Determination of functional RNA binding sites for MBNL proteins using antisense oligonucleotides

Określenie funkcjonalnych miejsc oddziaływania białek MBNL z RNA z zastosowaniem oligonukleotydów antysensowych

Piotr Cywoniuk

Supervisor: prof. UAM, dr hab. Krzysztof Sobczak

Department of Gene Expression
Institute of Molecular Biology and Biotechnology
Faculty of Biology
Adam Mickiewicz University in Poznań
Poznań, 2018

I would like to thank:

***Prof. Krzysztof Sobczak** – for 7 years of teaching me what science is*

***Prof. dr hab. Jerzy Ciesiołka and dr hab. Marcin Nowotny** – for your time*

Łukasz Sznajder, Kasia Taylor, Michał Sekrecki, Arek Kajdasz,
Magda Derbis, Agata Michalska, Magda Matłoka,
***Paulina and Piotrek Machtel** – my "brothers in arms"*

Family and friends

Lab and institute members

“Now this is not the end.
It is not even the beginning of the end.
But it is, perhaps, the end of the beginning.”
Winston Churchill

to my wife,
because you are, no matter what

Contents

Summary

Alternative splicing (AS) is a co-transcriptional process providing a substantial increase of proteome diversity by binding of certain RNA-binding proteins (RBPs) to primary transcripts (pre-mRNAs) which leads to conditional inclusion/exclusion of particular exons. That in turn, gives rise to formation of mRNAs differing in sequence from one pre-mRNA variant and, consequently, production of proteins of different amino-acid composition and distinct activity. *Muscleblind*-like (MBNL) protein family gathering three paralogs (MBNL1, 2 and 3) is a group of alternative splicing regulators essential during the development of particular tissues, primarily, skeletal muscles and neuronal cells. MBNLs are also involved in the pathomechanism of myotonic dystrophy (DM), genetic, multi-systemic disorder in which occurs the sequestration of MBNLs on the mRNA of a mutated gene leading to massive misregulation of alternative splicing in patients' cells. Attempts to understand the role of MBNLs and DM pathomechanism have resulted in a number of scientific projects applying whole-transcriptomic analyses, e.g. in order to define novel binding sites recognized by MBNLs. Obtained results provided a substantial amount of data which is required to be experimentally validated. Moreover, these results may shed more light on the mechanism of alternative splicing regulated by MBNL proteins which was the main purpose of my PhD project.

I aimed to give a deeper insight into the role of MBNLs, e.g. by development of reliable protocol based on antisense strategy for verification of new potential MBNL-binding sites. For this purpose, I applied antisense oligonucleotides (AONs), short, chemically modified nucleic acids derivatives able to hybridize specifically to pre-mRNA which may lead to inhibition of RBPs binding and affect their properties including AS regulation.

In three original articles encompassed in my dissertation I presented combination of AONs and hybrid-type minigenes as an efficient method for validation of potential MBNL-binding sites extracted from high-throughput RNA sequencing experiments. Using this approach I also confirmed the hypothesis that all MBNL paralogs recognize the same sequence in RNAs which may lead to inclusion or skipping of alternative exon depending on the binding site localization. Hybrid minigene-based constructs were also used in research on MBNL-binding sites spatial arrangement as well as in the development of assay for searching or validation of potential MBNL sequestration inhibitors in DM. Furthermore, I participated in experiments revealing antagonistic role of MBNL1 and SRSF1 proteins in alternative splicing regulation of exons *via*, either, structural rearrangements or competition for binding sites in RNA. In another project I analyzed the distribution of three alternative exons in the transcripts of *MBNL* genes which revealed differential inclusion among particular human tissues at different developmental stage and in DM patients' samples. In the project which focused on the impact of binding sites organization on MBNLs' activity I provided results quantifying expression level of particular MBNL protein isoforms.

Streszczenie

Splicing alternatywny (ang. alternative splicing, AS) jest kotranskrypcyjnym procesem, w wyniku którego istotnie wzrasta różnorodność proteomu. Wzrost ten jest efektem wiązania się białek wiążących RNA (ang. RNA-binding protein, RBP) do pierwotnych transkryptów (pre-mRNA), które prowadzi do warunkowego włączania lub wyłączania alternatywnych eksonów. Efektem tego zjawiska jest synteza z jednego pre-mRNA wielu białek różniących się składem aminokwasowym i aktywnością. Rodzinę białek *Muscleblind*-like (MBNL) tworzą trzy paralogi (MBNL1, 2 i 3) będące regulatorami alternatywnego splicingu kluczowymi podczas rozwoju wielu tkanek, głównie mięśni i komórek nerwowych. Białka MBNL mają również związek z patomechanizmem dystrofii miotonicznej (DM), choroby degeneracyjnej mięśni, w której dochodzi do sekwestracji białek MBNL na mRNA zmutowanego genu, co prowadzi do masowej deregulacji alternatywnego splicingu w komórkach pacjentów. W celu zgłębienia roli białek MBNL oraz etiologii DM przeprowadzono wiele projektów naukowych polegających m. in. na zastosowaniu analiz cało-transkryptomowych w celu określenia naturalnych miejsc oddziaływania białek MBNL ze specyficznymi regionami w cząsteczkach RNA. Badania te dostarczyły dużej ilości danych wskazujących potencjalne miejsca w RNA wiążących MBNL. Wyniki te wymagają jednak dalszej eksperymentalnej weryfikacji oraz mogą być wykorzystane do zrozumienia mechanizmu regulacji alternatywnego splicingu przez białka MBNL.

To właśnie stanowiło główny cel niniejszej rozprawy. Starałem się poszerzyć wiedzę na temat roli, jaką pełnią białka MBNL m. in. poprzez opracowanie efektywnego protokołu do weryfikacji potencjalnych miejsc ich oddziaływania z RNA z zastosowaniem strategii antysens. W tym celu wykorzystałem oligonukleotydy antysensowe (ang. antisense oligonucleotides, AON), będące krótkimi, modyfikowanymi chemicznie analogami kwasów nukleinowych, które poprzez specyficzną hybrydyzację do pre-mRNA mogą blokować wiązanie się białek RBP i wpływać w ten sposób na ich funkcje, w tym na proces alternatywnego splicingu.

W trzech publikacjach składających się na moją rozprawę doktorską przedstawiłem połączenie strategii AON i minigenów hybrydowych, jako wydajnej metody do weryfikacji potencjalnych miejsc wiązania dla MBNL wyłonionych na drodze eksperymentów z wykorzystaniem głębokiego sekwencjonowania RNA. Zastosowanie wspomnianego podejścia pozwoliło mi również na pozytywne zweryfikowanie hipotezy o rozpoznawaniu tych samych fragmentów RNA przez wszystkie paralogi MBNL, co ze względu na rozmieszczenie miejsca wiązania względem alternatywnego eksonu może prowadzić do promocji włączania lub wyłączania tego eksonu z mRNA. Konstrukty bazujące na przygotowanym przeze mnie minigenie hybrydowym zostały także wykorzystane w badaniach nad aranżacją strukturalną RNA miejsc wiązania dla białek MBNL jak i opracowaniu testu dla poszukiwania lub walidowania potencjalnych inhibitorów sekwestracji MBNL w DM. Ponadto, zaangażowany byłem w prace, które wykazały antagonistyczną rolę białek MBNL1 i SRSF1 w regulacji alternatywnego splicingu wielu eksonów poprzez wiązanie się do tych samych miejsc w RNA lub rearanżacje strukturalne miejsc wiązania dla obu białek. Przeprowadzona przeze mnie w innym projekcie analiza dystrybucji trzech alternatywnych eksonów w transkryptach genów *MBNL* wykazała ich zróżnicowane włączanie do mRNA w poszczególnych tkankach na różnych etapach rozwoju oraz zaburzenie tego procesu w mięśniach pacjentów z DM. Z kolei w projekcie wykazującym wpływ organizacji miejsc wiązania w RNA na aktywność regulatorową MBNL dostarczyłem danych o poziomie ekspresji poszczególnych badanych paralogów i izoform białek MBNL.

Preface

Introduction

Significant disproportion between the number of proteins expressed in human cells (over 100,000[1]) and the genes encoded in human genome (~20,300[2]) is driven by co-transcriptional process called alternative splicing (AS). More than 90% of human genes contain exons which are conditionally included or skipped during pre-mRNA splicing depending on temporal cellular requirements such as differentiation process, signal transduction, disease condition, etc.[3-7] (Fig. 1).

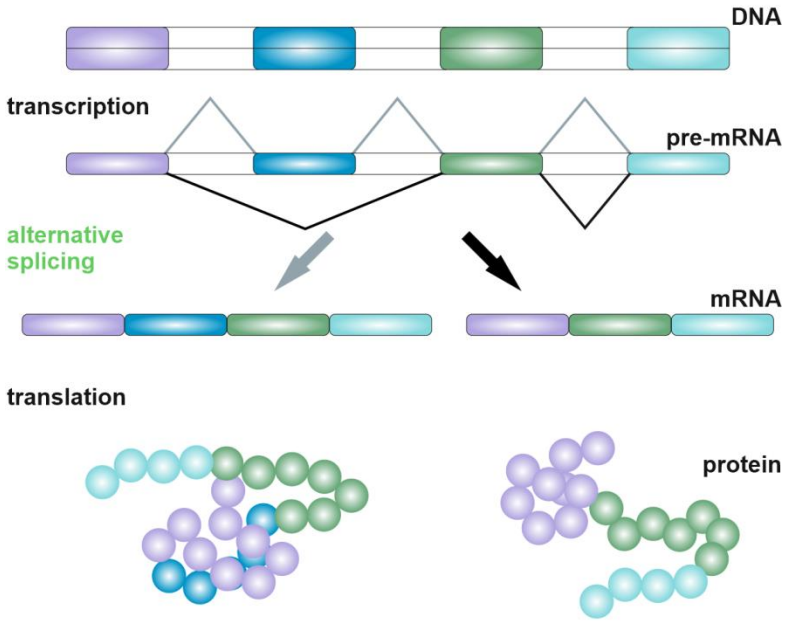


Figure 1. Scheme of alternative splicing. Colored boxes and circles correspond to protein-coding elements. White boxes correspond to introns.

Alternative splicing is regulated by *trans* factors - specific RNA-binding proteins (RBPs) which belong to different protein families such as Muscleblind-like (MBNL)[8], *CUG* repeat-binding protein, elav-like family (CELF)[9], RNA-binding fox-1 homolog (RBFOX)[10], etc. Alternative splicing factors affect spliceosome, multi-protein complex responsible for exon joining and excision of introns, by binding to *cis* elements - RNA fragments containing specific sequence motif recognized by particular RBP. Depending on the localization and function of *cis* elements, they can be divided into exonic or intronic splicing enhancer or silencer (ESE, ESS, ISE, ISS, respectively)[11,12]. Alternative splicing regulators often act antagonistically competing for the regulation of the same alternative exon[13-15]. Therefore, their general level and regulatory properties provide specific balance between mRNA isoforms differing in presence of alternative exon sequence processed from a particular pre-mRNA. That in turn, implies increased diversity of the proteome.

MBNL is a family of multifunctional RBPs pivotal in the fetal-to-adult developmental transition[8]. There are three members-paralogs: MBNL1, MBNL2 and MBNL3 conserved among species[16]. Their expression level and distribution depends on tissue type and developmental stage with MBNL1 most abundant in adult skeletal muscles and heart, MBNL2 in adult brain and the highest level of MBNL3 at embryonic and early postnatal stages and in regenerating muscles[17]. It has been found that, besides AS regulation, MBNL proteins coordinate alternative polyadenylation[18,19], stabilize[20] and transport mature mRNAs to the cellular compartments[21]. All paralogs present high sequence similarity and organization and recognize the same RNA consensus motif 5'-YGCY-3' (where Y is pyrimidine) which they bind with N-terminal domains having four zinc-finger CCCH-type (where C is cysteine, H is histidine) organized in two tandems separated with a linker sequence[22-26] (Fig. 2).

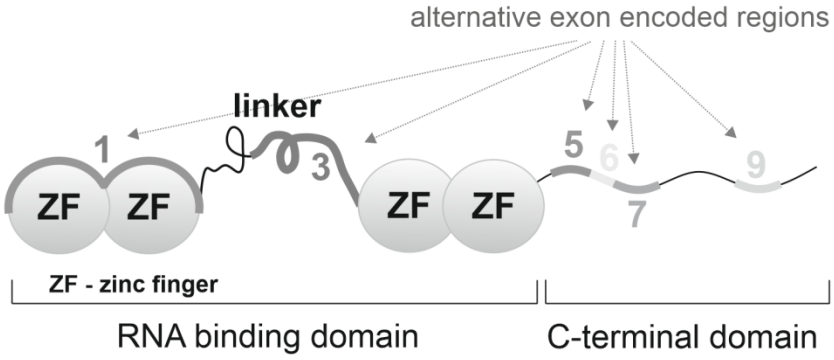


Figure 2. Schematic representation of MBNL protein organization.

Depending on the localization of consensus motif within regulated pre-mRNA MBNLs play a role of splicing activators when MBNL-binding motif is localized in downstream intron or repress alternative exon inclusion when bind to consensus motifs located within this exon or upstream intron[24,27] (Fig. 3).

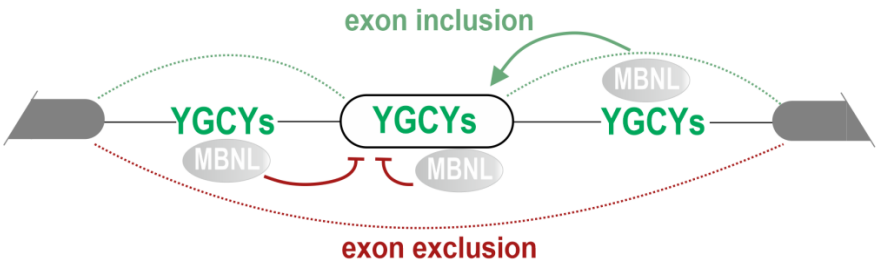


Figure 3. MBNL proteins' mode of action is dependent on consensus motif localization.

The mechanism of AS regulation by MBNL proteins is still poorly understood. Previous works have shown that MBNL1-mediated exon inclusion and exclusion in both cases is based on interaction with large subunit (65 kDa) of small nuclear U2 auxiliary factor (U2AF65). However, depending on motif localization, this interaction may occur in different manner. Warf and colleagues revealed that MBNL1 might inhibit alternative exon inclusion by direct competition with U2AF65 for binding site in vicinity of 5' region of regulated exon which leads to inhibition of U2 spliceosome component binding and skipping the exon[28]. On the other hand, another group showed that exon inclusion might be accelerated by indirect recruitment of U2AF65 by MNBL1 *via* RNA structure rearrangements or mobilization of other proteins in the 5' region of the downstream intron[29]. As many other primary transcripts, *MBNLs* contain alternative exons modulating their regulatory properties such as cellular localization or dimerization[30,31]. An interesting feature is MBNL1 autoregulation of two exons of its own pre-mRNA described recently by Konieczny and colleagues[32,33].

MBNL proteins are engaged in pathomechanism of myotonic dystrophy (DM) which is genetic, neuromuscular disorder affecting many organs and leading to muscle wasting, respiratory difficulties, cataracts, insulin resistance, etc.[17,34] There are two types of DM, type 1 (DM1) and type 2 (DM2) with less severe symptoms, where biosynthesis of toxic RNA occurs as a consequence of uncontrolled expansion of short nucleotide repeats in genome. In DM1 expansion of triplet 5'-CTG-3' occurs in 3'UTR region of *DMPK* (*Dystophia Myotonica Protein Kinase*) gene reaching up to few thousands[35,36]. DM2 characterizes with up to 11,000 of 5'-CCTG-3' repeats in *CNBP* (*CCHC-type zinc finger nucleic acid binding protein*) intron 1[37,38]. In both variants transcription of expanded repeats results in formation of toxic $(CUG)^{exp}$ or $(CCUG)^{exp}$ RNA which tends to form stable secondary structures and because of presence of repeating YGCY motifs (C<u>UGCU</u>G or CC<u>UGCC</u>UG) they sequester MBNL proteins very effectively[39,40]. Sequestration of MBNLs leads to massive misregulation of alternative splicing of MBNLs' target transcripts which drives subsequent changes in the proteome and induces DM phenotype[15,21,27,41-46].

A need to better understand the role of MBNL proteins has resulted in a number of works in which high-throughput sequencing techniques were applied. One of such approaches which enables to shed more light on group of transcripts regulated by MBNLs is Cross-Linking and Immuno Precipitation combined with deep sequencing (CLIP-seq)[47]. It is based on fixing of RNA-RBP interactions, purification of such complexes with antibody specific to particular protein and deep sequencing of co-purified RNA fragments. Data from CLIP-seq experiments for MBNLs have revealed hundreds of novel potential MBNL-binding sites which need to be validated experimentally[14,19,21,48,49]. One of possible approaches to verify CLIP-seq results is application of antisense oligonucleotides (AONs) strategy.

AONs are synthetic, short (up to ~25 nucleotides) derivatives of nucleic acids capable of binding to complementary sequences widely used to target specific transcripts, e.g. mutation-derived novel splice sites or therapeutic exon-skipping in Duchenne Muscular Dystrophy (DMD) by exon-intron junction masking[50-53]. Due to chemical modifications they are subjected, three generations of AON units can be distinguished: 1st with substitution of phosphor with sulfur in phosphate backbone, 2nd with moieties added to sugar ring and 3rd with structure of sugar ring modification[54-58] (Tab. 1). Depending on modification type, AONs reveal different properties such as stability, specificity to the targeted sequence or cytotoxicity. It also defines how AON affects the target. AONs of 1st generation are known to initiate degradation of the transcript *via* RNase H recruitment while AONs of 2nd and 3rd by hybridizing to the RNA target, can affect binding of regulatory factors, e.g. ribosome or RBPs[59,60]. Therefore, while some AONs can play a role of silencers *via* inhibition of transcription or translation, certain types can be applied as blockers to indicate binding site of particular protein or to modulate pre-mRNA metabolism, including splicing.

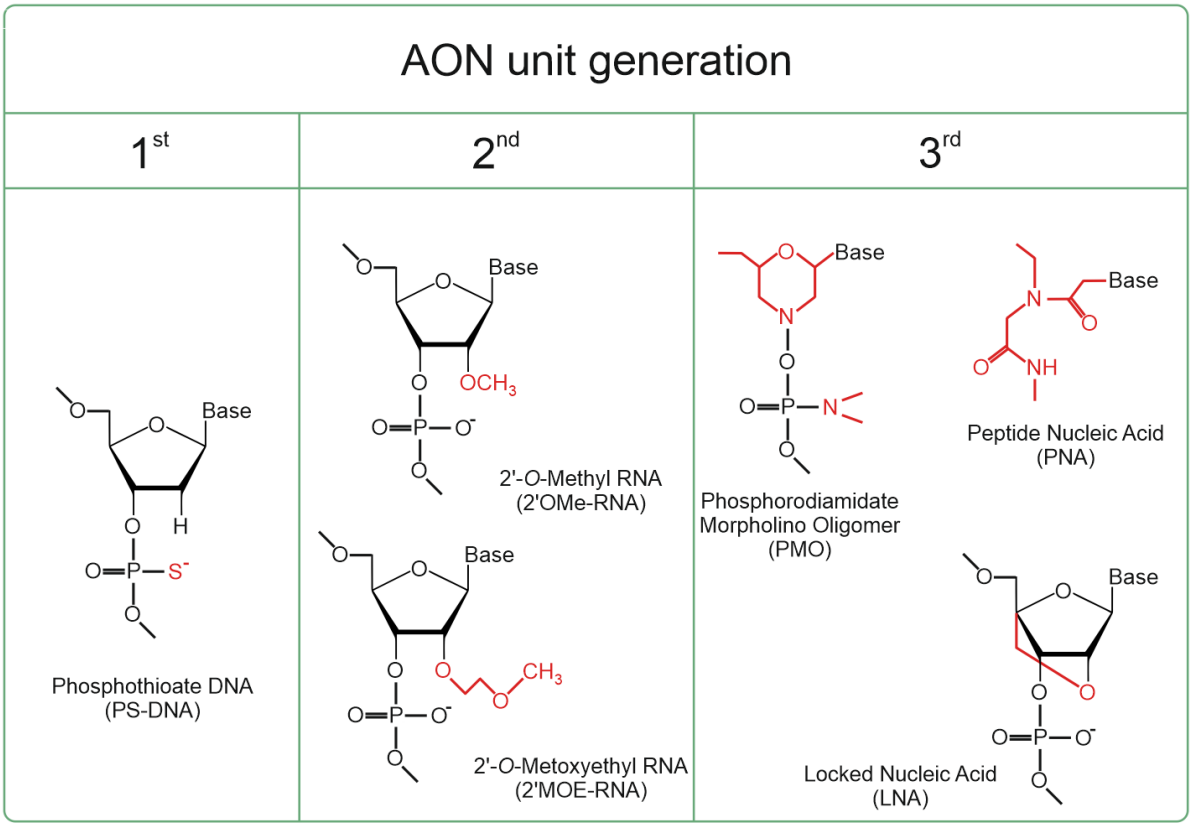


Table 1. Examples of DNA and RNA modifications (marked in red) in three generations of AON units.

Aims

Mechanism of action of MBNL proteins is still not entirely explored. The aim of my study was to better understand how MBNLs recognize their RNA targets in process of alternative splicing and what is the function of different splicing isoforms of MBNL proteins. Specific aims of my work were:

- to develop reliable and effective assays for verification of potential MBNL-binding sites predicted based on whole transcriptome-scale CLIP-seq experiments;
- to give a deeper insight into MBNL proteins' mode of action in regulation of alternative splicing and
- to characterize distribution of splicing isoforms of MBNLs in different tissues and pathologic stages.

Results

The results of my work were described in three experimental publications:

Cywoniuk, P., Taylor, K., Sznajder, L. J. & Sobczak, K. **Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions.** *Sci Rep* **7**, 17587, doi:10.1038/s41598-017-17816-x (2017).

Sznajder, L. J., Michalak, M., Taylor, K., Cywoniuk, P., Kabza, M., Wojtkowiak-Szlachcic, A., Matłoka, M., Konieczny, P. & Sobczak, K. **Mechanistic determinants of MBNL activity.** *Nucleic Acids Res* **44**, 10326-10342, doi:10.1093/nar/gkw915 (2016).

Taylor, K., Sznajder, L. J., Cywoniuk, P., Thomas, J. D., Swanson, M. S. & Sobczak, K. **MBNL splicing activity depends on RNA binding site structural context.** *Nucleic Acids Res* **46**, 9119-9133, doi:10.1093/nar/gky565 (2018).

In the article entitled "**Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions**"[61] we described sensitive assays for experimental validation of potential MBNL-binding sites based on application of antisense oligonucleotides combined with specifically designed "hybrid" splicing minigenes. In this work in a series of *in cellula* experiments, preceded by *in vitro* examination performed by dr Katarzyna Taylor, I found that about 20-nucleotides long 2'-*O*-Methyl RNAs with phosphorothioate backbone (2'*O*Me-PS RNAs) are the most effective AONs that block interaction between MBNL protein and targeted RNA sequence (Fig. 1d, g in[61]). Subsequently, we established several criteria for the selection of transcript regions containing potential MBNL-binding sites predicted based on CLIP-seq experiments performed by dr Łukasz Sznajder and dr Patryk Konieczny[62]. Based on that, I designed five 2'*O*Me-PS AONs targeting potential MBNL1-binding sites in intronic or exonic regions of *Atp2a1*, *Pphln1*, *Nfix*, *Ldb3* and *Mbnl1*pre-mRNAs (AON against *NASP* transcript was designed by dr Sznajder). Series of experiments in cells treated with these AONs revealed that only MBNL1-binding sites localized in introns could be verified properly using AON approach. In case of exonic targets AONs provoked general disruption of splicing of targeted alternative exons (Fig. 2d in[61]). Therefore, I prepared a series of *Atp2a1* gene-based minigenes in which original intronic MBNL-binding cassette was replaced with studied elements derived from other analyzed transcripts with potential MBNL1-binding sites. Proper definition of studied RNA regions and preservation of their native structure was supported by *in vitro* analysis performed by dr Taylor. My experiments have shown that transposition of particular pre-mRNA fragment derived from either exonic or intronic region into the context of downstream intron into the "hybrid" minigene restores response of alternative exon of this minigene to the MBNL1 activity. Moreover, I showed that subsequent treatment of AON specific to analyzed hybrid fragment strongly inhibits interaction of MBNL1 with targeted RNAs (Fig. 3 in[61]). Such approach combining AON strategy and system of hybrid minigenes is useful to validate potential binding sites for MBNL proteins but could also be applied for other RBPs. Furthermore, in order to explain unexpected splicing failure in AON-targeted exonic regions, together with dr Taylor, we identified serine/arginine-rich splicing factor 1 (SRSF1) as potential antagonist/modulator of MBNL activity in regulation of splicing of two alternative exons containing YGCY motifs located within these exons (Fig. 4 in[61]). Last but not least, dr Taylor revealed that obtained minigene with replaceable MBNL-binding element could be used for screenings for potential therapeutic agents inhibiting MBNL proteins sequestration on expanded CUG repeats.

