasmids into HeLa cells. (**G**) Luminescence obtained following administration of *EGFP_MBNL1* with either Ctrl or Δ1 plasmids. T, total cell extract; N, nuclear fraction; M, membrane fraction; C, cytosolic fraction devoid of polysomes. Bar graphs show mean values + SD. Statistical significance was evaluated by two-tailed Student's *t*-test; ***$P < 0.001$.

diminished content of endogenous MBNLs upon delivery of MBNLs with four ZnFs.

MBNL1 proteins containing only two ZnFs have no dominant-negative effect over full-length MBNL1 isoforms

Subsequently, we compared splicing activities of EGFPΔ_MBNL1 proteins containing two and four ZnFs, based on repression of MBNL1-regulated alternative exons in *MBNL1, MBNL2, NCOR2* and *NFIX* pre-mRNAs (Figure 7A and Supplementary Figure S8). Our direct comparison studies showed the highest splicing activity of the nuclear isoform containing four ZnFs and e5 (43.4; Figure 6B), despite its lower cellular content than 41.4 (Figure 6E). Interestingly we also detected residual splicing activity of the 43 isoform containing only 2 ZnFs, indicating that low levels of truncated isoforms are expressed (Figure 7A, compare splicing activities of 41.2 and 43.2).

The detected residual activity of MBNL1 proteins containing only two ZnFs as well as the observation that MBNL1 proteins can dimerize (11,13) and potentially form heterodimers led us to ask whether MBNL1 proteins containing only two ZnFs could have a dominant-negative effect over full-length MBNL1 isoforms. To answer this question, we first forced expressed EGFP_MBNL1 in HeLa cells and tested splicing activities of EGFP-stabilized isoforms lacking two ZnFs (Figure 7B-D and Supplementary Fig-

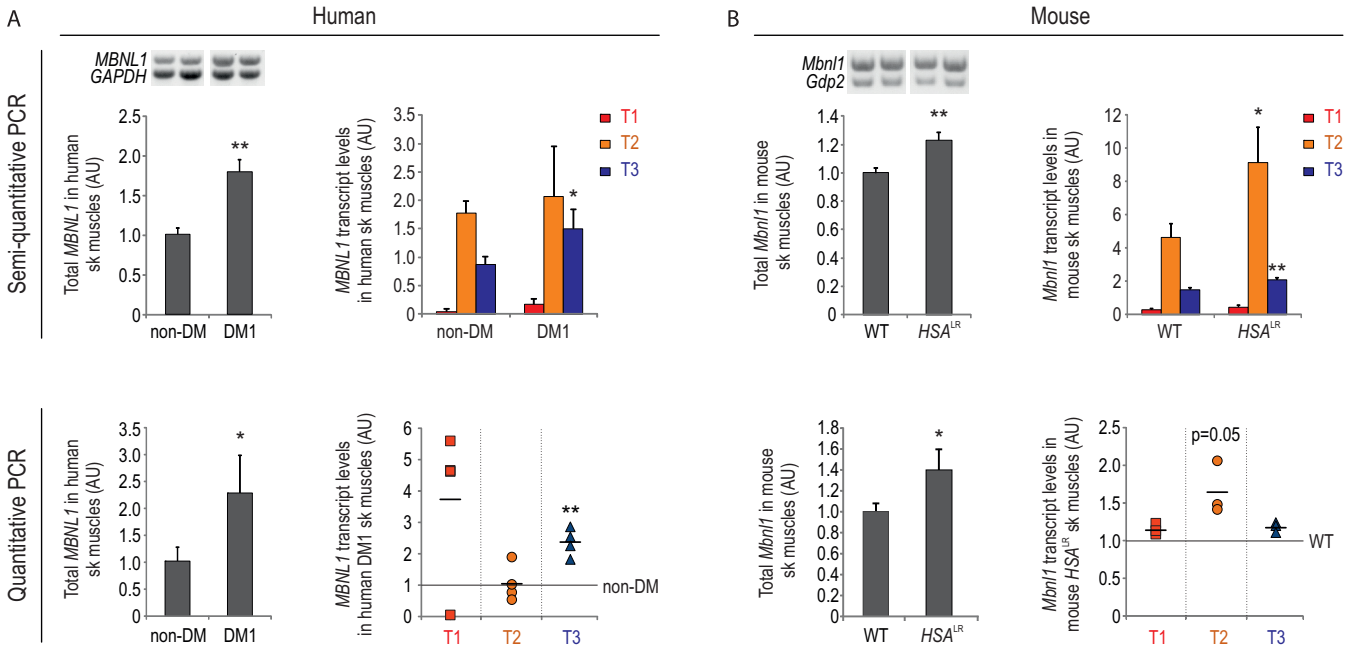

Figure 5. Increased content of *MBNL1* transcripts containing e1 in muscles showing loss of functional MBNLs. (**A** and **B**) Semi-quantitative multiplex RT-PCRs (upper panel) and quantitative RT-PCRs (lower panel) of total and transcription start site-specific (T1–T3) *MBNL1* mRNAs in either human (**A**) or mouse (**B**) skeletal muscles. $n = 3$, non-DM and DM1 (upper left panel), WT, HSA^{LR}; $n = 4$, non-DM and DM1 (upper right and lower panels). The signals were related to *GAPDH*/*Gdp2* mRNA and genomic DNA signals as described for data shown in Figure 2B (**A**) or to *GAPDH*/*Gdp2* mRNA alone (**B**). AU indicates arbitrary units. Bar graphs show mean values + SD. Statistical significance was evaluated by two-tailed Student's *t*-test; $*P < 0.05$, $**P < 0.01$.