Publication entitled "**Mechanistic determinants of MBNL activity**"[62] focused on properties of three MBNL paralogs and their splicing isoforms. Dr Łukasz Sznajder and other team members compared several features such as splicing regulatory potential, subcellular localization and contribution to DM pathomechanism of a number of MBNL proteins. Our results show that MBNLs differ in splicing activity whereas MBNL1 reveals as the strongest and MBNL3 the weakest splicing regulator. Presence of three analyzed alternative exons within *MBNLs*' transcripts has been shown to modulate activity of the proteins in different manner. Also subcellular distribution differs significantly among paralogs with MBNL1 and 2 dominant location in nucleus and MBNL3 in cytoplasm. In this work I compared the distribution of three alternative exons in *MBNL1*, *2* and *3* mRNAs among human tissues at fetal or adult developmental stage and in samples derived from DM1 and DM2 patients (Fig. 1B in[62]). Using described above assay combining AON strategy with hybrid splicing minigene I also confirmed that all three MBNL paralogs bind precisely the same RNA sequence but with different strength of effect (Fig. 3F, G in[62]).

An article "**MBNL splicing activity depends on RNA binding site structural context**"[63] was devoted for better understanding the impact of organization of YGCY consensus motifs in pre-mRNAs on MBNLs' affinity and, therefore, splicing activity. Dr Katarzyna Taylor using *in vitro* and *in cellula* systems, including hybrid minigene strategy, showed that MBNLs' binding preferences strongly depend on the number and composition of YGCY consensus motifs within secondary structure context of pre-mRNA. Obtained results indicate that MBNL1 affinity is significantly higher for single-stranded regions of RNA containing four such sequence motifs (at least two are necessary for MBNL1 binding). Furthermore, spatial arrangement of YGCY motifs is pivotal and significantly decreases the MBNL1 affinity when particular motifs are located on opposite sides of RNA hairpin structure. Additionally, we reported competitive relationship between activity of MBNL1 and MBNL3 which synergistic or antagonistic effect strongly depends on organization of YGCY motifs (overlapping or sequential, respectively). In this work I provided quantity-control experiments of particular MBNL paralog/isoform level using classic western blot as well as method based on visualization of fluorescence of eGFP-fused proteins directly in acrylamide gel (Fig. S2c, 6E, S10c, and S11a, b in[63]).

References

1 Savage, N. Mobile data: Made to measure. *Nature* **527**, S12-13, doi:10.1038/527S12a (2015).
2 Genome Reference Consortium Human Build 38 Patch Release 2, (2014).
3 Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat Rev Genet* **17**, 19-32, doi:10.1038/nrg.2015.3 (2016).
4 Sanchez, S. E., Petrillo, E., Kornblihtt, A. R. & Yanovsky, M. J. Alternative splicing at the right time. *RNA Biol* **8**, 954-959, doi:10.4161/rna.8.6.17336 (2011).
5 Hallegger, M., Llorian, M. & Smith, C. W. Alternative splicing: global insights. *FEBS J* **277**, 856-866, doi:10.1111/j.1742-4658.2009.07521.x (2010).
6 Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457-463, doi:10.1038/nature08909 (2010).
7 Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470-476, doi:10.1038/nature07509 (2008).

8 Pascual, M. *et al*. The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation* **74**, 65-80, doi:10.1111/j.1432-0436.2006.00060.x (2006).
9 Dasgupta, T. & Ladd, A. N. The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *Wiley Interdiscip Rev RNA* **3**, 104-121, doi:10.1002/wrna.107 (2012).
10 Jin, Y. *et al.* A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J* **22**, 905-912, doi:10.1093/emboj/cdg089 (2003).
11 Matlin, A. J., Clark, F. & Smith, C. W. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* **6**, 386-398, doi:10.1038/nrm1645 (2005).
12 Robberson, B. L., Cote, G. J. & Berget, S. M. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol* **10**, 84-94 (1990).
13 Kino, Y. *et al.* MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res* **37**, 6477-6490, doi:10.1093/nar/gkp681 (2009).
14 Wang, E. T. *et al.* Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res* **25**, 858-871, doi:10.1101/gr.184390.114 (2015).
15 Ohsawa, N., Koebis, M., Mitsuhashi, H., Nishino, I. & Ishiura, S. ABLIM1 splicing is abnormal in skeletal muscle of patients with DM1 and regulated by MBNL, CELF and PTBP1. *Genes Cells* **20**, 121-134, doi:10.1111/gtc.12201 (2015).
16 Oddo, J. C., Saxena, T., McConnell, O. L., Berglund, J. A. & Wang, E. T. Conservation of context-dependent splicing activity in distant Muscleblind homologs. *Nucleic Acids Res* **44**, 8352-8362, doi:10.1093/nar/gkw735 (2016).
17 Kanadia, R. N. *et al.* A muscleblind knockout model for myotonic dystrophy. *Science* **302**, 1978-1980, doi:10.1126/science.1088583 (2003).
18 Batra, R., Manchanda, M. & Swanson, M. S. Global insights into alternative polyadenylation regulation. *RNA Biol* **12**, 597-602, doi:10.1080/15476286.2015.1040974 (2015).
19 Batra, R. *et al.* Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell* **56**, 311-322, doi:10.1016/j.molcel.2014.08.027 (2014).
20 Masuda, A. *et al.* CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. *Sci Rep* **2**, 209, doi:10.1038/srep00209 (2012).
21 Wang, E. T. *et al.* Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150**, 710-724, doi:10.1016/j.cell.2012.06.041 (2012).
22 Teplova, M. & Patel, D. J. Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat Struct Mol Biol* **15**, 1343-1351, doi:10.1038/nsmb.1519 (2008).
23 He, F. *et al.* Solution structure of the RNA binding domain in the human muscleblind-like protein 2. *Protein Sci* **18**, 80-91, doi:10.1002/pro.17 (2009).
24 Goers, E. S., Purcell, J., Voelker, R. B., Gates, D. P. & Berglund, J. A. MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res* **38**, 2467-2484, doi:10.1093/nar/gkp1209 (2010).
25 Lambert, N. *et al.* RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell* **54**, 887-900, doi:10.1016/j.molcel.2014.04.016 (2014).
26 Cass, D. *et al.* The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements. *BMC Mol Biol* **12**, 20, doi:10.1186/1471-2199-12-20 (2011).
27 Du, H. *et al.* Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat Struct Mol Biol* **17**, 187-193, doi:10.1038/nsmb.1720 (2010).
28 Warf, M. B., Diegel, J. V., von Hippel, P. H. & Berglund, J. A. The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad Sci U S A* **106**, 9203-9208, doi:10.1073/pnas.0900342106 (2009).
29 Echeverria, G. V. & Cooper, T. A. Muscleblind-like 1 activates insulin receptor exon 11 inclusion by enhancing U2AF65 binding and splicing of the upstream intron. *Nucleic Acids Res* **42**, 1893-1903, doi:10.1093/nar/gkt1020 (2014).
30 Terenzi, F. & Ladd, A. N. Conserved developmental alternative splicing of muscleblind-like (MBNL) transcripts regulates MBNL localization and activity. *RNA Biol* **7**, 43-55 (2010).
31 Tran, H. *et al.* Analysis of exonic regions involved in nuclear localization, splicing activity, and dimerization of Muscleblind-like-1 isoforms. *J Biol Chem* **286**, 16435-16446, doi:10.1074/jbc.M110.194928 (2011).
32 Konieczny, P., Stepniak-Konieczna, E., Taylor, K., Sznajder, L. J. & Sobczak, K. Autoregulation of MBNL1 function by exon 1 exclusion from MBNL1 transcript. *Nucleic Acids Res*, doi:10.1093/nar/gkw1158 (2016).
33 Konieczny, P., Stepniak-Konieczna, E. & Sobczak, K. MBNL expression in autoregulatory feedback loops. *RNA Biol* **15**, 1-8, doi:10.1080/15476286.2017.1384119 (2018).
34 Udd, B. & Krahe, R. The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *Lancet Neurol* **11**, 891-905, doi:10.1016/S1474-4422(12)70204-1 (2012).
35 Brook, J. D. *et al.* 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* **69**, 385 (1992).
36 Fardaei, M., Larkin, K., Brook, J. D. & Hamshere, M. G. In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res* **29**, 2766-2771 (2001).
37 Ranum, L. P., Rasmussen, P. F., Benzow, K. A., Koob, M. D. & Day, J. W. Genetic mapping of a second myotonic dystrophy locus. *Nat Genet* **19**, 196-198, doi:10.1038/570 (1998).
38 Liquori, C. L. *et al.* Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864-867, doi:10.1126/science.1062125 (2001).

39 Miller, J. W. *et al.* Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* **19**, 4439-4448, doi:10.1093/emboj/19.17.4439 (2000).
40 Fardaei, M. *et al.* 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 (2002).
41 Lin, X. *et al.* Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum Mol Genet* **15**, 2087-2097, doi:10.1093/hmg/ddl132 (2006).
42 Kanadia, R. N. *et al.* Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc Natl Acad Sci U S A* **103**, 11748-11753, doi:10.1073/pnas.0604970103 (2006).
43 Yamashita, Y. *et al.* LDB3 splicing abnormalities are specific to skeletal muscles of patients with myotonic dystrophy type 1 and alter its PKC binding affinity. *Neurobiol Dis* **69**, 200-205, doi:10.1016/j.nbd.2014.05.026 (2014).
44 Klinck, R. *et al.* RBFOX1 cooperates with MBNL1 to control splicing in muscle, including events altered in myotonic dystrophy type 1. *PLoS One* **9**, e107324, doi:10.1371/journal.pone.0107324 (2014).
45 Carpentier, C. *et al.* Tau exon 2 responsive elements deregulated in myotonic dystrophy type I are proximal to exon 2 and synergistically regulated by MBNL1 and MBNL2. *Biochim Biophys Acta* **1842**, 654-664, doi:10.1016/j.bbadis.2014.01.004 (2014).
46 Suenaga, K. *et al.* Muscleblind-like 1 knockout mice reveal novel splicing defects in the myotonic dystrophy brain. *PLoS One* **7**, e33218, doi:10.1371/journal.pone.0033218 (2012).
47 Ule, J., Jensen, K., Mele, A. & Darnell, R. B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**, 376-386, doi:10.1016/j.ymeth.2005.07.018 (2005).
48 Goodwin, M. *et al.* MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain. *Cell Rep* **12**, 1159-1168, doi:10.1016/j.celrep.2015.07.029 (2015).
49 Charizanis, K. *et al.* Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron* **75**, 437-450, doi:10.1016/j.neuron.2012.05.029 (2012).
50 Chan, J. H., Lim, S. & Wong, W. S. Antisense oligonucleotides: from design to therapeutic application. *Clin Exp Pharmacol Physiol* **33**, 533-540, doi:10.1111/j.1440-1681.2006.04403.x (2006).
51 Friedman, K. J. *et al.* Correction of aberrant splicing of the cystic fibrosis transmembrane conductance regulator (CFTR) gene by antisense oligonucleotides. *J Biol Chem* **274**, 36193-36199 (1999).
52 Sierakowska, H., Sambade, M. J., Agrawal, S. & Kole, R. Repair of thalassemic human beta-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc Natl Acad Sci U S A* **93**, 12840-12844 (1996).
53 Aartsma-Rus, A. & van Ommen, G. J. Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications. *RNA* **13**, 1609-1624, doi:10.1261/rna.653607 (2007).
54 Eckstein, F. Phosphorothioate oligodeoxynucleotides: what is their origin and what is unique about them? *Antisense Nucleic Acid Drug Dev* **10**, 117-121, doi:10.1089/oli.1.2000.10.117 (2000).
55 Yu, R. Z. *et al.* Tissue disposition of 2'-O-(2-methoxy) ethyl modified antisense oligonucleotides in monkeys. *J Pharm Sci* **93**, 48-59, doi:10.1002/jps.10473 (2004).
56 Kurreck, J., Wyszko, E., Gillen, C. & Erdmann, V. A. Design of antisense oligonucleotides stabilized by locked nucleic acids. *Nucleic Acids Res* **30**, 1911-1918 (2002).
57 Vester, B. & Wengel, J. LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. *Biochemistry* **43**, 13233-13241, doi:10.1021/bi0485732 (2004).
58 Amantana, A. & Iversen, P. L. Pharmacokinetics and biodistribution of phosphorodiamidate morpholino antisense oligomers. *Curr Opin Pharmacol* **5**, 550-555, doi:10.1016/j.coph.2005.07.001 (2005).
59 Wu, H. *et al.* Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. *J Biol Chem* **279**, 17181-17189, doi:10.1074/jbc.M311683200 (2004).
60 Zalachoras, I. *et al*. Antisense-mediated RNA targeting: versatile and expedient genetic manipulation in the brain. *Front Mol Neurosci* **4**, 10, doi:10.3389/fnmol.2011.00010 (2011).
61 Cywoniuk, P., Taylor, K., Sznajder, Ł. & Sobczak, K. Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions. *Sci Rep* **7**, 17587, doi:10.1038/s41598-017-17816-x (2017).
62 Sznajder, L. J. *et al.* Mechanistic determinants of MBNL activity. *Nucleic Acids Res* **44**, 10326-10342, doi:10.1093/nar/gkw915 (2016).
63 Taylor, K. *et al.* MBNL splicing activity depends on RNA binding site structural context. *Nucleic Acids Res* **46**, 9119-9133, doi:10.1093/nar/gky565 (2018).

Publications
and
co-author statements

SCIENTIFIC REPORTS

OPEN

Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions

Received: 31 August 2017
Accepted: 1 December 2017
Published online: 14 December 2017

Piotr Cywoniuk[1], Katarzyna Taylor[1], Łukasz J. Sznajder[1,2] & Krzysztof Sobczak[1]

Alternative splicing is a complex process that provides a high diversity of proteins from a limited number of protein-coding genes. It is governed by multiple regulatory factors, including RNA-binding proteins (RBPs), that bind to specific RNA sequences embedded in a specific structure. The ability to predict RNA-binding regions recognized by RBPs using whole-transcriptome approaches can deliver a multitude of data, including false-positive hits. Therefore, validation of the global results is indispensable. Here, we report the development of an efficient and rapid approach based on a modular hybrid minigene combined with antisense oligonucleotides to enable verification of functional RBP-binding sites within intronic and exonic sequences of regulated pre-mRNA. This approach also provides valuable information regarding the regulatory properties of pre-mRNA, including the RNA secondary structure context. We also show that the developed approach can be used to effectively identify or better characterize the inhibitory properties of potential therapeutic agents for myotonic dystrophy, which is caused by sequestration of specific RBPs, known as *muscleblind*-like proteins, by mutated RNA with expanded *CUG* repeats.

Alternative splicing (AS) is a co-transcriptional process that leads to a significant increase in proteome diversity[1–3]. AS leads to the formation of different mRNA variants from the same precursor transcript due to conditional inclusion or exclusion of alternative exons or their parts depending on the tissue type, developmental stage or disease conditions[4]. The process is controlled by a group of tissue-specific *trans*-elements that act as alternative splicing regulators, including *muscleblind*-like proteins (MBNLs), *CUG* triplet repeat binding, elav-like family proteins (CELFs) and RNA-binding fox homolog proteins (RBFOXs). They define the splicing pattern of regulated alternative exons by binding to specific RNA sequence/structure motifs (*cis*-elements)[5–7]. The inclusion or exclusion of alternative exons is primarily determined by the interplay of several alternative splicing regulators and other types of regulatory proteins, including splicing enhancers and silencers (i.e., heterogeneous nuclear ribonucleoproteins (hnRNPs) or serine/arginine-rich splicing factors (SRSFs))[8].

MBNL proteins are key developmental regulators that are involved in different human diseases—e.g., myotonic dystrophy (DM)[5,9–12] or Fuchs endothelial corneal dystrophy[13]. Interestingly, the MBNL1, MBNL2 and MBNL3 paralogs possess both redundant and unique functions in RNA processing and disease pathogenesis[14]. The best studied MBNL function is alternative splicing regulation[5,11,12,15]. All paralogs recognize RNA targets through an RNA-binding domain that contains four zinc fingers (ZF) organized into two tandems[11]. The MBNL-binding RNA motif is specified as a 5′-YGCY-3′ sequence (Y represents a pyrimidine)[16,17] with preferentially single-stranded pyrimidines[18]. Depending on the location of the MBNL-binding motifs, whether upstream or downstream of the alternatively spliced exon, MBNLs serve either as repressors or activators of alternative exon inclusion, respectively[17,19]. A similar mode of action was previously described also for RBFOXs[20] and CELFs[21]. Although many functional MBNL-binding motifs within pre-mRNAs (e.g., *Clcn1*, *cTNT*, *Atp2a1* and *Mbnl1*) have been well described, most remain unknown[14,22–25].

[1]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, 61-614, Poznan, Poland. [2]Present address: Center for NeuroGenetics and the Genetics Institute, Department of Molecular Genetics and Microbiology, College of Medicine, University of Florida, Gainesville, Florida, 32610-3610, USA. Piotr Cywoniuk and Katarzyna Taylor contributed equally to this work. Correspondence and requests for materials should be addressed to K.S. (email: ksobczak@amu.edu.pl)

The last decade has introduced several technologies for high-throughput analysis of protein/RNA interactions, including MBNL/RNA[21,26], especially cross-linking and immunoprecipitation combined with deep sequencing (CLIP-seq)[27] with subsequent modifications: individual-nucleotide resolution (iCLIP)[28] and photoactivatable ribonucleoside-enhanced (PAR-CLIP)[29]. In general, CLIP-seq is a multistep method that starts from the physical cross-linking of RBPs with RNA sequences *in situ* using UV light. The protein/RNA complexes are subsequently purified by immunoprecipitation and molecular mass-based separation. Eventually, short RBP-bound RNA fragments are high-throughput sequenced[30,31], followed by data filtration. The obtained overlapped reads form CLIP-seq clusters, which are considered to be RBP-bound regions. This method generates an enormous amount of data, enabling deeper insight into the proteins' regulatory mode of action but also provides a high percentage of non-functional RNA binding sites and/or false-positive results. Thus, the potential RNA targets and RBP-binding motifs must still be verified in individual tests.

Herein, we described a new verification strategy that uses antisense oligonucleotides (AONs) and highly redundant hybrid splicing minigenes as a genetic construct carrying an alternative exon as well as interchangeable MBNL-sensitive regulatory elements within an upstream intron. The strategy is designed for the rapid, simple, reproducible, and efficient determination of functional protein/RNA interactions based on CLIP-seq or other whole transcriptome experiments. It also allows evaluation of the regulatory properties of MBNL-binding sites. Moreover, we show that the hybrid minigenes can be used to reveal and elucidate potential crosstalk between different RBPs in the regulation of alternative splicing. This method was tested for MBNL-specific RNA-binding regions but could be applicable for other RBPs. Finally, we emphasize the usefulness of hybrid minigenes to study the efficacy of potential therapeutic agents for inhibiting the interaction between MBNLs and expanded *CUG* repeats (CUG^{exp}) associated with myotonic dystrophy type 1 (DM1) development.

Results and Discussion

Two AON types effectively block functional MBNL/RNA interactions. First, we set out to establish the most effective strategy to block protein/RNA interactions. Thus, we designed several types of AONs that target a well-defined MBNL-binding region in the *Atp2a1* transcript and tested their inhibitory properties *in vitro* and *in cellulo*. *Atp2a1* pre-mRNA contains alternative exon 22 (ex22), which as confirmed by mutagenesis, is positively regulated by the binding of all three MBNL paralogs to two YGCY motif-containing regions (region #1 and #2) that are localized within intron 22, ~110 nucleotides (nt) downstream of ex22 (Fig. 1a,c)[14,24]. We confirmed the MBNL-dependent alternative splicing of ex22 by the silencing of both *MBNL1* and *MBNL2* (Fig. 1b). Because the efficiency of MBNL binding may be modulated by the RNA's structure[18], we determined the secondary structure for a 145-nt-long fragment of *Atp2a1* RNA containing regions #1 and #2 through limited cleavage with two enzymatic probes that recognize single-stranded RNA[32] (Fig. 1c, Supplementary Fig. S2a). We observed that both YGCY-rich regions were localized in the 5′-part of the semi-stable hairpin structure, surrounded by several internal loops (Fig. 1c).

To elaborate the interaction of MBNL1 with crucial YGCY motifs as well as the capacity of AONs to inhibit MBNL1 *in vitro*, we used a filter binding assay (FBA). First, we confirmed the sensitivity and specificity of this assay for a few RNA fragments that have previously been described in the literature as targets of MBNL1[24,25,33]. We did not observe any interaction of this protein with four negative control RNAs (RNA-Ctrl1-4), with or without YGCY motifs (Supplementary Fig. S1a–c). Second, we designed RNA-based AONs that selectively blocked either individual region #1 or #2 of *Atp2a1*-RNA or both simultaneously. We tested AONs with either a locked nucleic acid (LNA) or 2′-*O*-methyl (2′OMe) modification of the RNA structure with or without a phosphorothioated backbone (PS) and DNA-based AONs. The LNAs (LNA(-PS)#1 and #2) contained 10-nt residues and were complementary to individual regions #1 or #2, while the 2′OMe- and DNA-type AONs were 21-nt-long and covered both regions (2′OMe(-PS) and DNA) (Fig. 1c). Having performed the *in vitro* assay, we observed that the short LNA#2 did not prevent formation of MBNL1/RNA complexes, LNA#1 only slightly deteriorated them, and 21-nt-long AONs suppressed MBNL1/RNA complexes by approximately 8-fold (Fig. 1d, Supplementary Fig. S2b). Consistently, we observed diminished MBNL1 affinity to *Atp2a1*-RNA mutants with nucleotide substitutions in the YGCY motifs within regions #1 and #2 (Fig. 1e), confirming their previously described functional significance[24]. The lack of complete inhibition of complex formation suggests binding of MBNL1 to other YGCY motifs present within this RNA, but with significantly lower affinity. To investigate this hypothesis, we used DNA-based AONs and applied mutagenesis. We found another YGCY motif upstream of region #1 that was important for MBNL1 binding *in vitro* (Supplementary Fig. S2c–e). We showed that short LNAs were the least effective *in vitro*, most likely due to covering only a part of the motifs bound by MBNL1.

Next, we assessed the inhibitory potential of the same RNA-based AONs on MBNL-dependent alternative splicing of ex22. First, we co-transfected cells with the *Atp2a1* minigene, which encompasses a pre-mRNA sequence from exons 21 to 23 (*Atp2a1WT*), and the MBNL1-overexpressing vector. RT-PCR analysis showed substantial MBNL1-dependent enhancement of alternative ex22 inclusion (Fig. 1f,g). With the addition of different *Atp2a1*-specific AONs to this system, we observed almost complete inhibition of ex22 inclusion in LNA-PS#1- and 2′OMe-PS-treated cells as well as a significant change in LNA-PS#2. All AONs without the PS modification revealed no inhibitory effect on ex22 regulation (Fig. 1g). Moreover, the impact of 2′OMe-PS on repression of ex22 inclusion was observed even at the low 25-nM concentration (Fig. 1g). Importantly, for both LNA-PSs, but not for 2′OMe-PS, we observed a strong effect on three other MBNL-dependent alternative exons (Supplementary Fig. S2f). We expect that this effect may have been caused by LNA-PSs due to the relatively short sequence length and possible binding to a prevalent MBNL-binding motif present in many other transcripts.

In summary, we defined the structural features of the MBNL1-binding regions within *Atp2a1*-RNA. We demonstrated that the use of AONs *in vitro*, regardless of their chemistry, was an efficient tool to define the best MBNL-binding motifs. However, for *in cellulo* assays, the type of chemical modification is important. Both parameters—the length and chemical modification of 2′OMe-PS—determine its efficacy to inhibit MBNL/RNA

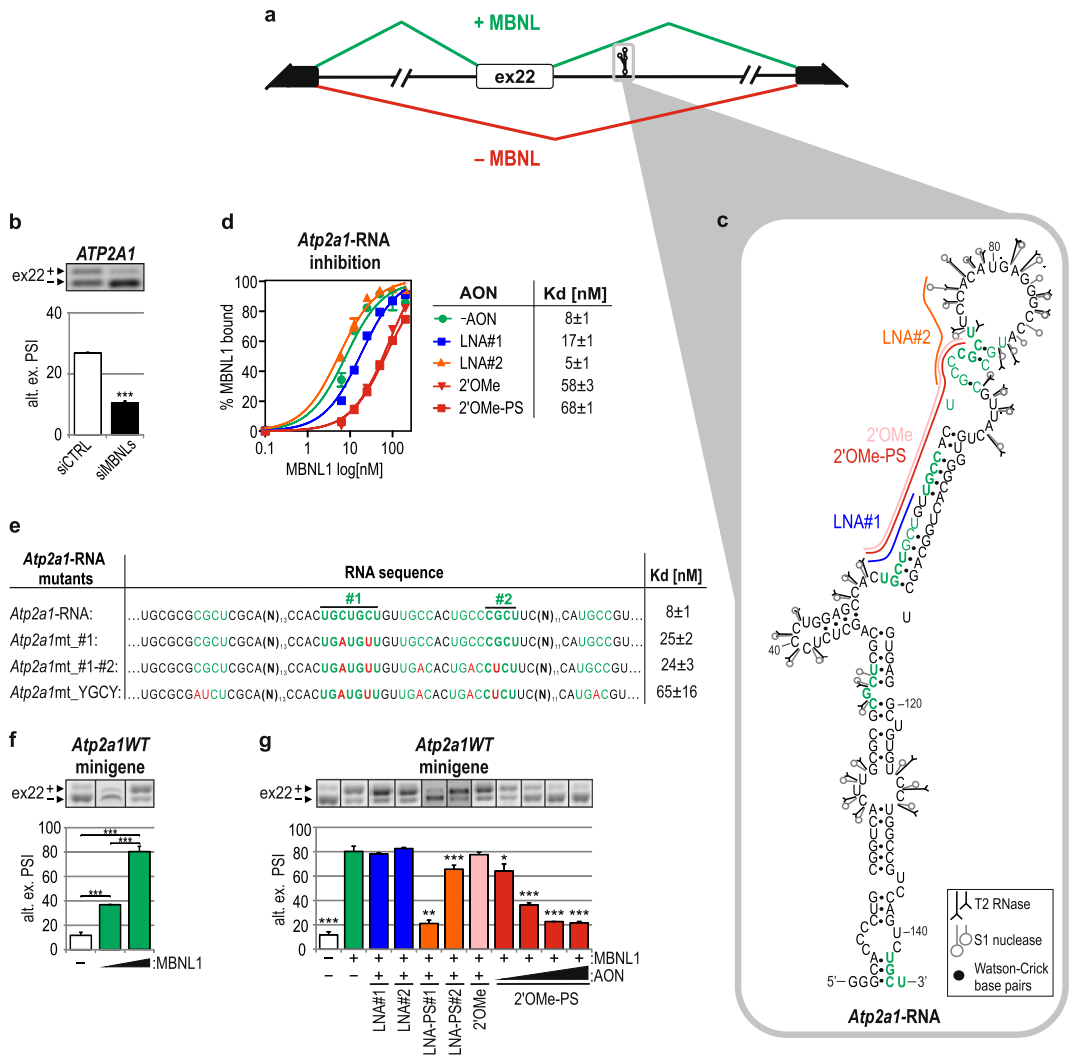


Figure 1. Alternative splicing of *Atp2a1* ex22 is MBNL dependent and efficiently distorted upon AON targeting of MBNL-binding regions. (**a**) A scheme illustrating the alternative splicing pattern of *Atp2a1* ex22 in the presence (green) or absence (red) of MBNLs. Black and white boxes represent constitutive and alternative exons, respectively. The RNA fragment, which was analyzed *in vitro*, comprising MBNL-binding regions #1 and #2 is marked with the gray frame. (**b**) Percentage of ex22 inclusion in *ATP2A1* endogenous mRNA upon *MBNL1,2* silencing in human cells. PSI, percent spliced in index, providing the inclusion level of an alternative exon; n = 3. (**c**) An experimentally determined secondary structure of *Atp2a1*-RNA with the YGCY motifs marked with green and in bold if present in humans and mice. Cleavages induced by RNase T2 and nuclease S1 are shown. The AONs bound *in vitro* to MBNL-binding regions #1 and #2 are accordingly indicated by blue (LNA#1), orange (LNA#2) and red lines (2′OMe/2′OMe-PS). AON, antisense oligonucleotide; 2′OMe-PS, phosphorothioated 2′-*O*-methyl; LNA, locked nucleic acid. The cleavage sites and intensities of the selected probes are shown using symbols explained in the inset. (**d**) Quantification of FBA showing the blocking properties of AONs *in vitro*. On the right, the dissociation constant (Kd) of MBNL1/*Atp2a1*-RNA complexes without (-AON) or with different AON applications; n = 4. (**e**) Quantification of FBA showing reduced MBNL1 affinity to *Atp2a1*-RNA with point mutations, especially YGCYs (marked with red). Regions #1 and #2, significant for MBNL binding, are also indicated; n = 4. (**f**) MBNL1 dose-dependent percentage of alternative ex22 inclusion in mRNA from the *Atp2a1WT* minigene in HeLa cells transfected with 200 or 500 ng of the MBNL1 expression vector per well; n = 3. (**g**) Percentage of alternative ex22 inclusion in *Atp2a1WT* mRNA upon MBNL1 overexpression and treatment with different AONs at 25–100 nM (2′OMePS) or 100 nM (others). The obtained results were compared with those of the control experiment with MBNL1 overexpression and no AON treatment (green bar); n = 3.

complex formation, consistent with previously published results that showed the applicability of 2′OMe-PS AONs for blocking many different *cis*-regulatory RNA elements, including intron-exon junctions or mutation-induced alternative splice sites[34–37]. Hence, the 2′OMe-PS type was chosen for further experiments.

AONs indicate functional MBNL-binding regions within introns but not exons. To confirm the applicability of AONs for the verification of RBP/RNA interactions, we selected five additional pre-mRNAs, besides *Atp2a1*, with potential MBNL-binding motifs within an intronic sequence downstream or upstream of *Pphln1* ex6 and *NASP ex7*, respectively, as well as within exonic sequences of ex10 of *Ldb3*, ex7 of *Nfix* and ex1 of *Mbnl1* (Fig. 2a). They were selected based on MBNL-specific CLIP-seq data deposited in MBNL Interactome Browser (MIB.amu.edu.pl)[14]. We confirmed their MBNL-dependent alternative splicing in cells upon *MBNL1,2* silencing (Fig. 2b). To date, functional MBNL-binding motifs have been well defined for three pre-mRNAs, *Atp2a1*, *Nfix* and *Mbnl1*[19,24,38]; however, these motifs remain unknown for the other selected transcripts. We established two criteria to emerge the most probable MBNL-binding motifs in pre-mRNA: (i) the presence of an MBNL-dependent alternative exon adjacent to putative binding motifs with CLIP-seq clusters either within the exon or upstream or downstream intron up to 250 nt from the exon; (ii) overlapping of at least three YGCY sequence motifs that are at least partially conserved in human and mouse.

For each selected transcript, the identified CLIP-seq clusters covered three or more YGCY motifs, and in some instances, their number reached up to 15. First, to specify their importance in MBNL1 binding, we generated ~100- to 200-nt-long intronic (*Pphln1*-, *NASP*-) and exonic (*Nfix*-, *Ldb3*-, *Mbnl1*-) RNA fragments and designed AONs that were complementary to these regions (*Mbnl1*-RNA described in[38]) (Supplementary Fig. S3a). In FBA, we observed that certain AONs, by blocking specific RNA fragments, significantly deteriorated the MBNL1 affinity, indicating one (within *Pphln1*-, *Atp2a1*-, *NASP*-RNAs) or two (within *Nfix*-, *Ldb3*-RNA) regions composed of a string of YGCY motifs as the most essential for MBNL1/RNA interactions (Fig. 2c and Supplementary Fig. S3b). In the case of *Nfix* and *Ldb3*, for further *in cellulo* studies, we selected AONs that bound to exonic sequences and showed the strongest regulatory potential.

To confirm the functionality of the selected YGCY motifs that are essential for binding, we introduced 2′OMe-PSs into cells at 125 nM. Subsequently, we used RT-PCR to analyze the alternative splicing pattern of targeted endogenous pre-mRNAs. AONs that target intronic sequences induced significant changes in alternative exon inclusion. The *Atp2a1*- and *Pphln1*-specific 2′OMe-PSs considerably suppressed ex22 and ex6 inclusion, respectively, whereas 2′OMe-PS, which is specific for *NASP*, significantly induced inclusion of ex7 (Fig. 2d). In those transcripts, the analyzed YGCY motifs were located in downstream (*Atp2a1* and *Pphln1*) or upstream (*NASP*) introns and induced MBNL-dependent alternative exon promotion or repression, respectively. In each case, the direction and strength of alternative splicing changes upon AON application were similar to the effect of siRNA-mediated *MBNL1,2* silencing (compare Fig. 2b and d).

We next sought to identify whether targeting MBNL-binding regions with specific 2′OMe-PS in MBNL-repressed exons of *Nfix*, *Ldb3* and *Mbnl1* would also trigger the same effect as siRNA against *MBNL1,2* and lead to their enhanced inclusion. Unexpectedly, we observed the opposite effect, and found that the inclusion of ex7, ex10 and ex1, respectively, was significantly repressed by specific AONs (Fig. 2d). Silencing of *MBNL1* and 2 and AON application induced the opposite effect on the inclusion of the analyzed exons. One possible explanation for this phenomenon is that AONs, due to their interaction within the exons of their targeted pre-mRNAs, may affect the binding of other splicing modulating factors, including splicing enhancers.

We can infer that DNA oligomers are an efficient tool to screen MBNL-binding regions *in vitro* because DNA and 2′OMe-PS AONs exert comparable inhibitory properties. On the other hand, only 2′OMe-PSs indicate functional intronic MBNL-binding regions in cells. However, the strength of the AON-induced change in alternative splicing strongly depends on the primordial distribution of the splicing isoforms and potency of MBNLs to regulate particular alternative exons[39]. Moreover, we encountered difficulties in analyzing all three MBNL-binding regions within exons where AONs induced exon skipping instead of the anticipated exon inclusion.

It remains unclear what mechanism drives alternative exon regulation by MBNL proteins bound to individual RNA *cis*-elements. Two previous works proposed involvement of MBNLs in the spliceosome formation *via* interacting with U2 auxiliary factor 65 kDa subunit (U2AF65) responsible for 3′ splice site definition. Depending on the binding site localization within introns, either downstream or upstream of alternative exon, MBNLs may play a role as an enhancer for U2AF65 binding[40] or its competitor by modifying the RNA secondary structure[41], respectively. The other possibility is that MBNLs associated with RNA affect binding of other *trans*-factors such as CELF1[42] and RBFOX[43].

Analogously, the binding of MBNLs to intronic *cis*-elements tested in this study could affect U2AF65 recruitment. Therefore, AONs blocking the RNA/MBNL interaction may contribute indirectly to observed alternative splicing changes. On the other hand, binding of AON itself could directly disrupt spliceosome organization *via* RNA structure rearrangements. However, due to a relatively long distance between MBNL-binding sites and alternative exons (~110–270 nt) it seems to be less possible. Blocking intronic or exonic MBNL-binding sites by AONs might also interfere with binding of other *trans*-factors pivotal for regulation of the alternative exon inclusion. Exon skipping induced by AON targeting either the 5′ and 3′ splice sites and/or exonic splicing enhancers (ESEs) has often been reported in studies on splicing manipulation in genetic, aberrant splicing-based disorders, including Duchenne Muscular Dystrophy (DMD)[44]. Deutekom and others showed that utilizing a set of AONs that cover many regions along with an exonic sequence leads to exon exclusion[45].

Highly redundant hybrid *Atp2a1* minigenes confirm the functional MBNL-binding motifs in intronic and exonic sequences. Previous observations led us to design hybrid minigenes as an alternative and complementary tool to test the sensitivity of RNA regulatory elements to MBNLs. To prepare the hybrid minigenes, we removed a natural 111-bp-long MBNL-binding cassette, which contains regions #1 and #2 with YGCY motifs, from intron 22 of the *Atp2a1WT* minigene (*Atp2a1Δ*). Removal of this cassette resulted in a strong reduction of the ex22 splicing response to MBNL1 overexpression (Fig. 3a). A similar effect was observed when we replaced the MBNL-binding cassette with a sequence that lacked YGCY motifs (*Atp2a1Δ-Ctrl*) (Fig. 3a). Subsequently, the studied RNA sequences from *Pphln1*, *NASP*, *Nfix, Ldb3* and *Mbnl1* transcripts, which

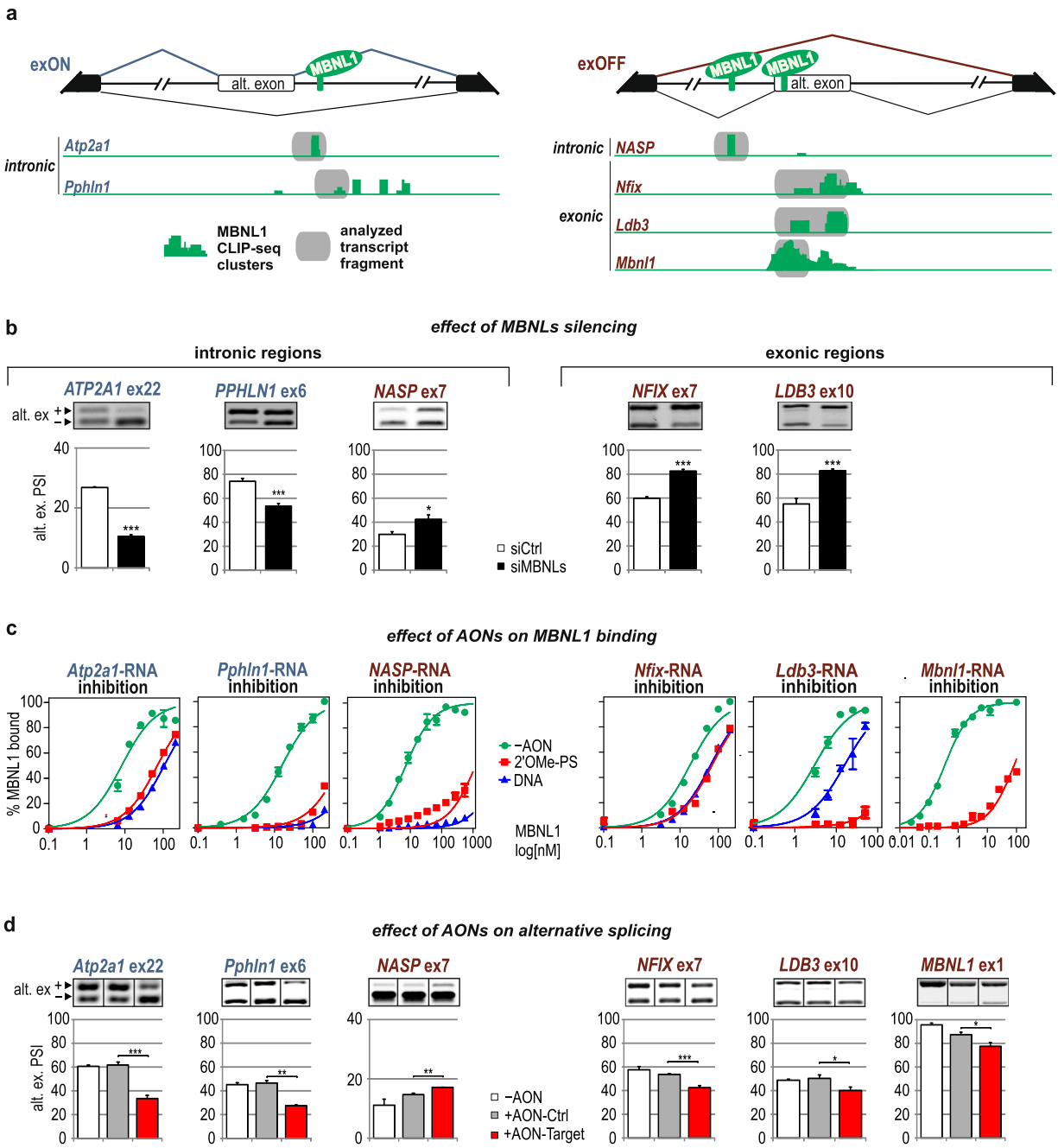


Figure 2. AONs indicate functional MBNL-binding regions within introns but not exons. (**a**) Schematic representation of the alternative splicing patterns of MBNL-dependent exons from *Atp2a1*, *Pphln1*, *NASP*, *Nfix*, *Ldb3* and *Mbnl1* pre-mRNAs, which depend on the localization of the MBNL-binding regions. MBNL1-specific CLIP-seq clusters are shown as green areas for each transcript. The splicing pattern for alternative exons spliced in (exON) upon MBNL binding within a downstream intron is depicted in dark blue, whereas the splicing pattern for alternative exons spliced out (exOFF) upon MBNL binding within an upstream intron or exon is depicted in burgundy. Constitutive and alternative exons are shown as black and white boxes, respectively. Intronic and exonic RNA fragments containing MBNL-binding regions targeted by AONs are marked with grey boxes. (**b**) RT-PCR analysis showing the percentage of alternative exon inclusion in the tested mRNAs in human cells after silencing of *MBNLs* and in cells treated with control siRNA (siCtrl). The results for *MBNL1* are presented in[38]; n = 3. (**c**) *In vitro* MBNL1 binding to RNA fragments derived from the studied transcripts based on FBA assays showing the blocking properties of 2′OMe-PS-AONs and DNA-AONs marked as red and blue curves, respectively; n = 2. (**d**) RT-PCR analysis showing changes in alternative splicing upon treatment with gene specific AON (AON-Target) or control AON (AON-Ctrl) in human or mouse cell lines; n = 3.

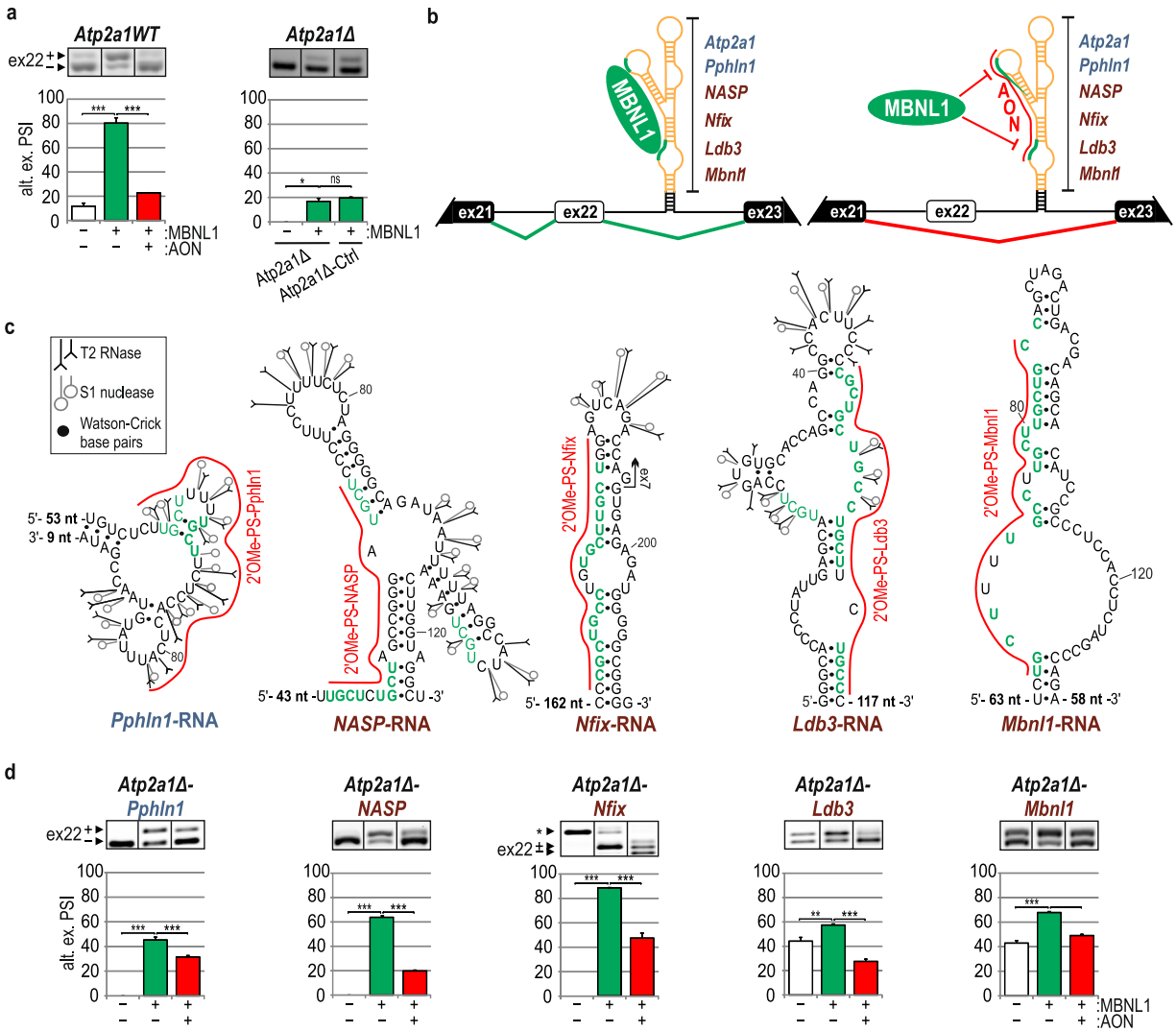


Figure 3. Hybrid *Atp2a1* minigenes confirm the functional intronic and exonic MBNL-binding regions. (**a**) Percentage of alternative ex22 inclusion in *Atp2a1WT* and *Atp2a1*Δ mRNA, lacking a fragment containing MBNL-binding regions, upon MBNL1 overexpression with or without 2′OMe-PS (AON) treatment; n = 3. *Atp2a1*Δ*-Ctrl*, a hybrid minigene with an inserted control sequence cassette lacking MBNL-binding regions; n = 2. (**b**) The schemes illustrate a general organization of hybrid *Atp2a1* minigenes containing an MBNL-responsive cassette derived from analyzed transcripts in the place of a natural *Atp2a1* cassette within intron 22. We expect to observe the promotion of ex22 through MBNLs binding to YGCY sequence motifs within inserted cassettes (green line) and ex22 exclusion upon AONs blocking MBNL-binding regions (red line). A thermodynamically stable structure having 14-nt-long complementary sequences derived from intron 22, restriction sites and 5-bp artificial helix, is marked with black. It is distant by 32–164 nucleotides from the MBNL-binding sites. (**c**) Experimentally determined secondary structures of selected areas of analyzed short intronic and exonic RNA fragments containing significant MBNL-binding regions complementary to specific AONs and marked with a red line. The entire structures of the analyzed *in vitro* RNA fragments are presented in Supplementary Fig. S4b. The number of not shown nucleotides is depicted on the 5′ and/or 3′ end of each structure. YGCY motifs are marked with green and in bold if present in humans and mice. *Mbnl1*-RNA was previously described[38]. (**d**) Percentage of alternative ex22 inclusion in the mRNA of a set of hybrid *Atp2a1* minigenes upon MBNL1 overexpression with or without treatment with specific 2′OMe-PS AONs. The asterisk indicates an artificial splicing isoform of the *Atp2a1*Δ*-Nfix* minigene; n = 3.

contained the analyzed MBNL-binding motifs, were inserted into the *Atp2a1*Δ minigene at the site of the deleted MBNL-sensitive cassette (Fig. 3b and Supplementary Fig. S3a). The sequences were cloned in a way to preserve their autonomous RNA secondary structures and reduce the potential effect of neighboring sequences on the generated RNA structure. For that purpose, the acceptor site in *Atp2a1* minigene, which constituted a naturally stable helical region derived from intron 22 within pre-mRNA, was slightly extended by a short artificial helix which was designed in a way to not affect the secondary structure of the MBNL-binding sites (Fig. 3b). To better understand the effect of the RNA structure of the MBNL-binding regions on MBNLs activity, we

experimentally determined the secondary structures of short intronic (*Pphln1*-, *NASP*-) and exonic (*Nfix*-, *Ldb3*-, *Mbnl1*-) RNA fragments representing the sequences inserted into the hybrid minigene, as described above for *Atp2a1*-RNA (Fig. 3c and Supplementary Fig. S4). RNA structure probing experiments showed that the significant MBNL-binding motifs, depicted in *in vitro* assays upon AON application, were predominantly located within semi-stable hairpin structures, either as strings of multiple YGCY motifs placed on one side of the hairpin structure with several unpaired nucleotide residues (*Ldb3*-RNA, *Mbnl1*-RNA) or within partially single-stranded regions (*Pphln1*-RNA, *NASP*-RNA). Only in *Nfix*-RNA the predicted MBNL-binding motifs were located within a stable hairpin structure with two symmetrical internal loops that were not recognized by any of the enzymatic structure probes.