ures S9 and S10). As judged by splicing repression of alternative e5 in *cTNT* minigene, EGFP_MBNL1_41.2 and 43.2 showed markedly reduced splicing activities (Figure 7B, and Supplementary Figures S9 and S10A). In agreement with this data, truncated proteins did not efficiently splice alternative exons in endogenous pre-mRNA targets (Figure 7B and Supplementary Figure S10B). We then co-expressed *EGFP_MBNL1_*41.4 with different amounts of *EGFP_MBNL1_*41.2 in HeLa cells together with the *cTNT* minigene (Figure 7D). Our results clearly indicate no adverse effect of e1-truncated MBNL1 on the full-length protein-induced alternative splicing.

DISCUSSION

In this study, we demonstrate that MBNLs regulate steady-state levels of MBNL1 (Figure 8) via an interaction with e1 of its own transcript. High MBNL cellular content induces an increase in *MBNL1* mRNA lacking e1 that is translated with lower efficiency and additionally, the truncated isoform is unstable and less active, presumably due to loss of the first two ZnFs shown to be essential for RNA target recognition [see also (57,58)]. In the adult cardiac and skeletal muscles characterized by high *MBNL1* mRNA content (19), we observed elevated amounts of *MBNL1* mRNA without e1. A reverse process takes place when the level of functional MBNLs is low, as in DM (8,22,29), where a highly functional, full-length MBNL1 is generated. This is in agreement with a recent study showing double amount of MBNL1 protein in DM1 patients relative to non-DM controls (59). Additionally, we observed no change in the splicing of *MBNL1* e1 in the adult brain, which corresponds with the relatively low level of MBNLs in this particular tissue (19).

One of the most surprising results of this study is that e1 exclusion negatively correlates with the decreasing distance between the transcription start site and e1. Particularly, in skeletal muscles of non-DM patients, in which we observed relatively high amounts of transcripts without e1 initiated from T2, there is only e1-containing *MBNL1* mRNA generated from T3. Importantly, T2 is located ∼30 kb further from e1 than T3, which is positioned in close proximity to e1. This indicates that the length of flanking introns could have a significant impact on alternative splicing [see (60,61)]. Importantly, the knowledge concerning MBNL1 expression level fluctuations depending on the transcription initiation start might offer a possible therapeutic strategy against DM aiming at the autoregulatory overexpression of MBNL1 from T2. So far, not much is known about transcription factors that drive expression of *MBNL1*, although MEF2s have been indicated as possible players (62). Unexpectedly, we observed no decrease in the content of e1 circRNA in either DM or HSA^{LR} skeletal muscles, opposite to the phenotype that could have been predicted based on the previous study (49) and preferential inclusion of e1 to mature *MBNL1* mRNA. One of the reasons could be a counteractive effect of enhanced transcription of *MBNL1* upon loss of functional MBNLs. In agreement with this hypothesis, increasing *MBNL1* transcription coincided with e1 circRNA generation during myoblast differentiation.

We demonstrate that all three MBNLs induce e1 skipping, which is in agreement with the current state of knowl-

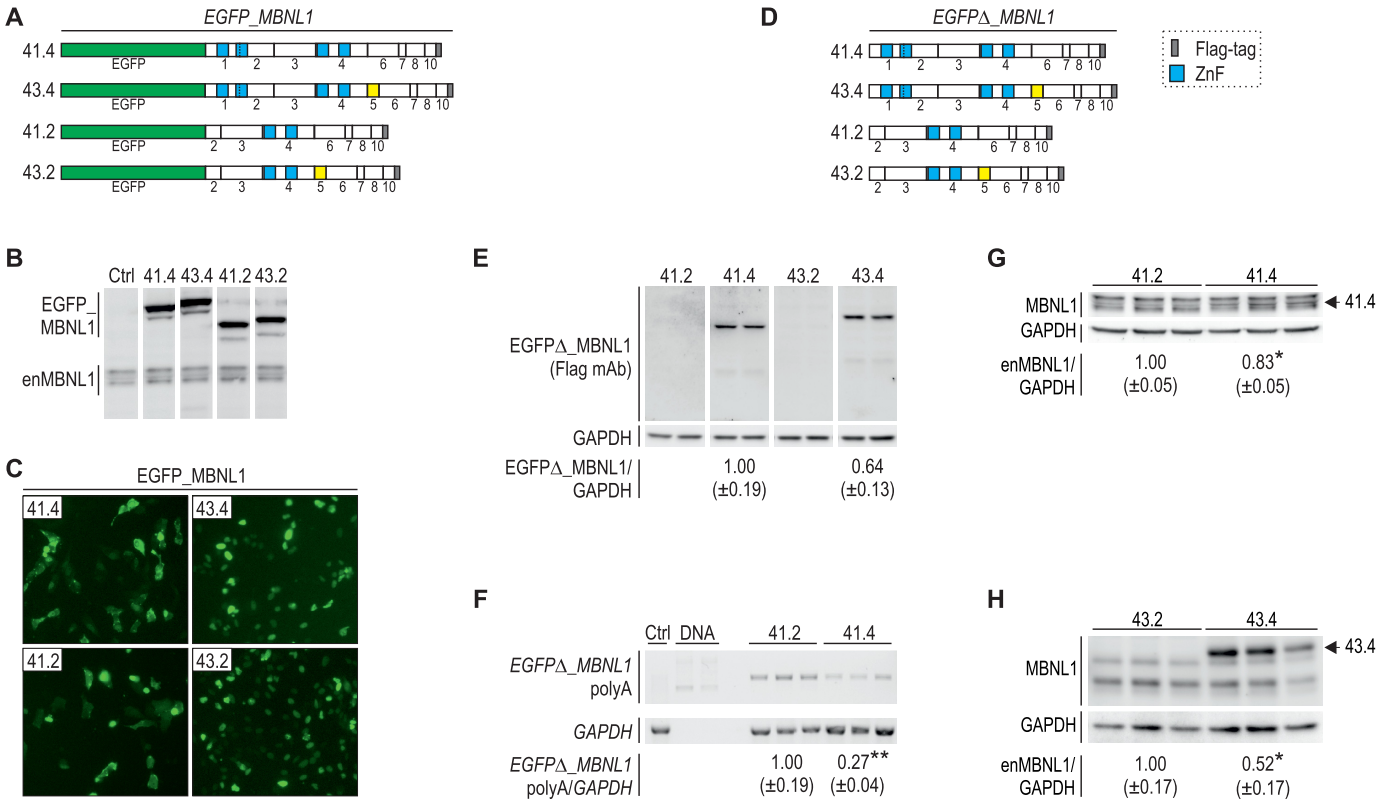

Figure 6. Autoregulatory function of MBNL1 expression. (**A**) Schematic representation of *MBNL1* constructs containing *EGFP* (*EGFP_MBNL1*). The constructs differ in the presence of e5 and the number of ZnFs. The *MBNL1* plasmid sequence corresponds to the coding sequence of *MBNL1* mRNA with (41.4/43.4) and without e1 (41.2/43.2), respectively. Green, blue, yellow and grey boxes indicate *EGFP*, ZnFs, e5 and Flag-tag, respectively. (**B**) Immunoblotting analysis for EGFP_MBNL1 and endogenous MBNL1 (enMBNL1) in HeLa cell lysates following transfection with *EGFP* (Ctrl) or *EGFP_MBNL1*-containing plasmids as indicated. (**C**) Immunofluorescence microscopy of HeLa cells transfected with *EGFP_MBNL1* constructs as indicated. (**D**) Constructs analogous to the ones in **A** but devoid of *EGFP* (*EGFPΔ_MBNL1*). (**E**) Immunoblotting assay for MBNL1 using antibody to Flag-tag in lysates from COS7 cells co-transfected with *EGFPΔ_MBNL1* constructs as indicated. Note lack of detectable signal of MBNL1 following delivery of 41.2 and 43.2. The lower panel shows quantification of EGFPΔ_MBNL1 relatively to GAPDH (mean values ± SD). $n = 3$ per each treatment group. (**F**) Semi-quantitative RT-PCR analyses of *EGFPΔ_MBNL1* and *GAPDH* mRNA following delivery of *EGFPΔ_MBNL1* constructs as indicated. cDNA from untreated HeLa cells (Ctrl) or plasmid DNA (41.2 and 41.4; DNA) were used as controls. (**G** and **H**) Immunoblotting analyses for MBNL1 following delivery of *EGFPΔ_MBNL1*_41.2 and 41.4 in HeLa cells (**G**) or 43.2 and 43.4 in COS7 cells (**H**). Quantification of enMBNL1 (as the sum of two endogenous MBNL1 bands) relative to GAPDH is shown below the blots (mean values ± SD). Statistical significance was evaluated by two-tailed Student's *t*-test; * $P < 0.05$, ** $P < 0.01$.

edge concerning how MBNLs interact with RNA targets (4,14,17,42). The mode of splicing regulation of *MBNL1* e1 is consistent with the negative splicing outcome of other alternative exons to which MBNL1 binds, such as *CLCN1* e7A and *NFIX* e7 (15,36), which is presumably a result of MBNL1-driven occlusion of sites to which other splicing factors can bind (19,63). Wang *et al.* (4) showed that MBNL1 and 2 determine subcellular compartmentalization of mRNA isoforms with distinct 3′UTRs depending on the presence of binding sites for MBNLs. Based on our data, we can conclude that MBNL1-dependent exon regulation can also be a factor influencing cellular localization of mRNAs.

As for the previously described MBNL target, *Tnnt3* (13), we performed a series of *in vitro* experiments to characterize MBNL1 binding to its e1. These analyses indicate that the observed high affinity interaction is facilitated by several sequential YGCY consensus motifs and their positioning within a semistable and relatively loosely organized hairpin. This is in agreement with a recent report showing that pairing of Y bases within the YGCY motif abolishes MBNL1 binding (18). E1 is not the only *MBNL1* exon that is sensitive to MBNL content. E5 and e7 are preferentially skipped in adult tissues when compared to the foetal (64), which leads to MBNL1 relocalization and dimerization, respectively (11,13). Exclusion of e5 is directly induced by MBNL1 binding to intron 4 that presumably inhibits the spliceosome from locating the 3′ splice-site (65). Our study revealed additionally that e3 is preferentially excluded from *MBNL1* transcripts upon high MBNL1 content in HeLa cells. This might be an additional regulative process directed at disabling MBNL1 function, as e3 encodes a flexible linker allowing MBNL1 to interact with a wide variety of RNA targets (7).