Next, we co-transfected HeLa cells with one of five obtained minigenes, *Atp2a1Δ-Pphln1, Atp2a1Δ-NASP, Atp2a1Δ-Nfix, Atp2a1Δ-Ldb3* and *Atp2a1Δ-Mbnl1*, and the MBNL1 overexpressing vector. As anticipated, we observed that for all of the precursor RNAs synthesized from hybrid minigenes, inclusion of ex22 was promoted by MBNL1, similar to that for the original *Atp2a1WT* pre-mRNA (Fig. 3a,d; green bars). For two minigenes (*Atp2a1Δ-Mbnl1* and *Atp2a1Δ-Ldb3*), the endogenous MBNLs already strongly elevated ex22 inclusion compared with *Atp2a1WT* (Fig. 3d; white bars). For cells transfected with *Atp2a1Δ-Nfix*, we observed an additional artificial splicing isoform due to the presence of the remaining 5′ splice site within the cloned *Nfix* exon 7 fragment (Fig. 3d, Supplementary Fig. S4c). However, MBNL1 overexpression revealed strong inclusion of the *Atp2a1* exon 22-containing isoform (Fig. 3d).

In cells that were additionally transfected with AONs that targeted specific MBNL-binding regions, we observed a reduction of the ex22 inclusion, consistent with the effect triggered by 2′OMe-PS in *Atp2a1WT* pre-mRNA (Fig. 3a,d; red bars). In most cases, upon AON usage, the decreased level of ex22 inclusion did not reach the level observed in *Atp2a1Δ* or *Atp2a1Δ-Ctrl* mRNAs (Fig. 3a,d). We suppose that this effect was caused by blocking only a single MBNL-binding region with one AON, whereas the remaining motifs could still recruit MBNLs, as we observed in *in vitro* assays, and had a slight impact on ex22 inclusion.

Currently, RBP/RNA interaction studies concerning alternative splicing regulators employ a minigene preparation established on a transcript of interest containing an RBP-regulated alternative exon and RBP-binding region with its mutated variants[19,46]. Additionally, the inconvenience of the distribution of splicing isoforms is a limiting parameter. The length of the intronic sequences might also make preparations difficult[47]. To verify the high-throughput data, several of such minigenes would need to be provided, extending the time of study. Moreover, AON application is limited to intronic regulatory elements because targeting exons that are abundant in *cis*-acting regulatory sites leads to unexpected splicing misregulation.

Here, we present a new approach based on a hybrid minigene composed of an unchangeable core with an alternative exon regulated by MBNLs and exchangeable analyzed RNA regions containing potential MBNL-binding motifs combined with specific AON treatment. Of note, while we decided to use AONs, mutagenesis of potential MBNL-binding sites in RNA could be applied instead. AON application provides us with rapid and reproducible results for many intronic- and exonic-derived regulatory regions as well as the ability to evaluate their regulatory properties. Despite unexpected complications with AON usage upon potential MBNL-binding sites verification in endogenous exonic regions we decided to proceed with experiments with AONs on hybrid minigenes background as a consequence of AON application at each step of the study from *in vitro* to *in cellulo*.

Our observations clearly show that regulatory regions transferred from the original context into a different transcript may still act as *cis*-regulatory elements and regulate alternative splicing according to general principles. Our observations also indicate that MBNLs are sensitive to each of the analyzed intronic or exonic RNA fragments. Moreover, these experiments revealed that a sequence that was primarily located within an exon and acted as an ESE, when it was moved into an intronic context, became an MBNL-dependent intronic splicing enhancer (ISE). This result sheds light on the question regarding the actual mechanism of MBNL-dependent exon repression. We hypothesize that MBNLs, when bound to exonic *cis*-elements, can affect the binding or activity of other *trans*-factors that act as alternative splicing enhancers. This may explain the perplexing splicing effects that we observed when testing AONs that targeted the exonic sequences of *Nfix*, *Ldb3* and *Mbnl1* pre-mRNAs (Fig. 2d). We showed that the hybrid minigene-based approach is a useful tool to analyze MBNL-binding regulatory elements and, with great probability, that it can be successfully applied to test the functionality of other RBP-bound *cis*-elements that are naturally located in any parts of precursor or mature forms of different RNA classes. Finally, the developed hybrid minigene assay, by providing an indiscrete background, may be applicable to compare the quality of different sequences that contain MBNL-binding regions regarding the number and localization of YGCY motifs.

Potential crosstalk between SRSF1 and MBNLs in the regulation of alternative splicing. Blocking the exonic regions of interest with complementary AONs may lead to the unintended interruption of the general splicing machinery activity but also other *trans*-acting factors, including a large group of splicing regulators, e.g. SRSFs or hnRNPs. We decided to partially elaborate this issue. First, we attempted to find potential splicing factors that were able to bind near the MBNL-binding region using *in silico* prediction software, including *ESE finder* and *SF Map*[48,49]. Based on these algorithms and published data[50–53], we anticipated that the SRSF1 protein would be a putative regulatory factor, which binds close to the exonic MBNL-binding regions tested in this study and be impeded by the examined AONs (Fig. 4a). Using an *in vitro* electrophoretic mobility shift assay (EMSA), we demonstrated that recombinant SRSF1 bound to the studied exonic *Mbnl1*-RNA and *Nfix*-RNA with high affinity, ranging from 13 to 60 nM, respectively (Fig. 4b), compared with negative control RNAs $(CAG)_{20}$ and $(AGG)_{20}$ (Supplementary Fig. S5). The use of 2′OMe-PS-Mbnl1 and 2′OMe-PS-Nfix in *in vitro* tests significantly deteriorated the affinity by 2 to 3 fold (Fig. 4b).

To further investigate the potential mechanism of competition between SRSF1 and MBNLs *in cellulo*, we tested the alternative splicing of ex1 of the *Mbnl1* fragment within the artificial *Atp2a1* minigene pre-mRNA (*Mbnl1* minigene). Overexpression of MBNL1 led to the exclusion of ex1 (Fig. 4c), while titration of SRSF1

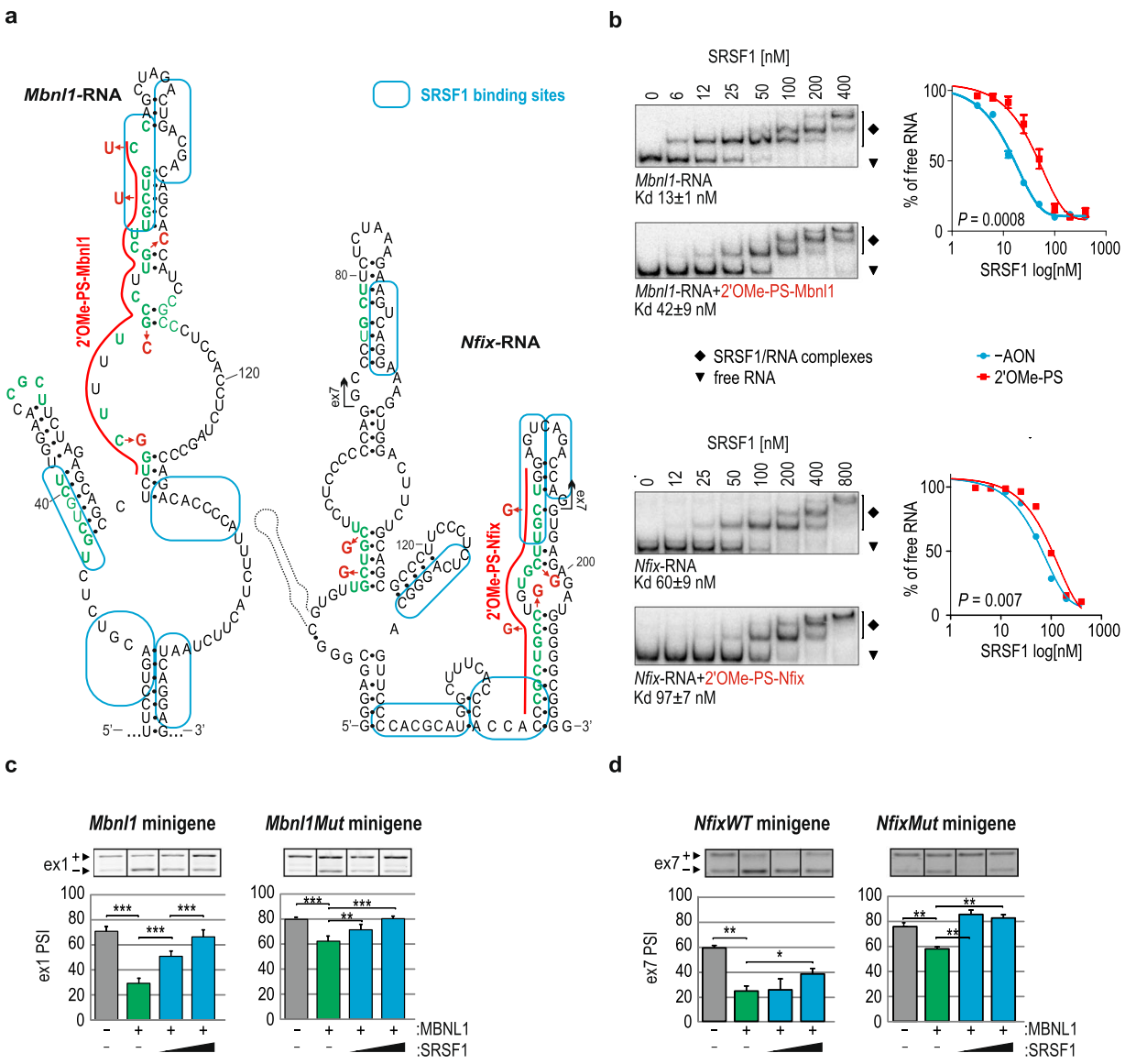


Figure 4. Potential crosstalk between SRSF1 and MBNLs. (**a**) Experimentally determined secondary structures of *Mbnl1*-RNA and *Nfix*-RNA with *in silico*-predicted SRSF1 binding regions marked with blue boxes[48,49]. 2′OMe-PS AONs blocking MBNL-binding regions are marked with a red line. Point mutations are marked with red. (**b**) Results of EMSA showing the interaction between SRSF1 at the indicated concentrations and 5′-^{32}P-labeled *Mbnl1*-RNA (upper panel) or *Nfix*-RNA (lower panel) with either intact MBNL-binding regions or the regions blocked by individual AONs. The calculated Kd values are indicated below each electrophoretogram; n = 4. On the right is the quantification of EMSA results based on the decline of the free RNA signal in favor of forming SRSF1/RNA complexes. Antagonistic role of SRSF1 and MBNL1 in the alternative splicing regulation of (**c**) ex1 of *Mbnl1* and (**d**) ex7 of *Nfix*. Note the SRSF1 dose-dependent promotion of both exon inclusion into the mRNA of wild-type *Mbnl1* or *Nfix* (left panels) or *Mbnl1Mut* or *NfixMut* minigenes (right panels). The amount of MBNL1 or SRSF1 expressing vector constituted 500 ng or ranged from 250 to 1000 ng per well, respectively; n = 2 (*Mbnl1*), n = 3 (*Nfix*).

in the MBNL1 background progressively promoted ex1 inclusion (Fig. 4c). The results for the pre-mRNA of *Mbnl1Mut* minigene with MBNL-binding motif mutations within ex1 showed only slightly distorted activity of SRSF1 (Fig. 4c). We observed similar but weaker effects of SRSF1 activity on the alternative ex7 inclusion level in mRNA of the *NfixWT* minigene and its variant with a mutated exonic YGCY motif - *NfixMut* (Fig. 4d).

These results suggest that SRSF1 and MBNLs are antagonistic splicing regulators of ex1 of *Mbnl1* and ex7 of *Nfix* pre-mRNAs. SRSF1 very likely interacts in the vicinity of but not within MBNL-binding motifs because mutations of YGCYs do not strongly affect its activity. Potentially, binding of the tested AONs may block SRSF1 binding sterically or *via* RNA structure alterations, resulting in repression of the targeted exons. Additional studies are required to elucidate the precise mechanism and interplay between these two splicing regulators.

To date, only a few mechanisms of MBNL-mediated alternative exon regulation that implicate other *trans*-acting factors have been proposed[40–42]. Although none of them is able to explain the MBNL-mediated regulation of alternative splicing through interactions within an exonic region, we can deduce that the mode of action of MBNLs involves direct competition of two proteins for the same binding region, as described for U2AF65[41], or indirect competition by the modulation of the RNA secondary structure and/or assistance of mediatory proteins, as shown for U2AF65[40] and CELF1[42]. Our initial results seem to support the latter mode of action; however, this finding needs to be clarified with more data.

The hybrid *Atp2a1* minigene is a useful tool to characterize inhibitors of adverse MBNL/RNA interactions. To emphasize the versatility of the *Atp2a1* hybrid minigene, we determined its usefulness for screening or characterizing the potency of different reagents that distort pathogenic protein/RNA interactions in myotonic dystrophy type 1 (DM1) or Fuchs endothelial corneal dystrophy.

DM1 is a neuromuscular, autosomal dominant disease that is caused by the expansion of three-nucleotide *CTG* repeats located in the 3′-untranslated region of the myotonic dystrophy protein kinase gene (*DMPK*)[54–56]. Expression of mRNA with *CUG* expansion ($(CUG)^{exp}$) leads to mRNA retention in the nucleus as well as pathogenic recruitment and sequestration of nuclear proteins, including MBNLs, to an RNA hairpin-like structure formed by the repeats[55]. The longer the repeats, the more severe the missplicing of hundreds of genes due to the functional insufficiency of MBNLs[57]. On the other hand, in Fuchs endothelial corneal dystrophy, toxic $(CUG)^{exp}$ RNA repeats occur within intron 2 of a transcription factor 4 pre-mRNA (*TCF4*)[13,58].

Among the therapeutic strategies aimed at ameliorating the toxic effect of $(CUG)^{exp}$, antisense oligonucleotides and small compounds have been frequently employed to specifically bind to repeats and prevent MBNL sequestration[59–62]. We have previously shown that an 8-nt-long oligonucleotide composed of LNA-PS moieties (LNA-PS-CAG-8)[63] and erythromycin (EM)[64] prevent MBNL sequestration on this toxic RNA by complementarily hybridizing to or potentially binding to U/U mismatches within the structure formed by $(CUG)^{exp}$, respectively. These molecules diminish adverse molecular outcomes, mainly abnormalities in alternative splicing in the cell and mouse models of DM1.

To investigate the usefulness of the hybrid minigene in the screening of potential therapeutics that target *CUG* repeats, we substituted a natural regulatory cassette of *Atp2a1* with an MBNL-binding region composed of 17 *CUG* repeats ($(CUG)_{17}$ (*Atp2a1Δ-*$(CUG)_{17}$) (Fig. 5a). Our previous experiments showed that this type of sequence forms a stable RNA hairpin structure that binds all MBNLs with high affinity[14]. In this study, MBNL1 overexpression induced a strong splicing response of ex22 in the minigene's pre-mRNA (Supplementary Fig. S6a). Additionally, we assumed that more than one MBNL1 molecule binds to the pre-mRNA of *Atp2a1Δ-*$(CUG)_{17}$. In support of our hypothesis, hybrid minigenes expressing the pre-mRNAs containing $(CUG)_4$ or $(CUG)_8$ undergo MBNL-mediated splicing with lower rate of ex22 inclusion compared to *Atp2a1Δ-*$(CUG)_{17}$ (Supplementary Fig. S6b).

Next, we chose LNA-PS-CAG-8 and EM as examples of efficient inhibitors of the MBNL/*CUG* interaction and used them in our minigene assay. Cells co-transfected with the *Atp2a1Δ-*$(CUG)_{17}$ minigene and MBNL1 overexpressing vector were treated with different concentrations of either LNA-PS-CAG-8 or EM. We observed a strong and gradual decrease in ex22 inclusion induced by LNA-PS-CAG-8 but not by the control LNA oligonucleotide, reaching below the level established by endogenous MBNLs at a concentration of 125 nM (Fig. 5b). The EM showed a weaker effect on ex22 inclusion over a broad range of this antibiotic concentration (Fig. 5c). This result was consistent with our previous findings that showed that these two compounds have different potencies as inhibitors of the MBNL/$(CUG)^{exp}$ interaction in both *in vitro* and *in vivo* models[63,64].

These results show that the *Atp2a1Δ-*$(CUG)_{17}$ hybrid minigene is a very sensitive tool for characterizing different MBNL/$(CUG)^{exp}$ inhibitors and that after certain modifications, such as reporter gene inclusion, the hybrid minigene can be successfully applied in high-throughput screening. A related screening system, based on the detection of alternative splicing pattern changes within the *Clcn1* minigene's pre-mRNA and including luciferase system for high-throughput analysis, has been described previously[65]. The advantage of our *Atp2a1Δ-*$(CUG)_{17}$ minigene is its sensitivity, reproducibility and simplicity because it can be introduced into non-DM1 cell models. Moreover, the assay provides not only qualitative but also quantitative information about the potency of the studied compounds and additionally enables to monitor the differences in the number of MBNLs associated with *CUG* repeats. Investigation of potential therapeutic agents for diseases mediated by toxic RNA that disrupts the activity of RBPs, including myotonic dystrophies, is an ongoing process. Our assay could facilitate the screening of molecules on *in vitro* stages.

Methods

Oligonucleotides. All of the DNA sequences were purchased from oligo.pl® (Institute of Biochemistry and Biophysics PAN, Warsaw, Poland) and are listed in Supplementary Tables. All RNA-based AONs were synthesized by RiboTask™ and are listed in Table 1.

Plasmid preparation. To obtain the *Atp2a1WT* minigene, a fragment of the *Atp2a1* gene between exons 21 and 23 containing MBNL-sensitive alternative exon 22 was amplified from mouse genomic DNA using Platinum® high fidelity *Taq* DNA polymerase (Invitrogen™) according to the manufacturer's protocol with a SrcMres_F/SrcMres_R primer set (Ta = 55 °C) whose sequences are provided in Supplementary Table S1. These primers introduced unique restriction sites for *Xho*I and *BamH*I. To obtain the *Atp2a1Δ* minigene, a 111-bp sequence within intron 22 in the *Atp2a1WT* minigene that included the MBNL-binding regions was substituted with a 34-bp linker sequence containing unique restriction sites for *Not*I and *Sal*I. First, two fragments of the final *Atp2a1Δ* insert were amplified using the pEGFP_F/NotSal_R and NotSal_F/pEGFP_R primer sets (Ta = 60 °C) to introduce the 34-bp linker in both PCR products. Second, based on their complementarity to 34-nt linker, the

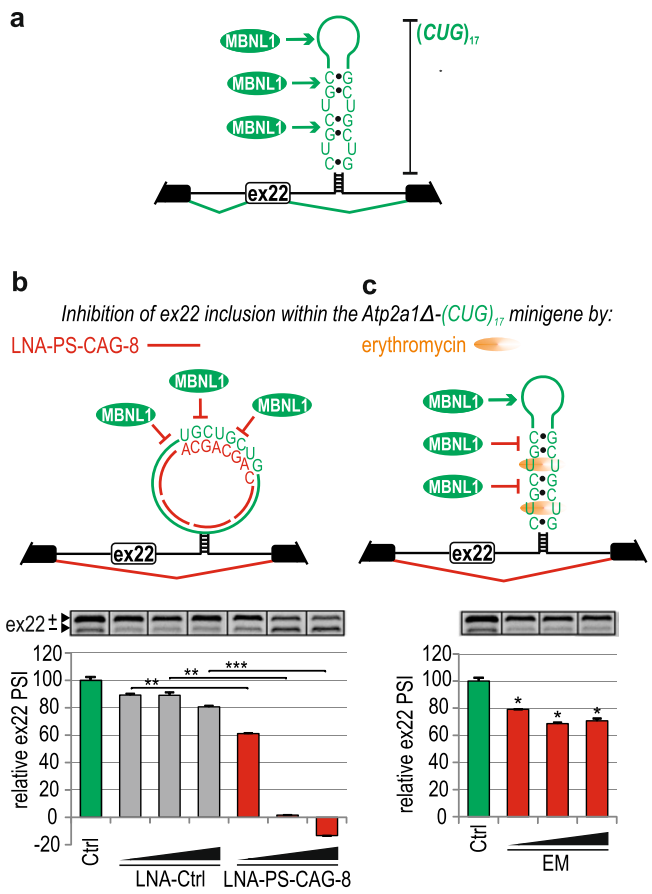


Figure 5. The hybrid minigene is useful to screen potential RNA-binding therapeutics. (**a**) A scheme illustrating MBNL-dependent ex22 inclusion to mRNA derived from the *Atp2a1Δ-(CUG)*$_{17}$ minigene by extensive MBNL1 binding to the hairpin structure formed by *(CUG)*$_{17}$. Constitutive and alternative exons are shown as black and white boxes, respectively. (**b**–**c**) *Top*; Schemes illustrating the inhibition of MBNL-dependent ex22 inclusion to mRNA derived from the *Atp2a1Δ-(CUG)*$_{17}$ minigene by antisense oligonucleotide LNA-PS-CAG-8 (left panel) and erythromycin (EM) by their binding to the hairpin structure formed by *(CUG)*$_{17}$ repeats. *Bottom*; reduction of alternative ex22 inclusion to *Atp2a1Δ-(CUG)*$_{17}$ mRNA in cells treated with (**b**) LNA-PS-CAG-8 or (**c**) EM. The results are normalized to the splicing response of ex22 in *Atp2a1Δ-(CUG)*$_{17}$ pre-mRNA without (0%) or with (100%) MBNL1 overexpression (Ctrl, green bar). The LNA-PS-CAG-8 and EM were used in a range of 25–125 nM or 50–500 μM concentrations, respectively; n = 2.

obtained PCR fragments constituted the template for PCR with the pEGFP_F/pEGFP_R primer set (Ta = 60 °C) to provide a full-length *Atp2a1Δ* sequence. To obtain *Atp2a1*-hybrid minigenes, the sequences of interest from the *Pphln1*, *NASP*, *Nfix*, *Ldb3*, *Mbnl1 and Ctrl (Capzb)* genes were amplified from human (*NASP*) or mouse (others) genomic DNA with the Pph1insF/Pph1insR, NASPinsF/NASPinsR, NfixinsF/NfixinsR, Ldb3insF/Ldb3insR, Mbnl1insF/Mbnl1insR, CTRLinsF/CTRLinsR primer sets (Supplementary Table S1), respectively. The primers introduce restriction sites for *Not*I or *Sal*I. They also contained 5 additional nucleotides at the 5′ ends to introduce a hairpin structure in transcribed pre-mRNAs to stabilize the secondary structure of the fragments of interest (except for *-Mbnl1* construct). To obtain the *Atp2a1-(CUG)*$_{17}$ minigene, oligodeoxynucleotides CTG17F and CTG17R with restriction sites for either *Not*I or *Sal*I, were used. Prior to their hybridization for one hour at RT, they were purified on a denaturing 10% polyacrylamide (PA) gel, followed by precipitation and denaturation at 95 °C for 5 min. All of the amplified DNA fragments of interest were ligated into the pEGFP-C1 vector (CloneTech) prior to transformation into DH5α bacteria cells. To obtain the *Mbnl1* minigene, a DNA sequence including the 3′ part of the intron (53 bp) and 5′ part of exon 1 (110 bp) of the *Mbnl1* gene was amplified from human genomic DNA using the Mbnl1F/Mbnl1R primer set (Ta = 55 °C), introducing restriction sites for *Not*I or *Sal*I (Supplementary Table S1). A *Mbnl1Mut* minigene was prepared in two rounds of PCR. First, two products were amplified with primers that introduced point mutations within the YGCY motifs (Mbnl1mtR/Mbnl1F as well as Mbnl1mtF/Mbnl1R; Supplementary Table S1) (Ta = 55 °C) and were purified on an agarose gel. Second, the purified PCR products were annealed and extended in 4 cycles of 92 °C for 2 min, 55 °C for 2 min, 72 °C for 2 min, and then, an Mbnl1F/Mbnl1R primer set that induced the restriction sites for *Not*I or *Sal*I was added and the reaction proceeded according to the manufacturer's protocol at Ta = 55 °C. The obtained final PCR products, after digestion with *Not*I or *Sal*I, were ligated into the *Atp2a1Δ* minigene and were transformed into DH5α bacteria cells. The pEGFP-C1-MBNL1-41 vector for MBNL1 overexpression was previously described[38]. pEGFP-SF2 (pEGFP-C1-SRSF1) was a gift from Dr. Tom Misteli (Addgene plasmid # 17990)[66].