Despite high amounts of e1-deficient transcript in cellular models upon delivery of *MBNL1* constructs and following forced expression of e1-truncated MBNL1, we were unable to detect protein isoforms lacking e1. Similarly, deletion of a genomic fragment encompassing a large portion of e1 to generate Mbnl1 knock-out mice resulted in the ab-

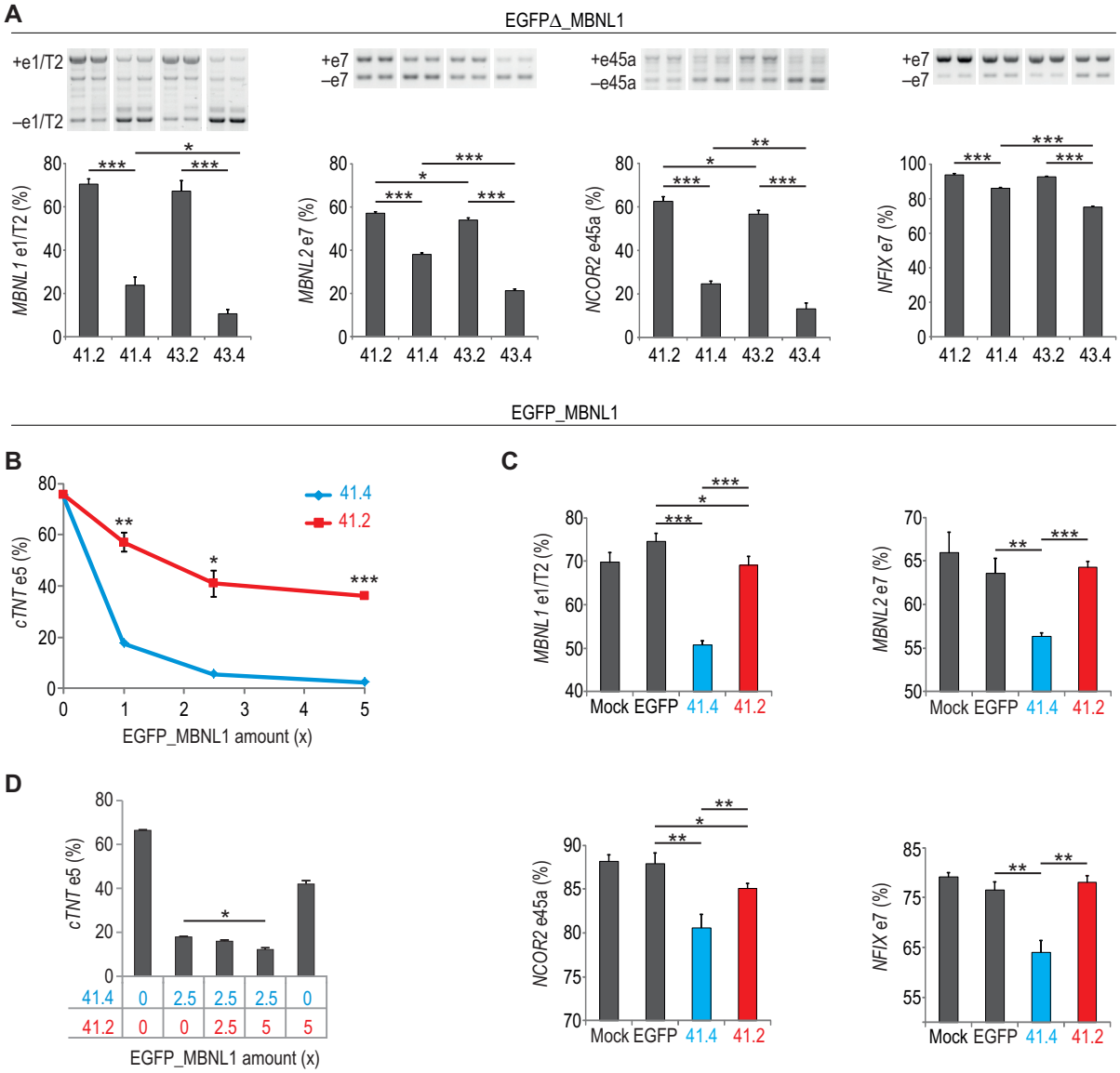

Figure 7. ZnFs 1 and 2 are essential for optimal MBNL1 splicing activity. (**A**) RT-PCR analyses of *MBNL1* e1, *MBNL2* e7, *NCOR2* e45a and *NFIX* e7 distribution in COS7 cells following delivery of *EGFPΔ_MBNL1* plasmids. (**B**) Comparison of splicing activities of MBNL1 proteins with four and two ZnFs in HeLa cells co-transfected with *EGFP_MBNL1* constructs and a *cTNT* minigene. (**C**) RT-PCR analyses of the distribution of the same alternative exons as in **A** but in HeLa cells upon administration of *EGFP* or *EGFP_MBNL1* vectors. Constructs were delivered in appropriate amounts to equalize for the reduced stability of EGFP_MBNL1_41.2. Bar graphs show mean values + SD. (**D**) Splicing activity of MBNL1 on *cTNT* e5 upon co-delivery of *EGFP_MBNL1*_41.4 and increasing amounts of *EGFP_MBNL1*_41.2 in HeLa cells. Mean values ± standard deviation (SD) are shown on the graph. $n = 3$ per each treatment group. Statistical significance was evaluated by a two-tailed Student's *t*-test; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

sence of the full-length as well as e1-deficient protein isoforms, despite preserved mRNA levels as judged by amplification of *Mbnl1* mRNA regions downstream from the deletion (66). One reason might be that the isoform containing only two ZnFs exhibits diminished stability, as revealed by direct comparison of the full-length and e1-truncated isoforms containing EGFP. Furthermore, the e1-deficient *MBNL1* transcript specifically concentrated in the cytosolic fraction devoid of polysomes, which would explain relatively low translation efficiency of *Luc2* fused to the corresponding *MBNL1* 5′UTR lacking e1. We also show that the truncated MBNL isoform lacking two of the four ZnFs, stabilized by EGFP, is highly inoperative regardless whether tested on a minigene or endogenous pre-mRNA targets. This data is in accordance with previously published studies demonstrating the crucial role of the ZnF1/2 tandem in effective splicing of majority of pre-mRNA targets (57,58,67).

In conclusion, our study uncovered a novel, autoregulatory function of MBNL proteins based on their binding to e1 of *MBNL1* transcript. This function might facilitate cellular protection from MBNL protein level fluctuations, which might otherwise lead to adverse effects caused by extreme MBNL content. This regulative process, particularly in the context of its promoter-dependent regulation, should be taken into the account in development of therapeutic strategies against DM grounded on MBNL overexpression.