AON name	Sequence 5′ → 3′[a]
for *Atp2a1* transcript	
LNA#1	CAGCAGCAGC
LNA#2	GUGGAAGCGG
LNA-PS#1	CAGCAGCAGC
LNA-PS#2	AAGCGGGCA
2′OMe	AGCGGGCAGUGGCAACAGCAG
2′OMe-PS	GCGGGCAGUGGCAACAGCAGC
for other transcripts	
2′OMe-PS-Pphln1	GGAGAAGCAAAAAGCAAGAGA
2′OMe-PS-NASP	AGCAUCCCGGCUAGCAGAGCA
2′OMe-PS-Nfix	CCAGCAAGCACAGGCAGCGGG
2′OMe-PS-Ldb3	CAGAAGCAGGCAGCAGCGGGG
2′OMe-PS-Mbnl1	GCAGCAAGCAAGGCAAAAGCA
LNA-PS-CAG-8	CAGCAGCA
LNA-Ctrl	ACGCACAAGG

Table 1. RNA-based AON sequences. [a]In each sequence, the 3′ end nucleotide was 2′OMe.

Cell culture and transfection. Human HeLa and mouse C2C12 cells were grown in high-glucose DMEM medium (Lonza) supplemented with 10% fetal bovine serum (Sigma) and 1x antibiotic/antimycotic (Sigma). Human skeletal myoblast (HSkM) cells were grown in HAM F-10 medium (Sigma) supplemented with 20% FBS, 1x antibiotic/antimycotic, 0.39 µg/ml dexamethasone (Sigma) and 10 ng/ml epidermal growth (Sigma). All cells were grown at 37 °C and in a 5% atmosphere of CO_2. HeLa, C2C12 or HSkM cells plated in 12-well plates were transfected at 50–60% confluence with Lipofectamine® 2000 (Invitrogen™) according to the manufacturer's protocol. Single transfection was conducted with an siRNA mix against *MBNL1* and *MBNL2* (25 nM each) (FUTURE synthesis and RiboTask™, respectively[67,68]), 50 nM AllStars negative control siRNA (Qiagen), or a specific AON at 125 nM, in which AON-Ctrl was 2′OMe-PS unspecific to a tested transcript. Co-transfection was conducted with 200 ng of the minigene and 500 ng (or as indicated in the figures) of the MBNL1, SRSF1 or GFP expressing vector. For verification of the MBNL-binding regions and inhibition of the MBNL/$(CUG)^{exp}$ interaction, the co-transfection was followed by a 4-hour incubation and transfection with selected AONs. Erythromycin was added directly to the medium. The cells were harvested 48 hours after transfection.

Splicing analyses of mature mRNA. Total RNA was isolated from cells using TRI-reagent (Sigma) according to the manufacturer's protocol. cDNA was synthesized using GoScript™ Reverse Transcriptase (Promega) with Random Primers (Promega) according to the manufacturer's protocol. RT-PCR was performed using GoTaq® Flexi DNA Polymerase (Promega) with the primer sets listed in Supplementary Table 1 at Ta = 55 °C. The PCR products were electrophoretically separated in 1–2% agarose gels containing 0.005% ethidium bromide, and images were captured using G:Box EF2 (Syngene). In the main figures, splicing analyses are represented by cropped sections of the original gels separated by black lines. All of the raw images of the RT-PCR analyses are shown in Supplementary Fig. 7.

Preparation of radiolabeled RNA fragments. The templates for the transcription reactions of RNA fragments were obtained in two rounds of PCR. In the first PCR, longer products were amplified with genomic mouse or human DNA templates using specific _F/_R primer sets (the primer sequences and annealing temperatures (Ta) are listed in Supplementary Table S2). The obtained PCR products constituted a template for a second round of PCR with specific _TF/_TR primer sets harboring a promoter for polymerase T7 (Supplementary Table S2) at Ta = 55 °C. Templates for *Atp2a1*-RNA mutants were prepared in two rounds of PCR similar to that for the *Mbnl1Mut* minigene. In the first step, specific Atp2a1mt_F/Atp2a1_TR and Atp2a1mt_R/Atp2a1_TF were used. In the second round of PCR, only _TF and _TR were used. The transcription reaction was performed in 50 µl containing 10 µl of DNA template (~0.5 µg), 0.15 mM of each ribonucleotide (Invitrogen™), 0.45 mM guanosine (Sigma-Aldrich), 50 U of T7 RNA Polymerase (Ambion), 1 × T7 transcription buffer (Ambion), and 40 U of RNasin®Plus RNase Inhibitor (Promega). The transcripts were purified on a denaturing 8% PA gel (19:1 acrylamide:bisacrylamide), followed by ethanol precipitation. For 5′-end radiolabeling, 2 pmol of RNA was incubated with 2 pmol (12 mCi) of [γ-P^{32}] ATP, 1 U of RNasin®Plus RNase Inhibitor (Promega) and 10 U of OPTI Kinase (Affymetrix) in 1 x reaction buffer (Affymetrix) at a total volume of 10 µl and at 37 °C for 45 min. The 5′-^{32}P-labeled RNAs were subsequently run on a denaturing 8% PA gel. The bands of RNAs were visualized *via* Storage Phosphor Screen BAS-IP (IP) and a laser scanner FLA-1500 (FujiFilm) and were then excised, followed by ethanol precipitation and resuspension in 20 µl of double-distilled water.

Chemical and enzymatic analysis of RNA structure. Analysis was conducted as previously described[38] with slight changes. Briefly, 5′-^{32}P-labeled RNA was first subjected to a denaturation and renaturation procedure in either buffer N (20 mM Tris-HCl pH 7.2, 80 mM NaCl, 2 mM $MgCl_2$) for ribonucleases T1 and T2 or in a commercial S1 buffer (Fermentas) for endonuclease S1. 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 and T2 at the concentrations indicated in

Supplementary Figs S2a and S4a. An equal volume of stop solution was added to stop all reactions. The products of cleavage were separated on a denaturing 10% PA gel along with the nucleotide ladders constituting products of alkaline hydrolysis (F), lead ion cleavage (Pb) and T1 nuclease digestion (T1L). The F ladder was generated by RNA cleavage in formamide at 99 °C for 10–15 min. The T1L ladder was generated by RNA cleavage in 50 mM sodium citrate pH 4.3 and 7 M urea by addition of 10 or 15 U/µl of T1 at 55 °C for 10 min. The Pb ladder was generated by RNA cleavage in buffer N by 0.25 mM of Pb^{2+} at 55 °C for 2.5 min. The bands of RNAs were visualized *via* IP and the laser scanner FLA-1500 after overnight (O/N) exposition. *In silico* prediction of the secondary structure was performed using the *mfold* program with experimentally defined constraints.

Recombinant proteins. The GST fusion construct with a truncated 41-kDa MBNL1 isoform sequence encoded by exons 1–4 of the MBNL1 gene with a His6 tag at the C-terminus was prepared as previously described[69].

To generate the construct for recombinant SRSF1, 558 bp of the SRSF1 coding sequence was amplified using the pEGFP-SF2 vector (Addgene) as a template and SRSF1_F/ SRSF1_R primer set (Supplementary Table S2) and was cloned into the pRSF-Duet vector using *BamH*I and *Not*I. Sequencing of the obtained construct was followed by transformation of the *E. coli* BL21(DE3)pLysS strain (Novagen). Production of 23-kDa SRSF1 containing a His6 tag at the C-terminus was performed as previously described[70].

Quantification of the MBNL1/RNA interaction and its inhibition by AONs *in vitro*. The filter binding assay (FBA) was performed to assess the affinity of MBNL1 to different RNAs. 5′-^{32}P-labeled RNAs (0.05 nM) were subjected to a denaturation and renaturation procedure by heating the samples at 95 °C for 1 min followed by a 10-min incubation at RT. Subsequently, the RNAs were incubated with the indicated protein concentrations (ranging from 0 to 200 nM) in buffer FB (250 mM NaCl, 15 mM KCl, 50 mM Tris-HCl pH 8.0, 0.05% Tween-20, 1 mM $MgCl_2$) at 37 °C for 15 min. Studying the inhibitory property of AONs, 0.05 nM of 5′-^{32}P-labeled RNAs mixed with 0.5 µM of RNA-based AONs or 20 µM of DNA-AONs (Supplementary Table S3) were subjected to the denaturation and renaturation procedure, followed by incubation at 37 °C for 15 min. Next, the indicated concentrations of MBNL1 (ranging from 0 to 200 nM) were added to each sample, followed by incubation at 37 °C for 15 min. Next, 25 µl of each sample was loaded onto the filter binding apparatus with nitrocellulose (Protran BA 85, Whatman®) and nylon (Hybond™ N+, Amersham) membranes that were previously wetted in buffer FB. The signal from the membranes was visualized *via* IP and laser scanner FLA-1500 after O/N exposition. The dissociation constant (Kd) of the MBNL1/RNA complexes treated or non-treated with different AONs was calculated in the GraphPad program using the following equation: one site specific binding curve (Y = Bmax*X/(Kd + X)).

Quantification of the SRSF1/RNA interaction and its inhibition by AONs *in vitro*. The electrophoretic mobility shift assay (EMSA) was carried out by incubation of 0.05 nM of 5′-^{32}P-labeled RNA fragments with the indicated SRSF1 concentrations (ranging from 0 to 800 nM). The reactions were performed in a volume of 10 µl in buffer A (100 mM NaCl, 20 mM HEPES-Na pH 8.0, 5% glycerol, 0.5 mM DTT, 0.2 mM EDTA, 50 µg/ml BSA (Invitrogen™)) at 37 °C for 20 min. To identify SRSF1-binding regions, 0.5 µM of 2′OMe-PS-Mbnl1 or -Nfix along with 0.05 nM of 5′-^{32}P-labeled RNAs were subjected to denaturation and renaturation, followed by incubation at 37 °C for 20 min in buffer A. The samples were mixed with loading buffer (30% glycerol, 0.25% of bromophenol blue and xylene cyanol) and run on a non-denaturing 8% PA gel. The gel was subsequently dried, and the signal was visualized *via* IP and the laser scanner FLA-1500 after O/N exposition. The Kd of the SRSF1/RNA interaction was calculated based on the signal of free RNA and using the equation: one phase decay curve equation (Y = (Y0 − Plateau)*exp(−K*X) + Plateau) in Graph Pad program.

Statistical analysis. Group data obtained *in vitro* and *in cellulo* are expressed as the means ± standard deviation (SD). The statistical significance of the RT-PCR and EMSA results was determined by two-tailed Student's t-test using Microsoft Excel (*$P < 0.05$; **$P < 0.01$ and ***$P < 0.001$). Statistical analysis was calculated using 2–4 biological repeats. Similar results were obtained from at least two independent experiments.

Data availability. All data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

References

1. Sanchez, S. E., Petrillo, E., Kornblihtt, A. R. & Yanovsky, M. J. Alternative splicing at the right time. *RNA Biol* **8**, 954–959, https://doi.org/10.4161/rna.8.6.17336 (2011).
2. Hallegger, M., Llorian, M. & Smith, C. W. Alternative splicing: global insights. *FEBS J* **277**, 856–866, https://doi.org/10.1111/j.1742-4658.2009.07521.x (2010).
3. Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457–463, https://doi.org/10.1038/nature08909 (2010).
4. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476, https://doi.org/10.1038/nature07509 (2008).
5. Pascual, M., Vicente, M., Monferrer, L. & Artero, R. The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation* **74**, 65–80, https://doi.org/10.1111/j.1432-0436.2006.00060.x (2006).
6. Jin, Y. *et al.* A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J* **22**, 905–912, https://doi.org/10.1093/emboj/cdg089 (2003).
7. Dasgupta, T. & Ladd, A. N. The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *Wiley Interdiscip Rev RNA* **3**, 104–121, https://doi.org/10.1002/wrna.107 (2012).
8. Twyffels, L., Gueydan, C. & Kruys, V. Shuttling SR proteins: more than splicing factors. *FEBS J* **278**, 3246–3255, https://doi.org/10.1111/j.1742-4658.2011.08274.x (2011).

9. Teplova, M. & Patel, D. J. Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat Struct Mol Biol* **15**, 1343–1351, https://doi.org/10.1038/nsmb.1519 (2008).
10. Cho, D. H. & Tapscott, S. J. Myotonic dystrophy: emerging mechanisms for DM1 and DM2. *Biochim Biophys Acta* **1772**, 195–204, https://doi.org/10.1016/j.bbadis.2006.05.013 (2007).
11. Fardaei, M. *et al*. 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 (2002).
12. Kino, Y. *et al*. Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum Mol Genet* **13**, 495–507, https://doi.org/10.1093/hmg/ddh056 (2004).
13. Du, J. *et al*. RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy. *J Biol Chem* **290**, 5979–5990, https://doi.org/10.1074/jbc.M114.621607 (2015).
14. Sznajder, L. J. *et al*. Mechanistic determinants of MBNL activity. *Nucleic Acids Res* **44**, 10326–10342, https://doi.org/10.1093/nar/gkw915 (2016).
15. Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol Cell* **13**, 91–100 (2004).
16. Warf, M. B. & Berglund, J. A. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA* **13**, 2238–2251, https://doi.org/10.1261/rna.610607 (2007).
17. Goers, E. S., Purcell, J., Voelker, R. B., Gates, D. P. & Berglund, J. A. MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res* **38**, 2467–2484, https://doi.org/10.1093/nar/gkp1209 (2010).
18. Lambert, N. *et al*. RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell* **54**, 887–900, https://doi.org/10.1016/j.molcel.2014.04.016 (2014).
19. Du, H. *et al*. Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat Struct Mol Biol* **17**, 187–193, https://doi.org/10.1038/nsmb.1720 (2010).
20. Zhang, C. *et al*. Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev* **22**, 2550–2563, https://doi.org/10.1101/gad.1703108 (2008).
21. Masuda, A. *et al*. CUGBP1 and MBNL1 preferentially bind to 3′ UTRs and facilitate mRNA decay. *Sci Rep* **2**, 209, https://doi.org/10.1038/srep00209 (2012).
22. Konieczny, P., Stepniak-Konieczna, E. & Sobczak, K. MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res* **42**, 10873–10887, https://doi.org/10.1093/nar/gku767 (2014).
23. Ho, T. H. *et al*. Muscleblind proteins regulate alternative splicing. *EMBO J* **23**, 3103–3112, https://doi.org/10.1038/sj.emboj.7600300 (2004).
24. Hino, S. *et al*. Molecular mechanisms responsible for aberrant splicing of SERCA1 in myotonic dystrophy type 1. *Hum Mol Genet* **16**, 2834–2843, https://doi.org/10.1093/hmg/ddm239 (2007).
25. Kino, Y. *et al*. MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res* **37**, 6477–6490, https://doi.org/10.1093/nar/gkp681 (2009).
26. Wang, E. T. *et al*. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150**, 710–724, https://doi.org/10.1016/j.cell.2012.06.041 (2012).
27. Ule, J., Jensen, K., Mele, A. & Darnell, R. B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**, 376–386, https://doi.org/10.1016/j.ymeth.2005.07.018 (2005).
28. Konig, J. *et al*. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* **17**, 909–915, https://doi.org/10.1038/nsmb.1838 (2010).
29. Hafner, M. *et al*. PAR-CliP–a method to identify transcriptome-wide the binding sites of RNA binding proteins. *J Vis Exp*, https://doi.org/10.3791/2034 (2010).
30. Ule, J. *et al*. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**, 1212–1215, https://doi.org/10.1126/science.1090095 (2003).
31. Jensen, K. B. & Darnell, R. B. CLIP: crosslinking and immunoprecipitation of *in vivo* RNA targets of RNA-binding proteins. *Methods Mol Biol* **488**, 85–98, https://doi.org/10.1007/978-1-60327-475-3_6 (2008).
32. Sobczak, K. & Krzyzosiak, W. J. RNA structure analysis assisted by capillary electrophoresis. *Nucleic Acids Res* **30**, e124 (2002).
33. Yuan, Y. *et al*. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res* **35**, 5474–5486, https://doi.org/10.1093/nar/gkm601 (2007).
34. Sierakowska, H., Sambade, M. J., Agrawal, S. & Kole, R. Repair of thalassemic human beta-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc Natl Acad Sci USA* **93**, 12840–12844 (1996).
35. Friedman, K. J. *et al*. Correction of aberrant splicing of the cystic fibrosis transmembrane conductance regulator (CFTR) gene by antisense oligonucleotides. *J Biol Chem* **274**, 36193–36199 (1999).
36. Dunckley, M. G., Manoharan, M., Villiet, P., Eperon, I. C. & Dickson, G. Modification of splicing in the dystrophin gene in cultured Mdx muscle cells by antisense oligoribonucleotides. *Hum Mol Genet* **7**, 1083–1090 (1998).
37. Bonifert, T. *et al*. Antisense Oligonucleotide Mediated Splice Correction of a Deep Intronic Mutation in OPA1. *Mol Ther Nucleic Acids* **5**, e390, https://doi.org/10.1038/mtna.2016.93 (2016).
38. Konieczny, P., Stepniak-Konieczna, E., Taylor, K., Sznajder, L. J. & Sobczak, K. Autoregulation of MBNL1 function by exon 1 exclusion from MBNL1 transcript. *Nucleic Acids Res*, https://doi.org/10.1093/nar/gkw1158 (2016).
39. Wagner, S. D. *et al*. Dose-Dependent Regulation of Alternative Splicing by MBNL Proteins Reveals Biomarkers for Myotonic Dystrophy. *PLoS Genet* **12**, e1006316, https://doi.org/10.1371/journal.pgen.1006316 (2016).
40. Echeverria, G. V. & Cooper, T. A. Muscleblind-like 1 activates insulin receptor exon 11 inclusion by enhancing U2AF65 binding and splicing of the upstream intron. *Nucleic Acids Res* **42**, 1893–1903, https://doi.org/10.1093/nar/gkt1020 (2014).
41. Warf, M. B., Diegel, J. V., von Hippel, P. H. & Berglund, J. A. The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad Sci USA* **106**, 9203–9208, https://doi.org/10.1073/pnas.0900342106 (2009).
42. Wang, E. T. *et al*. Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res* **25**, 858–871, https://doi.org/10.1101/gr.184390.114 (2015).
43. Klinck, R. *et al*. RBFOX1 cooperates with MBNL1 to control splicing in muscle, including events altered in myotonic dystrophy type 1. *PLoS One* **9**, e107324, https://doi.org/10.1371/journal.pone.0107324 (2014).
44. Aartsma-Rus, A. & van Ommen, G. J. Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications. *RNA* **13**, 1609–1624, https://doi.org/10.1261/rna.653607 (2007).
45. van Deutekom, J. C. *et al*. Antisense-induced exon skipping restores dystrophin expression in DMD patient derived muscle cells. *Hum Mol Genet* **10**, 1547–1554 (2001).
46. Dominski, Z. & Kole, R. Selection of splice sites in pre-mRNAs with short internal exons. *Mol Cell Biol* **11**, 6075–6083 (1991).
47. Kosaki, A., Nelson, J. & Webster, N. J. Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA. *J Biol Chem* **273**, 10331–10337 (1998).
48. Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q. & Krainer, A. R. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* **31**, 3568–3571 (2003).
49. Paz, I., Akerman, M., Dror, I., Kosti, I. & Mandel-Gutfreund, Y. SFmap: a web server for motif analysis and prediction of splicing factor binding sites. *Nucleic Acids Res* **38**, W281–285, https://doi.org/10.1093/nar/gkq444 (2010).
50. Dirksen, W. P., Li, X., Mayeda, A., Krainer, A. R. & Rottman, F. M. Mapping the SF2/ASF binding sites in the bovine growth hormone exonic splicing enhancer. *J Biol Chem* **275**, 29170–29177, https://doi.org/10.1074/jbc.M001126200 (2000).

51. Anczukow, O. *et al*. SRSF1-Regulated Alternative Splicing in Breast Cancer. *Mol Cell* **60**, 105–117, https://doi.org/10.1016/j.molcel.2015.09.005 (2015).
52. Sen, S. *et al*. Muscleblind-like 1 (Mbnl1) promotes insulin receptor exon 11 inclusion via binding to a downstream evolutionarily conserved intronic enhancer. *J Biol Chem* **285**, 25426–25437, https://doi.org/10.1074/jbc.M109.095224 (2010).
53. Sen, S., Talukdar, I. & Webster, N. J. SRp20 and CUG-BP1 modulate insulin receptor exon 11 alternative splicing. *Molecular and cellular biology* **29**, 871–880, https://doi.org/10.1128/MCB.01709-08 (2009).
54. Brook, J. D. *et al*. 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* **69**, 385 (1992).
55. Miller, J. W. *et al*. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* **19**, 4439–4448, https://doi.org/10.1093/emboj/19.17.4439 (2000).
56. Fu, Y. H. *et al*. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**, 1256–1258 (1992).
57. Chen, K. Y. *et al*. Length-dependent toxicity of untranslated CUG repeats on Caenorhabditis elegans. *Biochem Biophys Res Commun* **352**, 774–779, https://doi.org/10.1016/j.bbrc.2006.11.102 (2007).
58. Mootha, V. V. *et al*. TCF4 Triplet Repeat Expansion and Nuclear RNA Foci in Fuchs' Endothelial Corneal Dystrophy. *Invest Ophthalmol Vis Sci* **56**, 2003–2011, https://doi.org/10.1167/iovs.14-16222 (2015).
59. Wheeler, T. M. *et al*. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science* **325**, 336–339, https://doi.org/10.1126/science.1173110 (2009).
60. Sobczak, K., Wheeler, T. M., Wang, W. & Thornton, C. A. RNA interference targeting CUG repeats in a mouse model of myotonic dystrophy. *Mol Ther* **21**, 380–387, https://doi.org/10.1038/mt.2012.222 (2013).
61. Rzuczek, S. G. *et al*. Precise small-molecule recognition of a toxic CUG RNA repeat expansion. *Nat Chem Biol* **13**, 188–193, https://doi.org/10.1038/nchembio.2251 (2017).
62. Thornton, C. A., Wang, E. & Carrell, E. M. Myotonic dystrophy: approach to therapy. *Curr Opin Genet Dev* **44**, 135–140, https://doi.org/10.1016/j.gde.2017.03.007 (2017).
63. Wojtkowiak-Szlachcic, A. *et al*. Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res* **43**, 3318–3331, https://doi.org/10.1093/nar/gkv163 (2015).
64. Nakamori, M., Taylor, K., Mochizuki, H., Sobczak, K. & Takahashi, M. P. Oral administration of erythromycin decreases RNA toxicity in myotonic dystrophy. *Ann Clin Transl Neurol* **3**, 42–54, https://doi.org/10.1002/acn3.271 (2016).
65. Oana, K. *et al*. Manumycin A corrects aberrant splicing of Clcn1 in myotonic dystrophy type 1 (DM1) mice. *Sci Rep* **3**, 2142, https://doi.org/10.1038/srep02142 (2013).
66. Phair, R. D. & Misteli, T. High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604–609, https://doi.org/10.1038/35007077 (2000).
67. Childs-Disney, J. L. *et al*. Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules. *Nat Commun* **4**, 2044, https://doi.org/10.1038/ncomms3044 (2013).
68. Dansithong, W., Paul, S., Comai, L. & Reddy, S. MBNL1 is the primary determinant of focus formation and aberrant insulin receptor splicing in DM1. *J Biol Chem* **280**, 5773–5780, https://doi.org/10.1074/jbc.M410781200 (2005).
69. Wojciechowska, M., Taylor, K., Sobczak, K., Napierala, M. & Krzyzosiak, W. J. Small molecule kinase inhibitors alleviate different molecular features of myotonic dystrophy type 1. *RNA Biol* **11**, 742–754 (2014).
70. Caceres, J. F. & Krainer, A. R. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J* **12**, 4715–4726 (1993).