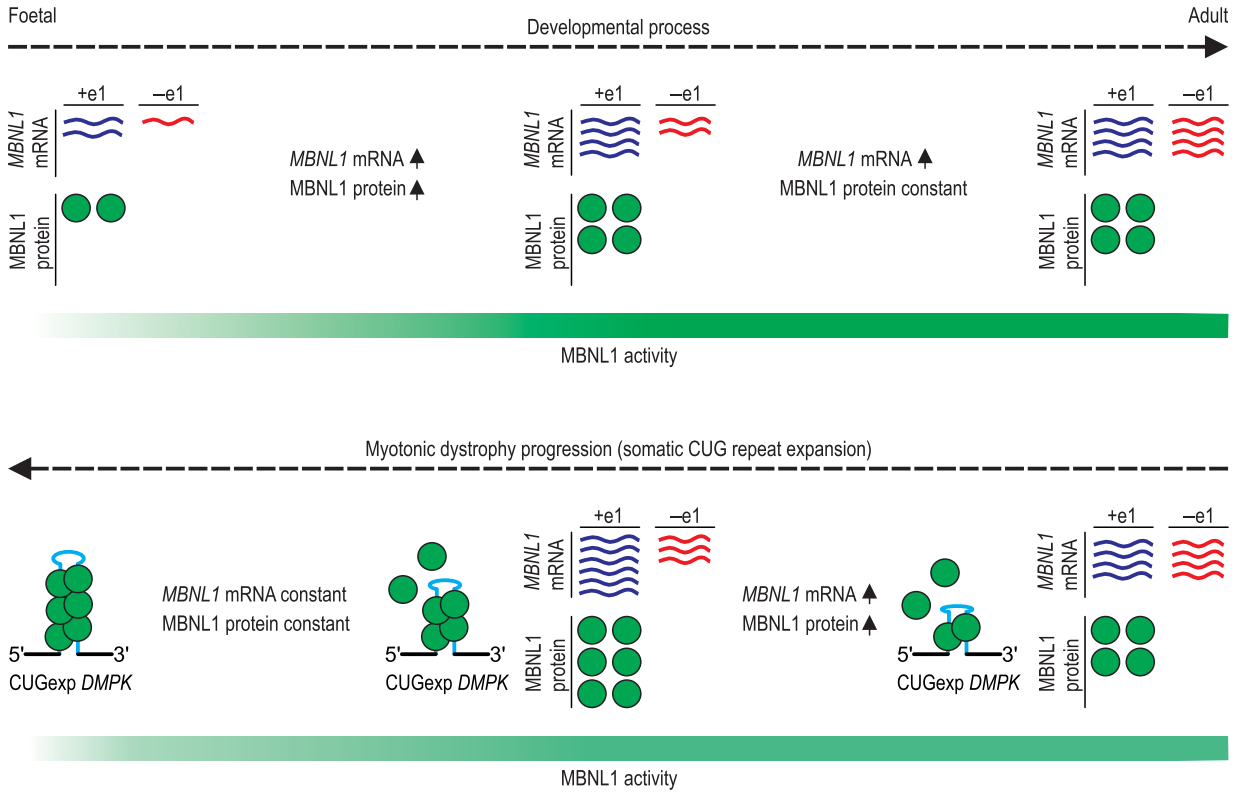

Figure 8. Proposed model of MBNL1 autoregulation dependent on e1 exclusion from *MBNL1* transcript. During development, elevated content of MBNL1 protein induces skipping of e1 in *MBNL1* pre-mRNA, which halts further increase of MBNL1. In myotonic dystrophy, sequestration of MBNLs on *DMPK* mRNA containing expanded CUG repeats (CUGexp) results in *MBNL1* e1-dependent compensatory process: MBNL1 content rises because first, the 5′UTR containing e1 determines higher translational activity of MBNL1 and second, MBNL1 with four ZnFs is more stable than its truncated counterpart originating from e1-devoid mRNA. With increasing expansion size, the e1-driven compensatory process overcomes MBNL1 sequestration less efficiently. Based on this assumption, condition of patients worsens over time due to somatic expansion of CUG repeats and a resulting drop in cellular activity of MBNL1.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Polish National Science Centre [2011/01/B/NZ1/01603, 2014/15/NZ2/02453 to K.S.]; Foundation for Polish Science-TEAM program co-financed by the European Union within the European Regional Development Fund (to K.S.); Polish National Science Centre [2014/15/B/NZ5/00142 to E.S.-K.]; Ministry of Science and Higher Education of the Republic of Poland, from the quality—promoting subsidy, under the Leading National Research Centre (KNOW) programme for the years 2012–2017 (KNOW RNA Research Centre in Poznan) [01/KNOW2/2014]. Funding for open access charge: Ministry of Science and Higher Education of the Republic of Poland, from the quality—promoting subsidy, under the Leading National Research Centre (KNOW) programme (KNOW RNA Research Centre in Poznan) [01/KNOW2/2014].
Conflict of interest statement. None declared.

REFERENCES

1. Batra,R., Charizanis,K., Manchanda,M., Mohan,A., Li,M., Finn,D., Goodwin,M., Zhang,C., Sobczak,K., Thornton,C. *et al.* (2014) Loss of MBNL Leads to Disruption of Developmentally Regulated Alternative Polyadenylation in RNA-Mediated Disease. *Mol. Cell*, **56**, 311–322.
2. Masuda,A., Andersen,H., Doktor,T., Okamoto,T., Ito,M., Andresen,B. and Ohno,K. (2012) CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. *Sci. Rep.*, **2**, 209.
3. Rau,F., Freyermuth,F., Fugier,C., Villemin,J.-P., Fischer,M.-C., Jost,B., Dembele,D., Gourdon,G., Nicole,A., Duboc,D. *et al.* (2011) Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.*, **18**, 840–845.
4. Wang,E., Cody,N., Jog,S., Biancolella,M., Wang,T., Treacy,D., Luo,S., Schroth,G., Housman,D., Reddy,S. *et al.* (2012) Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell*, **150**, 710–724.
5. Wang,E., Ward,A., Cherone,J., Giudice,J., Wang,T., Treacy,D., Lambert,N., Freese,P., Saxena,T., Cooper,T. *et al.* (2015) Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.*, **25**, 858–871.
6. Osborne,R., Lin,X., Welle,S., Sobczak,K., O'Rourke,J., Swanson,M. and Thornton,C. (2009) Transcriptional and post-transcriptional impact of toxic RNA in myotonic dystrophy. *Hum. Mol. Genet.*, **18**, 1471–1481.
7. Cass,D., Hotchko,R., Barber,P., Jones,K., Gates,D. and Berglund,J. (2011) The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements. *BMC Mol. Biol.*, **12**, 20.

8. Fardaei,M., Rogers,M., Thorpe,H., Larkin,K., Hamshere,M., Harper,P. and Brook,J. (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.
9. Teplova,M. and Patel,D. (2008) Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat. Struct. Mol. Biol.*, **15**, 1343–1351.
10. Fernandez-Costa,J. and Artero,R. (2010) A conserved motif controls nuclear localization of Drosophila Muscleblind. *Mol. Cells*, **30**, 65–70.
11. Tran,H., Gourrier,N., Lemercier-Neuillet,C., Dhaenens,C.-M., Vautrin,A., Fernandez-Gomez,F., Arandel,L., Carpentier,C., Obriot,H., Eddarkaoui,S. *et al.* (2011) Analysis of exonic regions involved in nuclear localization, splicing activity, and dimerization of Muscleblind-like-1 isoforms. *J. Biol. Chem.*, **286**, 16435–16446.
12. Kino,Y., Washizu,C., Kurosawa,M., Oma,Y., Hattori,N., Ishiura,S. and Nukina,N. (2015) Nuclear localization of MBNL1: splicing-mediated autoregulation and repression of repeat-derived aberrant proteins. *Hum. Mol. Genet.*, **24**, 740–756.
13. Yuan,Y., Compton,S., Sobczak,K., Stenberg,M., Thornton,C., Griffith,J. and Swanson,M. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.*, **35**, 5474–5486.
14. Charizanis,K., Lee,K.-Y., Batra,R., Goodwin,M., Zhang,C., Yuan,Y., Shiue,L., Cline,M., Scotti,M., Xia,G. *et al.* (2012) Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron*, **75**, 437–450.
15. Du,H., Cline,M., Osborne,R., Tuttle,D., Clark,T., Donohue,J., Hall,M., Shiue,L., Swanson,M., Thornton,C. *et al.* (2010) Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.*, **17**, 187–193.
16. Goers,E., Purcell,J., Voelker,R., Gates,D. and Berglund,J. (2010) MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.*, **38**, 2467–2484.
17. Poulos,M., Batra,R., Li,M., Yuan,Y., Zhang,C., Darnell,R. and Swanson,M. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. *Hum. Mol. Genet.*, **22**, 3547–3558.
18. Lambert,N., Robertson,A., Jangi,M., McGeary,S., Sharp,P. and Burge,C. (2014) RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell*, **54**, 887–900.
19. Konieczny,P., Stepniak-Konieczna,E. and Sobczak,K. (2014) MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res.*, **42**, 10873–10887.
20. Fernandez-Costa,J., Llamusi,M., Garcia-Lopez,A. and Artero,R. (2011) Alternative splicing regulation by Muscleblind proteins: from development to disease. *Biol. Rev. Camb. Philos. Soc.*, **86**, 947–958.
21. Han,H., Irimia,M., Ross,P., Sung,H.-K., Alipanahi,B., David,L., Golipour,A., Gabut,M., Michael,I., Nachman,E. *et al.* (2013) MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature*, **498**, 241–245.
22. Miller,J., Urbinati,C., Teng-Umnuay,P., Stenberg,M., Byrne,B., Thornton,C. and Swanson,M. (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.*, **19**, 4439–4448.
23. Lee,K.-Y., Li,M., Manchanda,M., Batra,R., Charizanis,K., Mohan,A., Warren,S., Chamberlain,C., Finn,D., Hong,H. *et al.* (2013) Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol. Med.*, **5**, 1887–1900.
24. Kanadia,R., Urbinati,C., Crusselle,V., Luo,D., Lee,Y.-J., Harrison,J., Oh,S. and Swanson,M. (2003) Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. *Gene Express. Patterns: GEP*, **3**, 459–462.
25. Lee,K.-S., Smith,K., Amieux,P. and Wang,E. (2008) MBNL3/CHCR prevents myogenic differentiation by inhibiting MyoD-dependent gene transcription. *Differ. Res. Biol. Divers.*, **76**, 299–309.
26. Lee,K.-S., Squillace,R. and Wang,E. (2007) Expression pattern of muscleblind-like proteins differs in differentiating myoblasts. *Biochem. Biophys. Res. Commun.*, **361**, 151–155.
27. Squillace,R., Chenault,D. and Wang,E. (2002) Inhibition of muscle differentiation by the novel muscleblind-related protein CHCR. *Dev. Biol.*, **250**, 218–230.
28. Udd,B. and Krahe,R. (2012) The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *Lancet Neurol.*, **11**, 891–905.
29. Fardaei,M., Larkin,K., Brook,J. and Hamshere,M. (2001) In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res.*, **29**, 2766–2771.
30. Nakamori,M., Sobczak,K., Puwanant,A., Welle,S., Eichinger,K., Pandya,S., Dekdebrun,J., Heatwole,C., McDermott,M., Chen,T. *et al.* (2013) Splicing biomarkers of disease severity in myotonic dystrophy. *Ann. Neurol.*, **74**, 862–872.
31. Kalsotra,A., Xiao,X., Ward,A., Castle,J., Johnson,J., Burge,C. and Cooper,T. (2008) A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 20333–20338.
32. Fugier,C., Klein,A., Hammer,C., Vassilopoulos,S., Ivarsson,Y., Toussaint,A., Tosch,V., Vignaud,A., Ferry,A., Messaddeq,N. *et al.* (2011) Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat. Med.*, **17**, 720–725.
33. Gao,Z. and Cooper,T. (2013) Reexpression of pyruvate kinase M2 in type 1 myofibers correlates with altered glucose metabolism in myotonic dystrophy. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 13570–13575.
34. Ho,T., Charlet-Berguerand,N., Poulos,M., Singh,G., Swanson,M. and Cooper,T. (2004) Muscleblind proteins regulate alternative splicing. *EMBO J.*, **23**, 3103–3112.
35. Kanadia,R., Johnstone,K., Mankodi,A., Lungu,C., Thornton,C., Esson,D., Timmers,A., Hauswirth,W. and Swanson,M. (2003) A muscleblind knockout model for myotonic dystrophy. *Science (New York, N.Y.)*, **302**, 1978–1980.
36. Kino,Y., Washizu,C., Oma,Y., Onishi,H., Nezu,Y., Sasagawa,N., Nukina,N. and Ishiura,S. (2009) MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res.*, **37**, 6477–6490.
37. Rau,F., Laine,J., Ramanoudjame,L., Ferry,A., Arandel,L., Delalande,O., Jollet,A., Dingli,F., Lee,K.-Y., Peccate,C. *et al.* (2015) Abnormal splicing switch of DMD's penultimate exon compromises muscle fibre maintenance in myotonic dystrophy. *Nat. Commun.*, **6**, 7205.
38. Savkur,R., Philips,A. and Cooper,T. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat. Genet.*, **29**, 40–47.
39. Tang,Z., Yarotskyy,V., Wei,L., Sobczak,K., Nakamori,M., Eichinger,K., Moxley,R., Dirksen,R. and Thornton,C. (2012) Muscle weakness in myotonic dystrophy associated with misregulated splicing and altered gating of Ca(V)1.1 calcium channel. *Hum. Mol. Genet.*, **21**, 1312–1324.
40. Jog,S., Paul,S., Dansithong,W., Tring,S., Comai,L. and Reddy,S. (2012) RNA splicing is responsive to MBNL1 dose. *PLoS One*, **7**.
41. Ule,J., Jensen,K., Mele,A. and Darnell,R. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods (San Diego, Calif.)*, **37**, 376–386.
42. Sznajder,Ł., Michalak,M., Taylor,K., Cywoniuk,P., Kabza,M., Wojtkowiak-Szlachcic,A., Matłoka,M., Konieczny,P. and Sobczak,K. (2016) Mechanistic determinants of MBNL activity, *Nucleic Acids Res.*, doi:10.1093/nar/gkw915.
43. Wojciechowska,M., Taylor,K., Sobczak,K., Napierala,M. and Krzyzosiak,W. (2014) Small molecule kinase inhibitors alleviate different molecular features of myotonic dystrophy type 1. *RNA Bio.l*, **11**, 742–754.
44. Sobczak,K. and Krzyzosiak,W. (2004) Imperfect CAG repeats form diverse structures in SCA1 transcripts. *J. Biol. Chem.*, **279**, 41563–41572.
45. Lin,X., Miller,J., Mankodi,A., Kanadia,R., Yuan,Y., Moxley,R., Swanson,M. and Thornton,C. (2006) Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.*, **15**, 2087–2097.
46. Philips,A., Timchenko,L. and Cooper,T. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science*, **280**, 737–741.
47. Childs-Disney,J., Stepniak-Konieczna,E., Tran,T., Yildirim,I., Park,H., Chen,C., Hoskins,J., Southall,N., Marugan,J., Patnaik,S. *et al.* (2013) Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules. *Nat. Commun.*, **4**, 2044.