Acknowledgements

The *Nfix* and *NfixMut* minigenes were gifts from M. Ares[19]. $(CAG)_{20}$ and $(AGG)_{20}$ RNA sequences were gifts from W. Krzyżosiak. We thank D. Zielińska for *Atp2a1Δ-(CUG)*$_4$ and *Atp2a1Δ-(CUG)*$_8$ minigenes preparation. Funding: Foundation for Polish Science TEAM program cofinanced by the European Union within the European Regional Development Fund [TEAM/2011-7/10 to KS]; Polish National Science Centre [2011/01/B/NZ1/01603, 2014/15/B/NZ2/02453 to KS, 2014/12/T/NZ2/00516 to ŁJS, 2017/24/C/NZ1/00112 to KT]; National Multidisciplinary Laboratory of Functional Nanomaterials NanoFun [POIG.02.02.00-00-025/09]. 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) program for the years 2014–2019.

Author Contributions

Conception and design of experiments: K.S., K.T., P.C. and Ł.J.S. Acquisition, analysis and interpretation of data: K.T. (all *in vitro* assays and *in cellulo* assays using *Mbnl1* and *Mbnl1Mut* minigenes and *Atp2a1Δ-(CUG)*$_{17}$ hybrid minigene), P.C. (the rest of the *in cellulo* assays). Identification of MBNL targets using CLIP-seq data: P.C. and Ł.J.S. Drafting and revision of the manuscript: K.S., K.T., P.C. and Ł.J.S.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-17816-x.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

SUPPLEMENTARY DATA

Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions

Piotr Cywoniuk[1,#], Katarzyna Taylor[1,#], Łukasz J. Sznajder[1,2], Krzysztof Sobczak[1,*]

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

[2] Current address Center for NeuroGenetics, Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, 2033 Mowry Road PO Box 103610, Gainesville, FL 32610-3610

[#] These authors contributed equally to this work.

* Correspondence should be addressed to K.S. (ksobczak@amu.edu.pl).

Supplementary Materials & Methods

Supplementary Table 1. Sequences of primers for *in cellulo* analysis

Supplementary Table 2. Sequences of primers for *in vitro* analysis

Supplementary Table 3. Sequences of DNA antisense oligonucleotides

Supplementary Table 4. Coordinates of analyzed alternative exons in human (hg19) and mouse (mm10) genomes

Supplementary Figures

Supplementary Figure S1. The presence of YGCY motifs is not sufficient for MBNL1 binding in *in vitro* tests**.**

Supplementary Figure S2. Indication of significant MBNL1-binding motifs within *Atp2a1*-RNA *via* DNA-AONs and mutagenesis.

Supplementary Figure S3. DNA-AONs sufficiently block MBNL1-binding regions *in vitro*.

Supplementary Figure S4. Enzymatic probing of short intronic and exonic RNA fragments.

Supplementary Figure S5. Specificity of SRSF1 binding *in vitro*.

Supplementary Figure S6. Splicing sensitivity of *Atp2a1-(CUG)*$_{17}$ pre-mRNA to MBNL1 overexpression.

Supplementary Figure S7. Raw images of RT-PCR analyses of alternative splicing.

Supplementary Materials & Methods

Supplementary Table S1

Primers' sequences for *in cellulo* analysis

Minigenes' preparation		
Primer name	**Sequence 5'→3'**	
SrcMres_F	ACTACTCGAGGATCTTCAAGCTCCGGGC	
SrcMres_R	ACTGGATCCAGCAATCAGCTAGTCAGTT	
SrcMins_F	GATCTTCAAGCTCCGGGCCCTG	
SrcMins_R	AGCAATCAGCTAGTCAGTTGCC	
pEGFP_F	AAAGACCCCAACGAGAAGCGCGATC	
pEGFP_R	CCTCTACAAATGTGGTATGGCTGATTATGATCAG	
NotSal_F	GCGGCCGCTAAGTCCAGTAGAGCTACTAGTCGACTGGCCGCAGGCTGCTGGGCTGG	
NotSal_R	GTCGACTAGTAGCTCTACTGGACTTAGCGGCCGCTGACCGCAGGCTGCGCATGGGCT	
Mbnl1insF	ACTGCGGCCGCTTTCCTGACGTCTCTGCTGC	
Mbnl1insR	ATCGTCGACTCTCCTGATTAGAAGTAGAAATGG	
NfixinsF	ACTGCGGCCGCAGGACGTTCCCCACGCATGGCTT	
NfixinsR	ATCGTCGACAGGACGCAGTGTCCCCTTAAGC	
Pph1insF	ACTGCGGCCGCAGGACGATGTCTGCTCACTCTGCGGT	
Pph1insR	ATCGTCGACAGGACTTTATATCGGTTACATAAATG	
Ldb3insF	ACTGCGGCCGCAGGACGGGCACCCCTATTGAGCATGCT	
Ldb3insR	ATCGTCGACAGGACTCTCGGATTCTCCACGGGACT	
NASPinsF	ACTGCGGCCGCAGGACCAATACCTAGTGTAAC	
NASPinsR	ATCGTCGACAGGACAGCCTACCAAGTTTCA	
CTRLinsF	ACTGCGGCCGCAGGACGGAGCCTTAGATGTGTCACT	
CTRLinsR	ATCGTCGACAGGACTGAGCAGGTGGGTTGCCTT	
Mbnl1F	ACTGCGGCCGCAGGAAACGGACATCAGTGAATGCTACTTAGTAA	
Mbnl1R	ATCGTCGACAGGAAACGGATGTGCTGTCGTCAGTCTAGCTGG	
Mbnl1mtF	CTGGTTTTCCCTTCCTTGTTGTCAGCTAGACT	
Mbnl1mtR	CTGACAACAAGGAAGGGAAAACCAGAGGCTGC	
CTG17F	GGCCGCGGCCC$(CTG)_{17}$GGGCCG	
CTG17R	TCGACGGCCC$(CAG)_{17}$GGGCCGC	
Splicing analyses		
Primer name	**Sequence 5'→3'**	**Organism**
Atp2A1_F	ATCTTCAAGCTCCGGGCCCT	*Mm*
Atp2A1_R	CAGCTTTGGCTGAAGATGCA	*Mm*
ATP2A1_F	GCTCATGGTCCTCAAGATCTCAC	*Hs*
ATP2A1_R	AGCTCTGCCTGAAGATGTGTCAC	*Hs*
Pphln1_F	ACACAGCAGATCCGGCTCCAGT	*Hs/Mm*
Pphln1_R	AGCCCACTTGCTTTCAGCCTCA	*Hs/Mm*
NASP_F1	GGTGTGCATGTGGAAGAGG	*Hs*
NASP_F2	AGCTGGAGATGGGGTTGATA	*Hs*
NASP_R1	TTTGGCATTTCTTCGGTCTT	*Hs*

NFIX_F	GAGCCCTGTTGATGACGTGTTCTA	*Hs*
NFIX_R	CTGCACAAACTCCTTCAGTGAGTC	*Hs*
Nfix_F	GAGTCCAGTAGATGATGTGTTCTA	*Mm*
Nfix_R	CTGCACAAACTCCTTCAGCGAGTC	*Mm*
LDB3_F	AGCCAGAAGATGAGGCTGACGA	*Hs*
LDB3_R	TCACTGTAGCTGGTGTGGGTGGCA	*Hs*
Ldb3_F	ACCCTGATGAAGAGGCTCTGCG	*Mm*
Ldb3_R	TGGACCCTCGCTGTAGCTGGTA	*Mm*
MBNL1ex1F	CAGCGACATGCAACAGTCTT	*Hs*
MBNL1ex1R	TGTCAGCAGGATGAGCAAAC	*Hs*
Mbnl1ex1F	GAGTCGCGATCCCACAATG	*Mm*
Mbnl1ex1R	TGTCAGCAGGATGAGCAAAC	*Mm*
MBNL2_F	TCCTTTACCAAAGAGACAAGCAC	*Hs*
MBNL2_R	CTCAATGCAGATTCTTGGCATTCC	*Hs*
NCOR2_F	ACACCCACAACCGGAATGAGCCTG	*Hs*
NCOR2_R	GGACTTGGCTTTTCGGCTGCTG	*Hs*
PHKA1_F	TGCACACACTTGAGCTTCATGGA	*Hs*
PHKA1_R	AAAGTCCACCTCCCCAGACTGGTC	*Hs*

Hs, *Homo sapiens*; *Mm*, *Mus musculus*

Supplementary Table S2

Primers' sequences for *in vitro* analysis

Primer name	**Sequence 5'→3'**	**Ta [°C]**
Atp2a1_F	GGTGGGATTTTCTCCCAACCT	60
Atp2a1_R	GCTGTATCTCCTTGCGCATC	
Atp2a1_TF	TAATACGACTCACTATAGGGCACCCCTGCGGTCACT	
Atp2a1_TR	AGCAGACTGGACGGCCA	
Atp2a1mt_upF	TAATACGACTCACTATAGGGCACCCCTGCGGTCACTTGCGCGATCTCGCA	
Atp2a1mt_upR	Atp2a1_TR	
Atp2a1mt_#1F	ATGTTGTTGCCACTGCCCGCTTC	
Atp2a1mt_#1R	GGGCAGTGGCAACAACATCAGTTG	
Atp2a1mt_#2F	ACACTGACCTCTTCCACATGAG	
Atp2a1mt_#2R	CATGTGGAAGAGGTCAGTGTCAACA	
Atp2a1mt_up#1F	Atp2a1mt_#1F on Atp2a1_up#1 template	
Atp2a1mt_up#1R	Atp2a1mt_#1R	
Atp2a1mt_#1-#2F	ATGATGTTGACACTGACCTCTTCCACATGA	
Atp2a1mt_#1-#2R	AGGTCAGTGTCAACATCATCAGTGGCTCCAG	
Atp2a1mt_YGCYF	CACATGAGGGCCATGACGTTA	
Atp2a1mt_YGCYR	TAACGTCATGGCCCTCATGTG	
Atp2a1mt_YGCY_TF	TAATACGACTCACTATAGGGCACCCCTGCGGTCACTTGCGCGATCTCGCA	
Pphln1_F	AGGTAGTTGGTTGTCTGGGC	60
Pphln1_R	CTCAGACTTCGCCTTCCAGA	
Pphln1_TF	TAATACGACTCACTATAGGGACGATGTCTGCTCACTCTGCGGT	
Pphln1_TR	GGGACTTTATATCGGTTACATAAATG	
NASP_F	GGTGTGCATGTGGAAGAGG	55

NASP_R	TCTCAACTCTTCCCTTGCTTC	
NASP_TF	TAATACGACTCACTATAGGCAATACCTAGTGTAACAAGC	
NASP_TR	AGCCTACCAAGTTTCAGCAGATGGC	
Nfix_F	CTTTATCGACCGTCAGGAGCA	55
Nfix_R	GCAGTGTCCCCTTAAGCCCC	
Nfix_TF	TAATACGACTCACTATAGGGAGCGGGCTGCTAGCTGGCTTCCT	
Nfix_TR	CCCCGCCCCCATCTCTC	
Ldb3_F	CCAGACTGTCCTTTCCACCC	60
Ldb3_R	ATCCAGCTGTCCACCAATCG	
Ldb3_TF	TAATACGACTCACTATAGGGCACCCCTATTGAGCATGCT	
Ldb3_TR	CTCGGATTCTCCACGGGACT	
Clcn1_F	CCTTCCATGTTTCCTCCTGTG	57
Clcn1_R	ACCAAGGTAGGGAGGAAGTG	
Clcn1_TF	TAATACGACTCACTATAGGCCCCCGTTCTTCTGTG	
Clcn1_TR	AAAGTAGCATCCCACGCCCA	
Tnnt3_F	GAAGGTAGAGACCGGCTTCG	57
Tnnt3_R	GTCACTGAGGAACCATGCCA	
Tnnt3_TF	TAATACGACTCACTATAGGACAGCAAGTATCCATGAAGG	
Tnnt3_TR	GCACAAGCGCACATGCAAGA	
Mbnl1_F	GATGGCTGGCTGCAATATGCC	57
Mbnl1_R	CCTAGGGCAATGGCAGATACTC	
Mbnl1_TF	TAATACGACTCACTATAGGGTTGGTACTAAGAAGTGCCT	
Mbnl1_TR	CCTAGGGCAATGGCAGATACTC	
Mfn-Ctrl1_F	GCAGGCAAGCACACTTACTG	57
Mfn-Ctrl1_R	GCTGGGTGAACCTGAACACT	
Mfn-Ctrl1_TF	TAATACGACTCACTATAGGGACTGGAGAGATGGCTCAGT	
Mfn-Ctrl1_TR	GCTGGGTGAACCTGAACACT	
Mfn-Ctrl3_2F	GCTAGCAACTTATTTGAGGAGCC	55
Mfn-Ctrl3_2R	AGCAGGTGGGTTGCCTT	
Mfn-Ctrl3_2TF	TAATACGACTCACTATAGGAGCCTTAGATGTGTCACT	
Mfn-Ctrl3_2TR	AGCAGGTGGGTTGCCTT	
Atp2a1-Ctrl2_F	TGTTCCCCCTGCTTTGTTCA	57
Atp2a1-Ctrl2 _R	TCTCCATGACGGTCTGTGAC	
Atp2a1-Ctrl2 _TF	TAATACGACTCACTATAGGGCTTCCTGCTGTCATCACC	
Atp2a1-Ctrl2_TR	CTGACATCTGGTTGGTGGTG	
Capzb-Ctrl4_F	CCTTCTCTCTGCGGTCGGTC	57
Capzb-Ctrl4_R	CTCCCCAGAAAGGGGCAAAC	
Capzb-Ctrl4_TF	TAATACGACTCACTATAGGGACTTCTCACGACTTAGCCT	
Capzb-Ctrl4_TR	CCAGAGGCTGCCCAGAAAG	
SRSF1_F	ATCGGGATCCAATGTCGGGAGGTGGTGTGATTCGT	
SRSF1_R	AACTCGGCGGCCGCTCATACCTCATGAGATCTAAACTTAGTG	

Supplementary Table S3

Sequences of DNA antisense oligonucleotides

AONs' name	**Sequence 5'→3'**
DNA	GCGGGCAGTGGCAACAGCAGC
DNA-up	GCTGCGAGCGCGCGCAAGTGACCGC
DNA-dw	AGTAACGGCATGGCCCTCAT
DNA-Pphln1	GTGGAGAAGCAAAAAGCAAGA
DNA-Pphln1-2	AGCAAAAACCGCAGAGTGAGCA
DNA-NASP	AGCATCCCGGCTAGCAGAGCA
DNA-NASP-2	AGCCTACCAAGTTTCAGCAGATGGC
DNA-Nfix	CCAGCAAGCACAGGCAGCGGG
DNA-Nfix-2	GCAGGGCCTGGGGGGAGGAAGCAG
DNA-Ldb3	GGCAGAAGCAGGCAGCAGCG
DNA-Ldb3-2	CTGGTGCACACTGGAGCATG
DNA-Ldb3-3	CGAGGCCGCAATGGGAGAGGCAGCAG
DNA-Ldb3-4	CGGCAGCAGCATGGGTGGCAGCAG
DNA-Ldb3-5	GCGGCAGGGCCTGCAGCA

Supplementary Table S4

Coordinates of analyzed alternative exons in human (hg19) and mouse (mm10) genomes

gene name & alternative exon	**human genome coordinates**	**mouse genome coordinates**
ATP2A1 ex22	hg19 chr16:28,915,020-28,915,064	mm10 chr7:126,446,540-126,446,594
PPHLN1 ex6	hg19 chr12:42,384,939-42,384,997	mm10 chr15:93,452,117-93,452,174
NASP ex7	hg19 chr1:45,607,312-45,608,363	not analyzed
NFIX ex7	hg19 chr19:13,078,607-13,078,736	mm10 chr8:84,723,727-84,723,851
LDB3 ex10	hg19 chr10:88,466,287-88,466,478	mm10 chr14:34,555,340-34,555,529
MBNL1 ex1	hg19 chr3:152,299,398-152,300,377	mm10 chr3:60,528,755-60,529,811
MBNL2 ex7	hg19 chr13:97,365,135-97,365,172	not analyzed
NCOR2 ex45	hg19 chr12:124,327,386-124,327,638	not analyzed
PHKA1 ex19	hg19 chrX:72,620,713-72,620,927	not analyzed

Supplementary Figures

Supplementary Figure S1

a

Fragments of transcripts of previously confirmed MBNL1 targets:

Atp2a1:
GGGCACCCCUGCGGUCACUUGCGCGCGCUCGCAGCUCUCCCUGGAGCCACUGCUGCUGUUGCCACUGCCCGCUUCC
ACAUGAGGGCCAUGCCGUUACUGUGGGCACUGGCAGCUGUGAGGCUGUGUCCUGGCCGUCCAGUCUGCUA

Clcn1:
GGCCCCCGUUCUUCUGUGCUUCCUGACACCCAUCCACCUGGUUUACAUACCACCUGUCUGUCCCCCUCUGCCACCUG
CCUCGCCCGUCGUGCUUCUCUGUUGCAGACCGUGCCUGGGCAGCUUGAUCUCCUGGUGCCAGCCUGUGCAGUGGGC
GUGGGAUGCUACUUUA

Tnnt3:
GGACAGCAAGUAUCCAUGAAGGGGAGGAGACCUAUGGGAAGGUGUAGGAGCUGCCCUGCCACCCUCAGGCAGGGGG
CUCUGUGCGUGCCUGCCUCUGCUGUUCCAUGUGGCCACUGCUGCUGUGUGGCCUCUGCACUUCUGCAGCUGCGUGG
CUUUUUUCUUGCAUGUGCGCUUGUGC

Fragments of transcripts - negative controls:

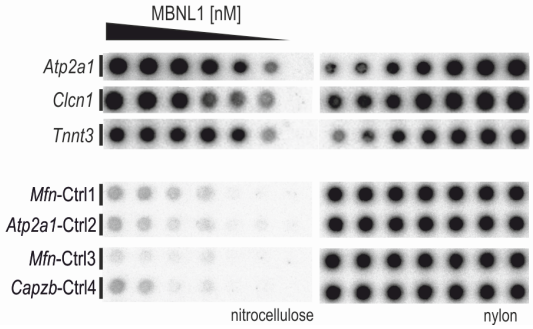
***Mfn*-Ctrl1**:
GGGACUGGAGAGAUGGCUCAGUGGUUAAGAGCGCUGGCUGCUCUUGCAGAGUCCAUAAGUUCAGUUCCCAGCUCAUC
AUCAUCUGUAGUUGGAUCUGAUGCUCUCUUCCAGUGUGCCAUAAAUAAAAUUUUUUUUUAAAGUACUGAAGAGUCAGU
GUUCAGGUUCACCCAGCA

***Atp2a1*-Ctrl2**:
GGGCUUCCUGCUGUCAUCACCACCUGCUUGGCUUUGGGUACCCGCCGGAUGGCAAAGAAGAACGCCAUCGUGAGGAG
UCUGCCCUCUGUGGAGACCCUGGGCUGUACCUCUGUCAUCUGUUCUGACAAAACAGGGACCCUCACCACCAACCAGA
UGUCAGA

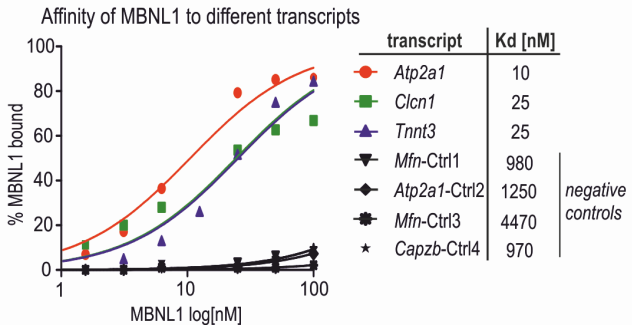
***Mfn*-Ctrl3**:
GGCACCUAUCAGUGACCUUUUUCCUUCUACCAAGUCACUUGUGGGCUGGGUACUACUUAUCUCCUCCAGCUGACCUG
UUUAUUCUGUUUCCAUGGUAGCAGGUGUAAGGCAUGCAGGAUGUUAGUAGGGUUACUAACCAACAUUGCA

***Capzb*-Ctrl4**:
GGGAGCCUUAGAUGUGUCACUUCCCUCCUCCCCCAUACCCCUGCACCCCCUAAUCUACCUGGAUACCUGCAGAGCAG
GCCCACCAUCUUUGGUGGUGGAAGGGCCAGAGAACUACCUCACAUGGUCAAGGAAGCCUUGUAAGGUAGGCUGAGAG
GGAAGGCAACCCACCUG

b



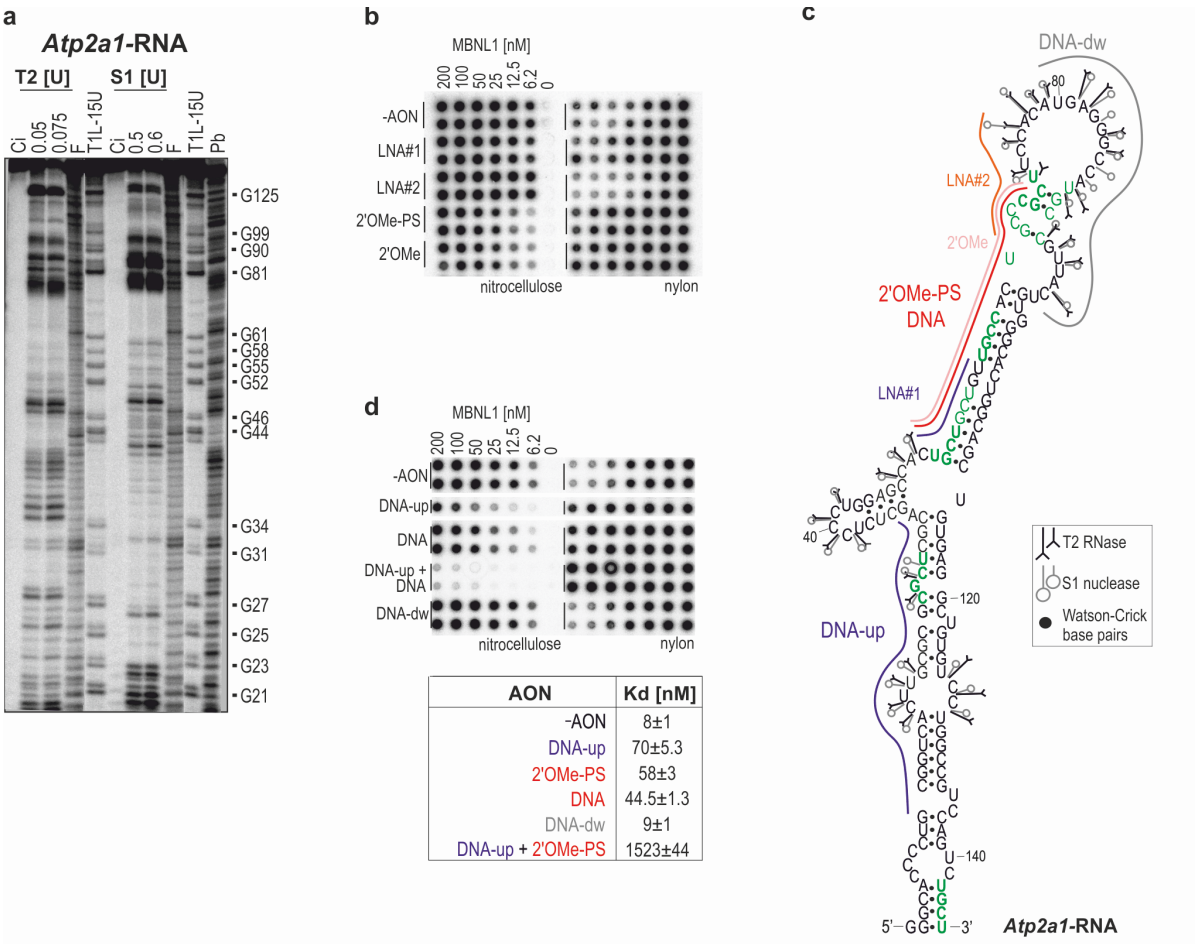
c



Supplementary Figure S1

The presence of YGCY motifs is not sufficient for MBNL1 binding in *in vitro* tests. (**a**) The sequences of transcript fragments previously described in literature as MBNL1 bound: *Atp2a1*[1], *Clcn1*[2], *Tnnt3*[3] and ~150 nt control RNAs (RNA-Ctrl) with or without YGCY motifs in their sequence. YGCY motifs are marked with green and in bold if conservatively maintained. (**b**) Filter binding assay (FBA) represents the affinity of recombinant MBNL1 to particular transcripts. Nitrocellulose membrane holds signals derived from the MBNL1/RNA complexes, whereas the nylon membrane keeps signals of free RNA. MBNL1 concentration used for *Atp2a1*, *Clcn1*, *Mfn*-Ctrl1, *Atp2a1*-Ctrl2, *Mfn*-Ctrl3, *Capzb*-Ctrl4: 100, 50, 25, 6.25, 3.1, 1.5, 0 nM and for *Tnnt3*: 100, 50, 25, 12.5, 6.25, 3.1, 0 nM; the data coming from different parts of the same membrane are divided with white space; n = 1 (**c**) The quantification of FBA results. Dissociation constant (Kd) for each previously examined fragment of transcript oscillates between 10-25 nM, whereas for control samples the Kd value is at least 38 fold higher.

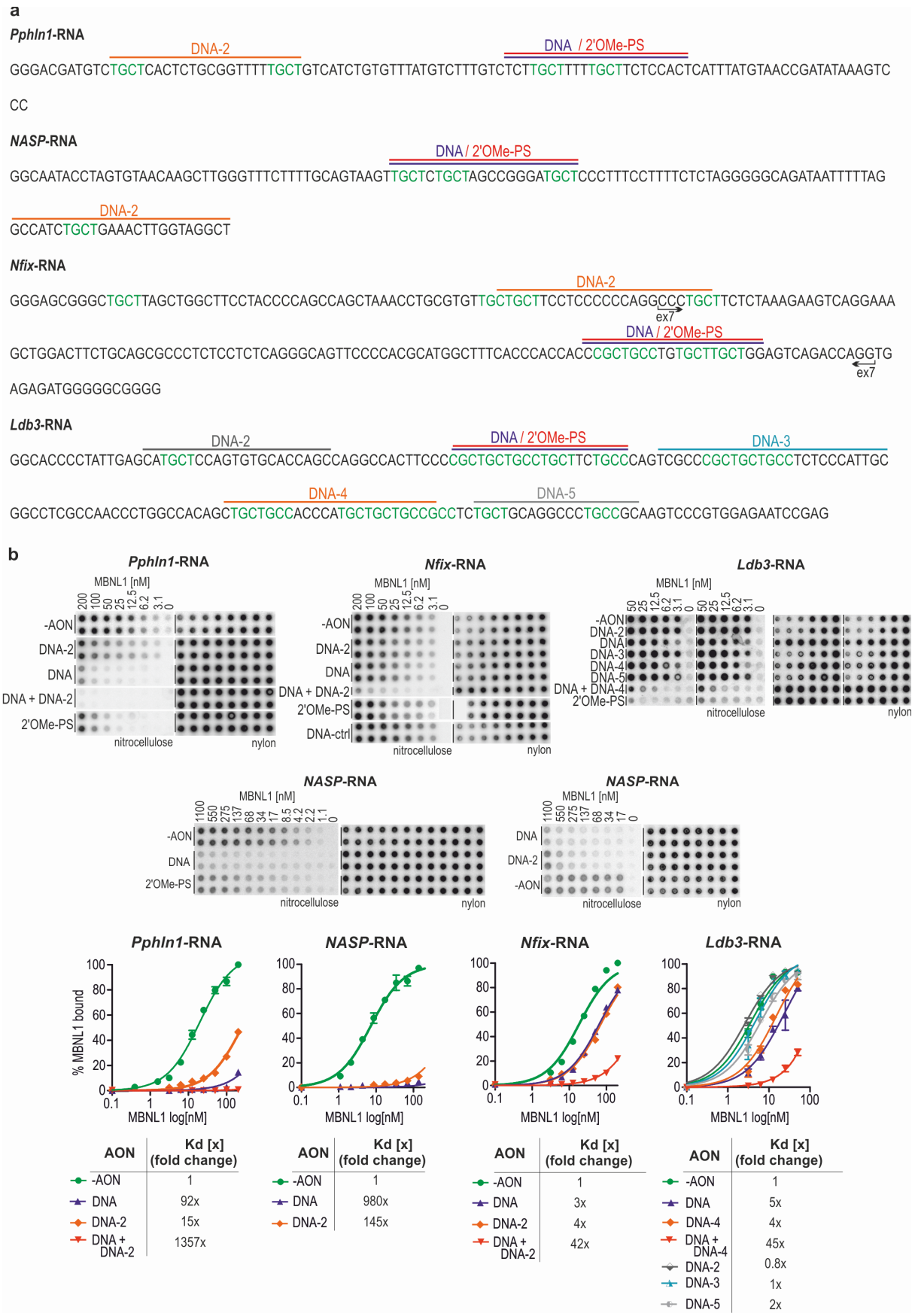
Supplementary Figure S2



Supplementary Figure S2

Indication of significant MBNL1-binding motifs within *Atp2a1*-RNA *via* DNA-AONs and mutagenesis. (**a**) The structure probing of analyzed 5'-^{32}P-labeled *Atp2a1*-RNA fragment using RNases. Ci- untreated RNA samples; F – formamid ladder; Pb – lead ion ladder; T1L - G ladder formed by RNase T1 cleavage in denaturing conditions; T2 and S1 - endonucleases which cleave single stranded regions; particular guanosines are marked as G on the right side of the picture. The cleavage sites, and intensities for the selected probes are shown by using symbols explained in the inset. (**b**) Representative raw data from FBA showing differently modified AON's inhibitory properties of MBNL1/RNA complexes. (**c**) As in **Figure 1c** but with additional DNA-AONs bound *in vitro* to regions upstream (DNA-up, marked with dark blue) and downstream (DNA-dw, marked with grey) of major MBNL1 binding regions #1 and #2. (**d**) *Top*, representative raw data from FBA showing the blocking properties of AONs (DNA-based) and revealing additional MBNL1 bound region upstream region #1; the data coming from different parts of the same membrane are divided with white space. *Bottom,* quantification of FBA. The Kd values are presented with SD; n = 2. (**e**) Top, representative raw data from FBA showing reduced affinity of MBNL1 to *Atp2a1*-RNA mutants of significant binding sites; the data coming from different parts of the same membrane are divided with white space. *Bottom*, the sequences of Atp2a1-RNA mutants; YGCY motifs are marked with green and in bold if conservatively maintained. The major MBNL1-binding regions are marked as #1 and #2; up, the YGCY motif upstream from region #1; dw, the YGCY motif downstream from region #2. On the right, quantification of FBA. The Kd values are presented with SD; n = 4. (**f**) Percentage of MBNL-dependent alternative exons inclusion in *NCOR2*, *PHKA1* and *MBNL2* transcripts upon MBNL1 overexpression and treatment with 100 nM of *Atp2a1*-specific AONs; n = 3.

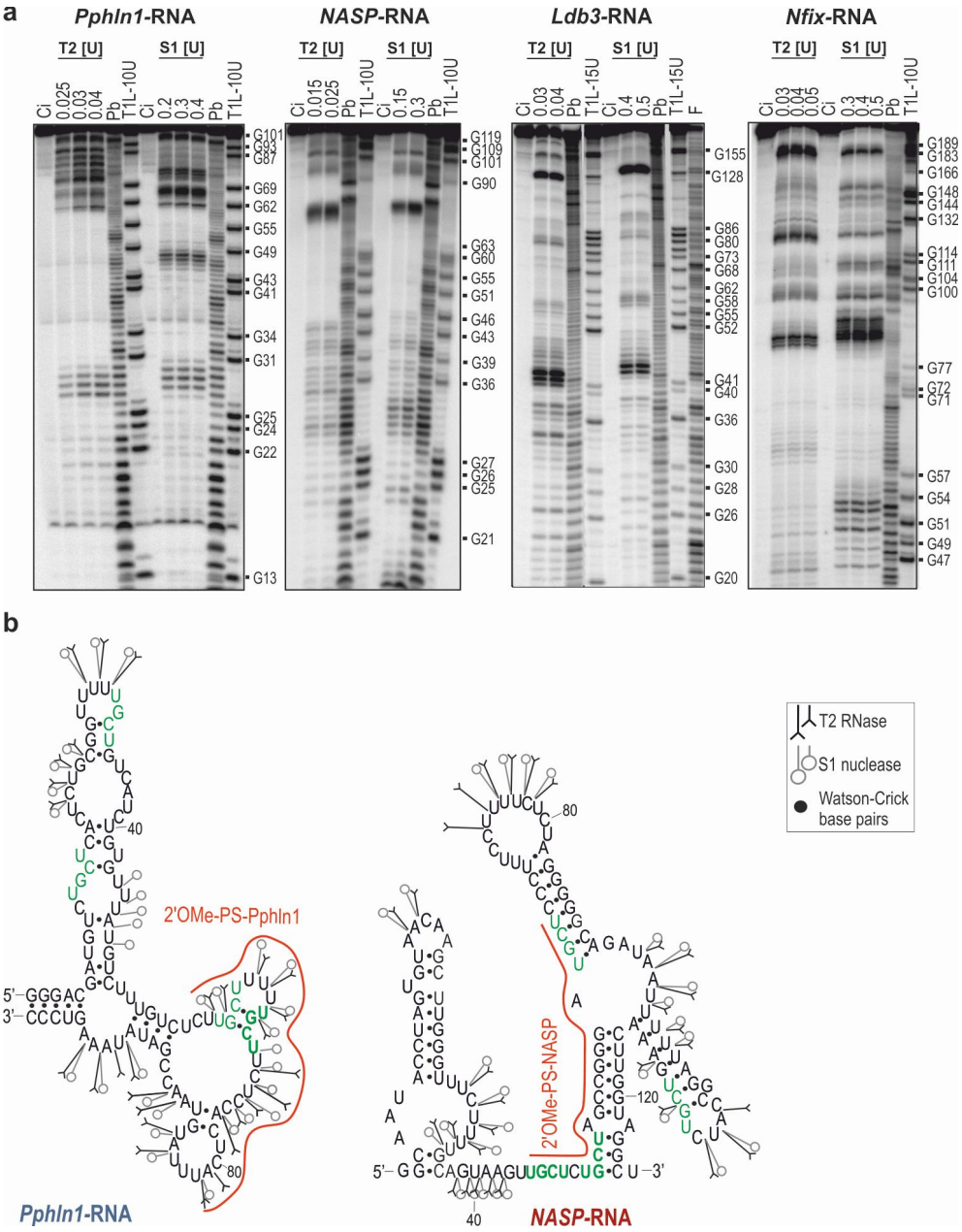
Supplementary Figure S3

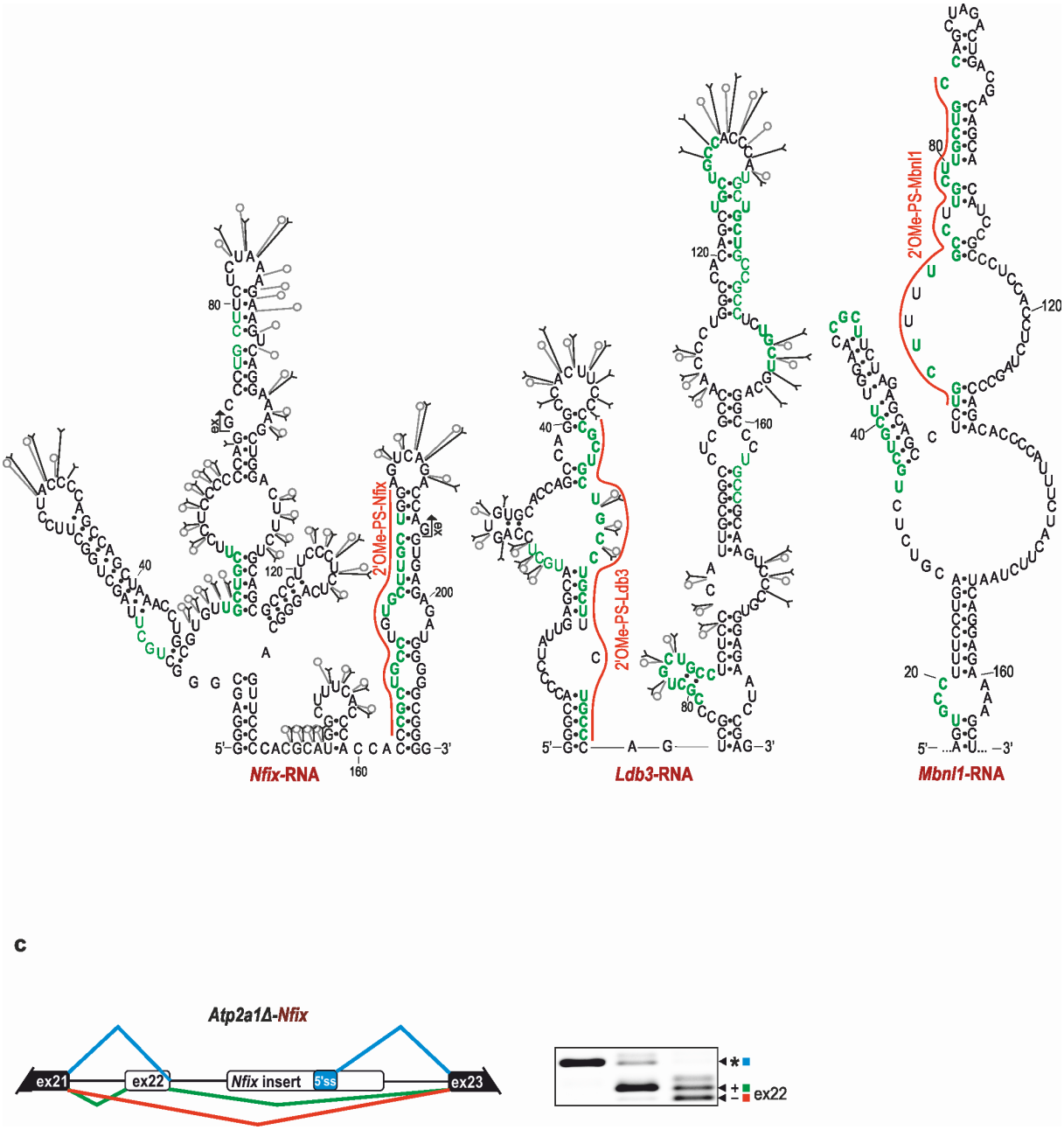


Supplementary Figure S3

DNA-AONs sufficiently block MBNL1 binding motifs *in vitro*. (**a**) The sequences of *Atp2a1*, *Pphln1*, *NASP*, *Nfix*, *Ldb3* and *Mbnl1* transcript fragments with highlighted YGCY containing regions bound by specific DNA-AONs or RNA-based AONs *in vitro*. (**b**) *Top*, representative raw data from FBA showing the blocking properties of AONs (DNA- and RNA-based) and indicating significant MBNL1-binding regions; the data coming from different parts of the same membrane are divided with white space. *Bottom*, quantification of FBA; below each chart, there is a fold change of Kd values normalized to Kd for MBNL1/RNA complexes without AON application (-AON); n = 2.

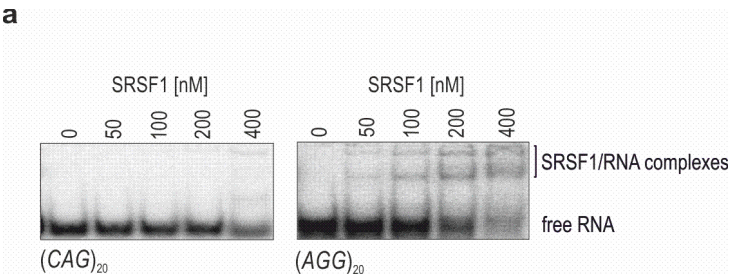
Supplementary Figure S4





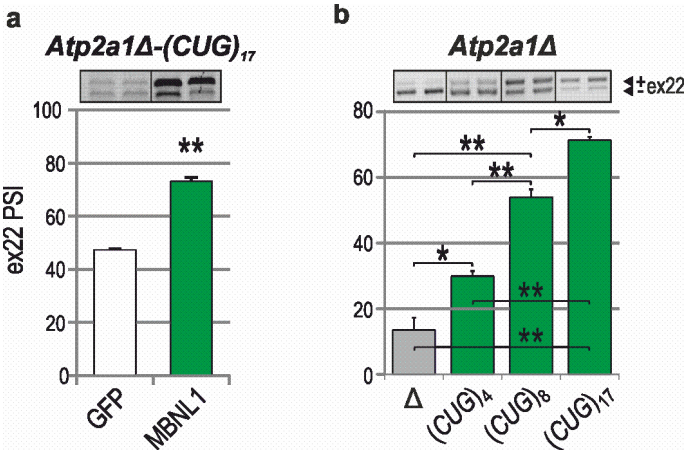
Enzymatic probing of short intronic and exonic RNA fragments. (**a**) As in **Supplementary Figure S2a** but for 5'-^{32}P-labeled *Pphln1*-, *NASP*-, *Nfix*-, *Ldb3*-RNA using specific RNases. Ci- untreated RNA samples, Pb – lead ion ladder, F – formamid ladder, T1L - G ladder formed by RNase T1 cleavage in denaturing conditions; particular guanosines are marked as G on the right of each electrophoregram. (**b**) Experimentally determined secondary structures of short intronic and exonic RNA fragments containing YGCY motifs marked with green and in bold if present in human and mouse. Significant for MBNL1-binding regions and complementary to particular AONs are indicated with a red line. (**c**) A schematic representation of splicing isoforms' distribution of *Atp2a1Δ-Nfix* mRNA. An asterisk points to an artificial isoform composed of *Nfix* insert and *Atp2a1* intron 22 fragment.

Supplementary Figure S5



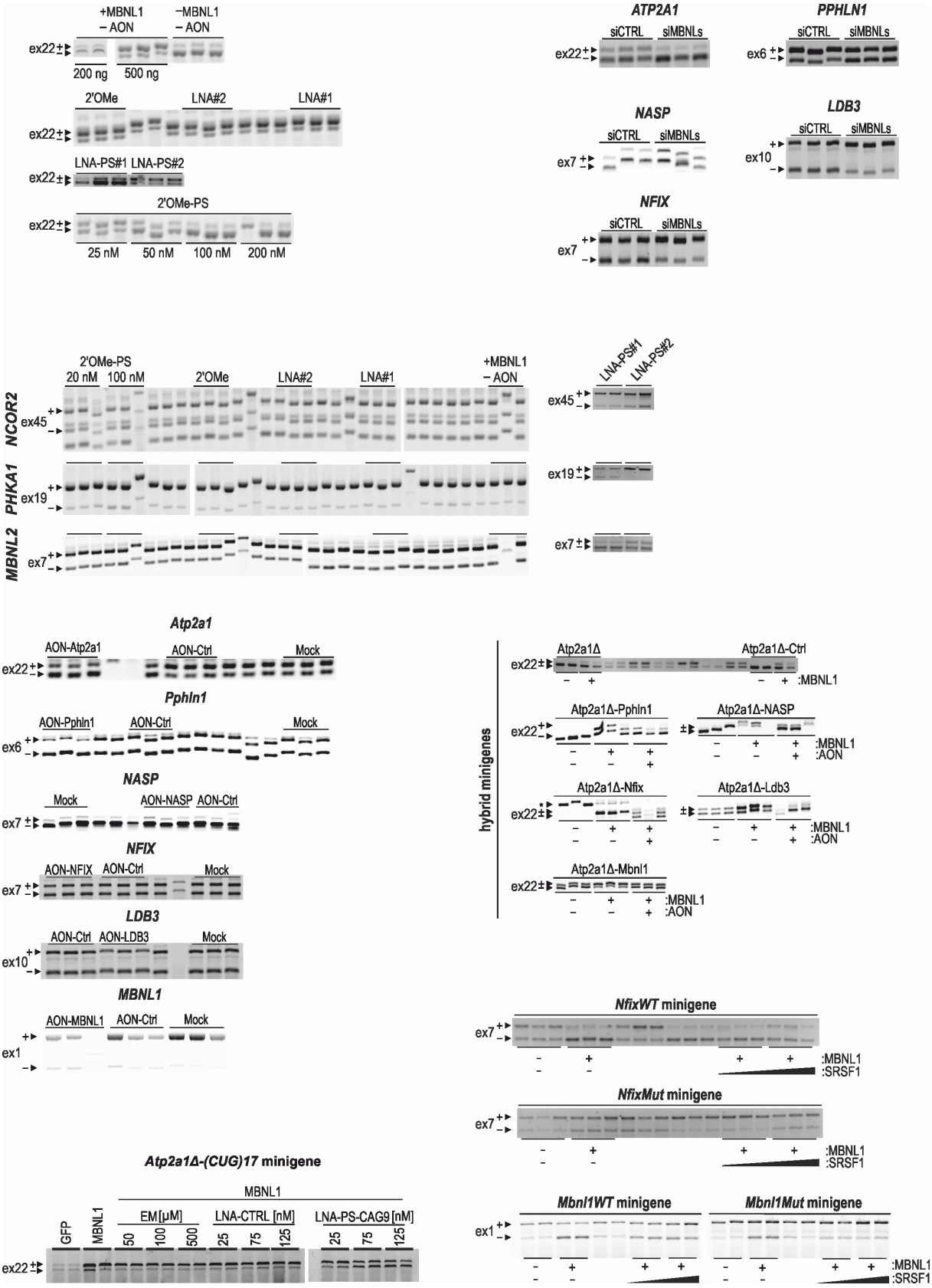
Specificity of SRSF1 binding *in vitro*. (**a**) EMSA on a native PA gel showing no or weak interaction between 5'-^{32}P-labeled negative control RNA sequences $(CAG)_{20}$ and $(AGG)_{20}$ and SRSF1 at indicated concentrations.