48. 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.
49. Ashwal-Fluss,R., Meyer,M., Pamudurti,N., Ivanov,A., Bartok,O., Hanan,M., Evantal,N., Memczak,S., Rajewsky,N. and Kadener,S. (2014) circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell*, **56**, 55–66.
50. Kishore,S., Jaskiewicz,L., Burger,L., Hausser,J., Khorshid,M. and Zavolan,M. (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods*, **8**, 559–564.
51. Rosenbloom,K., Sloan,C., Malladi,V., Dreszer,T., Learned,K., Kirkup,V., Wong,M., Maddren,M., Fang,R., Heitner,S. *et al.* (2013) ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.*, **41**, D56–D63.
52. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
53. Forrest,A., Kawaji,H., Rehli,M., Baillie,J., de Hoon,M., Haberle,V., Lassmann,T., Kulakovskiy,I., Lizio,M., Itoh,M. *et al.* (2014) A promoter-level mammalian expression atlas. *Nature*, **507**, 462–470.
54. Rosenbloom,K., Armstrong,J., Barber,G., Casper,J., Clawson,H., Diekhans,M., Dreszer,T., Fujita,P., Guruvadoo,L., Haeussler,M. *et al.* (2015) The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res.*, **43**, D670–D681.
55. Mankodi,A., Logigian,E., Callahan,L., McClain,C., White,R., Henderson,D., Krym,M. and Thornton,C. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science (New York, N.Y.)*, **289**, 1769–1773.
56. Jagannathan,S., Nwosu,C. and Nicchitta,C.V. (2011) Analyzing mRNA localization to the endoplasmic reticulum via cell fractionation. *Methods Mol. Biol.*, **714**, 301–321.
57. Edge,C., Gooding,C. and Smith,C. (2013) Dissecting domains necessary for activation and repression of splicing by muscleblind-like protein 1. *BMC Mol. Biol.*, **14**, 29.
58. Purcell,J., Oddo,J., Wang,E. and Berglund,J. (2012) Combinatorial mutagenesis of MBNL1 zinc fingers elucidates distinct classes of regulatory events. *Mol. Cell. Biol.*, **32**, 4155–4167.
59. Malatesta,M., Giagnacovo,M., Costanzo,M., Cisterna,B., Cardani,R. and Meola,G. (2013) Muscleblind-like1 undergoes ectopic relocation in the nuclei of skeletal muscles in myotonic dystrophy and sarcopenia. *Eur. J. Histochem.*, **57**, e15.
60. Fox-Walsh,K., Dou,Y., Lam,B., Hung,S., Baldi,P. and Hertel,K. (2005) The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 16176–16181.
61. Roy,M., Kim,N., Xing,Y. and Lee,C. (2008) The effect of intron length on exon creation ratios during the evolution of mammalian genomes. *RNA*, **14**, 2261–2273.
62. Bargiela,A., Llamusi,B., Cerro-Herreros,E. and Artero,R. (2014) Two enhancers control transcription of Drosophila muscleblind in the embryonic somatic musculature and in the central nervous system. *PLoS One*, **9**, e93125.
63. Witten,J. and Ule,J. (2011) Understanding splicing regulation through RNA splicing maps. *Trends Geneti.: TIG*, **27**, 89–97.
64. Terenzi,F. and Ladd,A. (2010) Conserved developmental alternative splicing of muscleblind-like (MBNL) transcripts regulates MBNL localization and activity. *RNA Biol.*, **7**, 43–55.
65. Gates,D., Coonrod,L. and Berglund,J. (2011) Autoregulated splicing of muscleblind-like 1 (MBNL1) Pre-mRNA. *J. Biol. Chem.*, **286**, 34224–34233.
66. Kanadia,R., Johnstone,K., Mankodi,A., Lungu,C., Thornton,C., Esson,D., Timmers,A., Hauswirth,W. and Swanson,M. (2003) A muscleblind knockout model for myotonic dystrophy. *Science (New York, N.Y.)*, **302**, 1978–1980.
67. Grammatikakis,I., Goo,Y.-H., Echeverria,G. and Cooper,T. (2011) Identification of MBNL1 and MBNL3 domains required for splicing activation and repression. *Nucleic Acids Res.*, **39**, 2769–2780.