Supplementary Figure S6



The hybrid minigene is useful for screening of potential RNA binding therapeutics. (**a**) The efficacy of alternative ex22 inclusion into *Atp2a1-(CUG)$_{17}$* mRNA induced by endogenous MBNLs (white bar) or MBNL1 overexpression (green bar); n = 2. (**b**) A gradual increase in the number of MBNL1-binding sites in *Atp2a1Δ-(CUG)n* (n = 4, 8, 17) minigenes positively correlates with the enhancement of alternative ex22 inclusion upon MBNL1 overexpression; n = 2.

Supplementary Figure S7



Supplementary Figure S7

Raw images of RT-PCR analyses of alternative splicing. (**a**) Original agarose gels presenting all RT-PCR results of alternative splicing analyses on which calculations and statistical data were based.

Supplementary references

1 Hino, S. *et al.* Molecular mechanisms responsible for aberrant splicing of SERCA1 in myotonic dystrophy type 1. *Hum Mol Genet* **16**, 2834-2843, doi:10.1093/hmg/ddm239 (2007).

2 Kino, Y. *et al.* MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res* **37**, 6477-6490, doi:10.1093/nar/gkp681 (2009).

3 Yuan, Y. *et al.* Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res* **35**, 5474-5486, doi:10.1093/nar/gkm601 (2007).

CO-AUTHOR STATEMENT

I declare that the work in the paper ,,**Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions** " Piotr Cywoniuk, Katarzyna Taylor, Łukasz J. Sznajder & Krzysztof Sobczak in ***Scientific Reports*** 2017, of which I am a co-author, is a part of my PhD dissertation (P.C.). I designed, prepared and analyzed all *in cellulo* experiments (except assays using Mbnl1 and Mbnl1Mut minigenes and Atp2a1Δ-$(CUG)_{17}$ hybrid minigene) including Atp2a1WT, Atp2a1Δ, -Pphln1, -NASP, -Nfix, -Ldb3, -Mbnl1 development/composition as well as identification of MBNL targets using CLIP-seq data and design of all AON sequences against particular transcripts' regions (except 2'OMe-PS-NASP) (Fig. 1a, b, f, g, S2f, 2a, d, S4c, 3a, b, d, S7). I performed PCR analysis of alternative splicing of MBNL targets in MBNL1,2-silenced HSkM cell line (silencing performed by Ł.J.S.) and *in cellulo* experiments with NfixWT and NfixMut minigenes (Fig. 2b and 4d, respectively). I prepared a figure legend of Fig. 1a, b, f, g, S2f, 2a, b, d, S4c, 3a, b, d, 4d, S7 and materials and method section describing "Oligonucleotides", "Plasmid preparation" (with K.T.), "Cell culture and transfection" (with K.T.) and "Splicing analyses of mature mRNA". With K.T. and K.S. I prepared draft of whole manuscript.

K.T. (Katarzyna Taylor, first author ex aequo) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) with various RNA fragments and rMBNL1 as well as antisense oligonucleotides (AONs) as MBNL:RNA complex inhibitors and all recombinant proteins purification (Fig.1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a, b, S5). KT designed and prepared Mbnl1, Mbnl1Mut and Atp2a1Δ-$(CUG)_{17}$ minigenes and performed and analyzed *in cellulo* assays using them (Fig. 4c, S6, 5). K.T. prepared a figure legend of Fig. 1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a-c, 5, S5, S6 and materials and method section describing "Plasmid preparation Plasmid preparation" (with P.C.), "Cell culture and transfection" (with P.C.), "Preparation of radiolabeled RNA fragments", "Quantification of the MBNL1/RNA interaction and its inhibition by AONs *in vitro*", "Quantification of the SRSF1/RNA interaction and its inhibition by AONs *in vitro*" and "Statistical analysis".

Ł.J.S. (Łukasz J. Sznajder) identified MBNL targets in *NASP* transcript using CLIP-seq data and designed 2'OMe-PS-NASP AON (used in Fig. 2a, c, d, 3c, d, S3a, c, S4b, S7).and prepared MBNL1 and 2 knock down in HSkM cells (used in Fig. 2b). Ł.J.S. drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 15.10.2018, Poznań

Name: Piotr Cywoniuk

Signature: Piotr Cywoniuk

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper ,,**Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions** '' Piotr Cywoniuk, Katarzyna Taylor, Łukasz J. Sznajder & Krzysztof Sobczak in ***Scientific Reports*** 2017, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who designed, prepared and analyzed all *in cellulo* experiments (except assays using Mbnl1 and Mbnl1Mut minigenes and Atp2a1Δ-$(CUG)_{17}$ hybrid minigene) including Atp2a1WT, Atp2a1Δ, -Pphln1, -NASP, -Nfix, -Ldb3, -Mbnl1 development/composition as well as identification of MBNL targets using CLIP-seq data and design of all AON sequences against particular transcripts' regions (except 2'OMe-PS-NASP) (Fig. 1a, b, f, g, S2f, 2a, d, S4c, 3a, b, d, S7). P.C. performed PCR analysis of alternative splicing of MBNL targets in MBNL1,2-silenced HSkM cell line (silencing performed by Ł.J.S.) and *in cellulo* experiments with NfixWT and NfixMut minigenes (Fig. 2b and 4d, respectively). P.C. prepared a figure legend of Fig. 1a, b, f, g, S2f, 2a, b, d, S4c, 3a, b, d, 4d, S7 and materials and method section describing "Oligonucleotides", "Plasmid preparation" (with K.T.), "Cell culture and transfection" (with K.T.) and "Splicing analyses of mature mRNA". P.C. with K.T. and K.S. prepared draft of whole manuscript.

K.T. (Katarzyna Taylor, first author ex aequo) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) with various RNA fragments and rMBNL1 as well as antisense oligonucleotides (AONs) as MBNL:RNA complex inhibitors and all recombinant proteins purification (Fig.1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a, b, S5). KT designed and prepared Mbnl1, Mbnl1Mut and Atp2a1Δ-$(CUG)_{17}$ minigenes and performed and analyzed *in cellulo* assays using them (Fig. 4c, S6, 5). K.T. prepared a figure legend of Fig. 1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a-c, 5, S5, S6 and materials and method section describing "Plasmid preparation Plasmid preparation" (with P.C.), "Cell culture and transfection" (with P.C.), "Preparation of radiolabeled RNA fragments", "Quantification of the MBNL1/RNA interaction and its inhibition by AONs *in vitro*", "Quantification of the SRSF1/RNA interaction and its inhibition by AONs *in vitro*" and "Statistical analysis".

Ł.J.S. (Łukasz J. Sznajder) identified MBNL targets in *NASP* transcript using CLIP-seq data and designed 2'OMe-PS-NASP AON (used in Fig. 2a, c, d, 3c, d, S3a, c, S4b, S7).and prepared MBNL1 and 2 knock down in HSkM cells (used in Fig. 2b). Ł.J.S. drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 20.09. 2018, Poznań

Name: Katarzyna Taylor

Signature: Taylor

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions** " Piotr Cywoniuk, Katarzyna Taylor, Łukasz J. Sznajder & Krzysztof Sobczak in ***Scientific Reports*** 2017, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who designed, prepared and analyzed all *in cellulo* experiments (except assays using Mbnl1 and Mbnl1Mut minigenes and Atp2a1Δ-$(CUG)_{17}$ hybrid minigene) including Atp2a1WT, Atp2a1Δ, -Pphln1, -NASP, -Nfix, -Ldb3, -Mbnl1 development/composition as well as identification of MBNL targets using CLIP-seq data and design of all AON sequences against particular transcripts' regions (except 2'OMe-PS-NASP) (Fig. 1a, b, f, g, S2f, 2a, d, S4c, 3a, b, d, S7). P.C. performed PCR analysis of alternative splicing of MBNL targets in MBNL1,2-silenced HSkM cell line (silencing performed by Ł.J.S.) and *in cellulo* experiments with NfixWT and NfixMut minigenes (Fig. 2b and 4d, respectively). P.C. prepared a figure legend of Fig. 1a, b, f, g, S2f, 2a, b, d, S4c, 3a, b, d, 4d, S7 and materials and method section describing "Oligonucleotides", "Plasmid preparation" (with K.T.), "Cell culture and transfection" (with K.T.) and "Splicing analyses of mature mRNA". P.C. with K.T. and K.S. prepared draft of whole manuscript.

K.T. (Katarzyna Taylor, first author ex aequo) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) with various RNA fragments and rMBNL1 as well as antisense oligonucleotides (AONs) as MBNL:RNA complex inhibitors and all recombinant proteins purification (Fig.1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a, b, S5). KT designed and prepared Mbnl1, Mbnl1Mut and Atp2a1Δ-$(CUG)_{17}$ minigenes and performed and analyzed *in cellulo* assays using them (Fig. 4c, S6, 5). K.T. prepared a figure legend of Fig. 1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a-c, 5, S5, S6 and materials and method section describing "Plasmid preparation Plasmid preparation" (with P.C.), "Cell culture and transfection" (with P.C.), "Preparation of radiolabeled RNA fragments", "Quantification of the MBNL1/RNA interaction and its inhibition by AONs *in vitro*", "Quantification of the SRSF1/RNA interaction and its inhibition by AONs *in vitro*" and "Statistical analysis".

Ł.J.S. (Łukasz J. Sznajder) identified MBNL targets in *NASP* transcript using CLIP-seq data and designed 2'OMe-PS-NASP AON (used in Fig. 2a, c, d, 3c, d, S3a, c, S4b, S7).and prepared MBNL1 and 2 knock down in HSkM cells (used in Fig. 2b). Ł.J.S. drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 27.09.2018, Gainesville, FL

Name: Łukasz J. Sznajder

Signature:

Łukasz Sznajder

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions** " Piotr Cywoniuk, Katarzyna Taylor, Łukasz J. Sznajder & Krzysztof Sobczak in ***Scientific Reports*** 2017, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who designed, prepared and analyzed all *in cellulo* experiments (except assays using Mbnl1 and Mbnl1Mut minigenes and Atp2a1Δ-$(CUG)_{17}$ hybrid minigene) including Atp2a1WT, Atp2a1Δ, -Pphln1, -NASP, -Nfix, -Ldb3, -Mbnl1 development/composition as well as identification of MBNL targets using CLIP-seq data and design of all AON sequences against particular transcripts' regions (except 2'OMe-PS-NASP) (Fig. 1a, b, f, g, S2f, 2a, d, S4c, 3a, b, d, S7). P.C. performed PCR analysis of alternative splicing of MBNL targets in MBNL1,2-silenced HSkM cell line (silencing performed by Ł.J.S.) and *in cellulo* experiments with NfixWT and NfixMut minigenes (Fig. 2b and 4d, respectively). P.C. prepared a figure legend of Fig. 1a, b, f, g, S2f, 2a, b, d, S4c, 3a, b, d, 4d, S7 and materials and method section describing "Oligonucleotides", "Plasmid preparation" (with K.T.), "Cell culture and transfection" (with K.T.) and "Splicing analyses of mature mRNA". P.C. with K.T. and K.S. prepared draft of whole manuscript.

K.T. (Katarzyna Taylor, first author ex aequo) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) with various RNA fragments and rMBNL1 as well as antisense oligonucleotides (AONs) as MBNL:RNA complex inhibitors and all recombinant proteins purification (Fig.1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a, b, S5). KT designed and prepared Mbnl1, Mbnl1Mut and Atp2a1Δ-$(CUG)_{17}$ minigenes and performed and analyzed *in cellulo* assays using them (Fig. 4c, S6, 5). K.T. prepared a figure legend of Fig. 1c-e, S1, S2a-e, 2c, 3c, S3, S4a, b, 4a-c, 5, S5, S6 and materials and method section describing "Plasmid preparation Plasmid preparation" (with P.C.), "Cell culture and transfection" (with P.C.), "Preparation of radiolabeled RNA fragments", "Quantification of the MBNL1/RNA interaction and its inhibition by AONs *in vitro*", "Quantification of the SRSF1/RNA interaction and its inhibition by AONs *in vitro*" and "Statistical analysis".

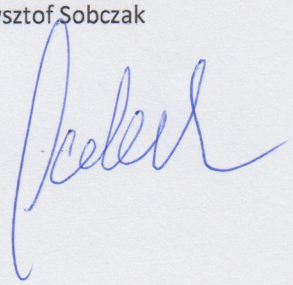
Ł.J.S. (Łukasz J. Sznajder) identified MBNL targets in *NASP* transcript using CLIP-seq data and designed 2'OMe-PS-NASP AON (used in Fig. 2a, c, d, 3c, d, S3a, c, S4b, S7).and prepared MBNL1 and 2 knock down in HSkM cells (used in Fig. 2b). Ł.J.S. drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 20.09. 2018, Poznań

Name: Krzysztof Sobczak

Signature:



Nucleic Acids Research Advance Access published October 12, 2016

doi: 10.1093/nar/gkw915

Mechanistic determinants of MBNL activity

Łukasz J. Sznajder[1,*], Michał Michalak[2], Katarzyna Taylor[1], Piotr Cywoniuk[1], Michał Kabza[3], Agnieszka Wojtkowiak-Szlachcic[1], Magdalena Matłoka[1], Patryk Konieczny[1] and Krzysztof Sobczak[1,*]

[1]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland, [2]Department of Molecular and Cellular Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland and [3]Department of Bioinformatics, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

Received September 19, 2016; Accepted October 05, 2016

ABSTRACT

Muscleblind-like (MBNL) proteins are critical RNA processing factors in development. MBNL activity is disrupted in the neuromuscular disease myotonic dystrophy type 1 (DM1), due to the instability of a non-coding microsatellite in the *DMPK* gene and the expression of CUG expansion (CUG^{exp}) RNAs. Pathogenic interactions between MBNL and CUG^{exp} RNA lead to the formation of nuclear complexes termed foci and prevent MBNL function in pre-mRNA processing. The existence of multiple MBNL genes, as well as multiple protein isoforms, raises the question of whether different MBNL proteins possess unique or redundant functions. To address this question, we coexpressed three *MBNL* paralogs in cells at equivalent levels and characterized both specific and redundant roles of these proteins in alternative splicing and RNA foci dynamics. When coexpressed in the same cells, MBNL1, MBNL2 and MBNL3 bind the same RNA motifs with different affinities. While MBNL1 demonstrated the highest splicing activity, MBNL3 showed the lowest. When forming RNA foci, MBNL1 is the most mobile paralog, while MBNL3 is rather static and the most densely packed on CUG^{exp} RNA. Therefore, our results demonstrate that MBNL paralogs and gene-specific isoforms possess inherent functional differences, an outcome that could be enlisted to improve therapeutic strategies for DM1.

INTRODUCTION

Cell development and fate is guided by a multitude of RNA binding proteins (RBPs) that affect the processing, localization, translation and turnover of RNAs. Tissue-specific proteome complexity arises from a relatively low number of about 20 000 protein coding genes due to the production of multiple mRNA isoforms generated by alternative splicing (AS) from >90% of protein coding genes (1). In many cases, the profile of alternative isoforms is modified during the course of tissue development and cell-specific transcriptome maturation is adjusted by RBPs functioning as *trans*-acting splicing factors. Proper AS depends on normal RBP function and abnormalities in the activity of some splicing factors lead to a number of human diseases (1).

Muscleblind-like (MBNL) proteins are conserved multifunctional RBPs which influence AS and alternative polyadenylation (APA), mRNA stability and trafficking as well as microRNA biogenesis (2–9). In mammals, there are three MBNL paralogs, MBNL1, MBNL2 and MBNL3 (10,11). All MBNL paralogs contain two N-terminal tandem zinc finger (ZnF) domains which bind preferentially to specific RNA sequences and/or structures containing two or more clustered GC steps flanked by pyrimidines (YGCY) (12–15). Each MBNL paralog may contain variable amino acid sequences encoded by alternative exons that modulate its cellular localization, the number of ZnF domains and the distance between them, multimerization capacity, affinity to RNA sequence motifs and AS activity (16–19). A comparison of MBNL activities might be reflected by splicing activities. However, AS events are affected by the total expression level of MBNL paralogs and the distribution of multiple splicing isoforms. Both depends on tissue

*To whom correspondence should be addressed. Tel: +48 61 829 5958; Fax: +48 61 829 5949; Email: ksobczak@amu.edu.pl
Correspondence may also be addressed to Łukasz J. Sznajder. Tel: +1 352 273 8155; Fax: +1 352 273 8284; Email: lukasz.j.sznajder@gmail.com
Present addresses:
Łukasz J. Sznajder, Department of Molecular Genetics and Microbiology, Center for NeuroGenetics and the Genetics Institute, University of Florida, College of Medicine, 2033 Mowry Road, Gainesville, FL 32610, USA.
Magdalena Matłoka, Sorbonne Universités, UPMC Univ Paris 06, INSERM UMRS974, CNRS FRE3617, Center for Research in Myology, Institut de Myologie, GH Pitié-Salpêtrière, F-75013 Paris, France.

type and developmental stage (4,6,10,20,21). Mutual down-regulation of MBNL1 and MBNL2 causes more prominent changes in the pattern of specific AS events and the selection of APA sites compared to the depletion of a single paralog (4,6,20). The analysis of MBNL3 is more difficult due to its low expression level in the majority of adult tissues (10). *Mbnl3* isoform knockout mice ($Mbnl3^{\Delta E2}$) show abnormalities in muscle regeneration and functionality but not in the AS pattern (18). In contrast, both *Mbnl1* and *Mbnl2* knockout mice ($Mbnl1^{\Delta E3/\Delta E3}$, $Mbnl2^{\Delta E2/\Delta E2}$) exhibit global AS changes mainly in muscles and brain, respectively (3,6,22,23). The activity of recombinant MBNL paralogs has been analyzed in cellular models but few studies have compared their splicing activity in the context of selected exogenous AS events (24–27). Nevertheless, the activities of the MBNL1, MBNL2, MBNL3 proteins and their isoforms have never been directly compared and understanding the differences between MBNL activities might shed light on their impact on RNA metabolism in normal and pathological stages.

In myotonic dystrophy type 1 (DM1) and type 2 (DM2), the three MBNL paralogs are specifically bound to, and sequestered by, expanded CUG (CUG^{exp}) and CCUG repeats ($CCUG^{exp}$), respectively (10,28,29). These toxic RNAs are composed of multiple UGCU that *in vitro* form stable hairpin structures and concentrate in nuclear foci in multiple cell types of DM patients (29–37). In living cells, these dynamic structures undergo formation and dispersion events that are affected by MBNL proteins (38,39) and other factors (40). Depletion of MBNL1 and MBNL2 decreases CUG^{exp} foci size and MBNL proteins bind to these structures (24,38,41). The reduction of free MBNL in the nucleoplasm leads to global AS and APA changes (3,4,6,20,22,23,42). The mouse model that overexpresses ∼250 CUG repeats (HSA^{LR}) shares more than 80% of the AS and APA abnormalities with the $Mbnl1^{\Delta E3/\Delta E3}$ model (3,4,22,43), but the nature of the interaction between CUG^{exp} RNA and MBNL proteins in nuclear foci remains unclear.

In this study, we compared the activity of MBNL paralogs and their splicing isoforms in different cellular models and discovered novel features of MBNL proteins responsible for their function. MBNL paralogs and splicing isoforms differed significantly in subcellular localization and they bound the same regulatory sequences in pre-mRNAs *in vitro* and *in vivo* evoking splicing changes of the same AS events with different strengths. MBNL paralogs also bound to toxic CUG^{exp} RNA with high affinity forming densely packed complexes and associated/dissociated from CUG^{exp} foci freely but to different extents. We also identified several factors that have an impact on both AS activity and CUG^{exp} foci formation, including MBNL sequences encoded by alternative exons.

MATERIALS AND METHODS

MBNL1, MBNL2 and MBNL3 constructs were generated using pEGFP-C1 and MB1-41, MB1-42, MB1-43 and MB1-N used in this study will be described elsewhere (Konieczny *et al.*, unpublished data). MB1-40 was produced by partial digestion and cloning of an EcoRI-Alw44I fragment from MB1-42 into MB1-41 while MB1-C was generated by the deletion of a region between MB1-41 HindII restriction sites. MBNL2 (MB2-38, MB2-39, MB2-40, MB2-41) and MBNL3 (MB3-37, MB3-39) isoforms were amplified from a human adult kidney and liver cDNA library, respectively (Human MTC Panel I, Clontech cat no. 636742) and amplified fragments were inserted into BamHI and EcoRI sites. The sequences of these constructs were verified by sequencing. For more information, see Supplementary Table S1 and Supplementary Material & Methods. The coding sequences of mCherry or Dendra2 were amplified and cloned between NheI-XhoI (MBNL1) and NheI-HindIII (MBNL2 and MBNL3) restriction sites instead of enhanced green fluorescent protein (EGFP or GFP). For generation of an EGFP-mCherry construct, an amplified mCherry sequence was inserted downstream of EGFP into HindIII-EcoRI sites. The final constructs were sequenced. All primers and polymerase chain reaction (PCR) conditions for Pfx50 high fidelity DNA polymerase amplifications (Invitrogen) are specified below. To generate recombinant MBNL1, MBNL2 and MBNL3 proteins, the construct for MBNL1 expression was described previously (44) while MBNL2 and MBNL3 were amplified (using primers listed in Supplementary Material & Methods) digested with BamHI and EcoRI, inserted into pGEX-6P-GST-His12 and the final constructs were sequenced. Purification of recombinant GST and His12-tagged MBNL1, MBNL2 and MBNL3 was performed as described (45) and protein concentration was measured using both the Bradford Assay and Sypro Ruby staining on 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis gels (Supplementary Figure S6A).

Minigenes

Human *TNNT2* ex.4 minigene and a DT960 construct encoding CUG^{exp} were a gift (Thomas Cooper, Baylor College of Medicine) and have been described previously (39,46). The mouse *Atp2a1* ex.22 minigene was prepared on the pEGFP-C1 background as previously described (47). The generation of mutated *Atp2a1*, *Nfix* also *Ldb3* minigenes will be described elsewhere (Cywoniuk *et al.*, unpublished data).

Cell cultures and transfection

HeLa cells were grown in Eagle's minimal essential medium (MEM) supplemented with 10% fetal bovine serum, 1× MEM non-essential amino acid solution (Sigma), 1× antibiotic and antimycotic (Sigma) at 37°C with 5% CO_2. Before transfection cells were seeded on 12-well plates filled with 1 ml of medium and allowed to reach up to 60-70% confluence. For endogenous splicing and protein expression analysis, HeLa cells were transfected with 2 μg of MBNL constructs using X-tremeGENE HP DNA Transfection Reagent (1:2 ratio, Roche) and harvested after 42 h. For exogenous splicing analysis, HeLa cells were co-transfected with 100 ng of splicing minigene, and 75, 150 and 300 ng of MBNL constructs, which were supplemented by empty pEGFP to the highest total DNA amount and cells were harvested after 24 h. In order to study pre-mRNA-MBNL interactions, HeLa cells were cotransfected with 500 ng of

MBNL constructs and 200 ng of the *Atp2a1* minigene. After 4 h, cells were transfected with 125 nM AONs (Supplementary Material & Methods). For both transfections, Lipofectamine 2000 (Invitrogen) was used. In fluorescent *in situ* hybridization (FISH) and fluorescence recovery after photobleaching (FRAP) experiments, HeLa cells were transfected with 200 ng of DT960 and 500 ng of MBNL constructs. For the Dendra2 analysis, the amount of MBNL1 reached up to 1 μg while for fluorescence-lifetime imaging microscopy (FLIM), 200 ng of DT960 and 375 ng of each GFP and mCherry fused to MBNL were used.

RNA isolation and RT-PCR

Total RNA from HeLa cells was isolated using TRI Reagent (Sigma) and total RNA (2 μg) was reverse transcribed using GoScript Reverse Transcription System (Promega) and random primers (Promega) according to the manufacturer's protocol. Samples transfected with splicing minigenes were treated with RQ1 DNase (Promega). For human tissues, Human MTC Panel I and Human Fetal MTC Panel (Clontech cat no. 636742 and 636747) were while muscle samples were a gift (Charles Thornton, University of Rochester). All PCRs were conducted using GoTaq Flexi DNA Polymerase (Promega) and detailed PCR conditions are described in Supplementary Table S2. PCR products were resolved on agarose gels with USB dye (Syngene) and gels were visualized on G:BOX and analyzed using GeneTools software (Syngene).

Western blotting

HeLa cells were lysed with RIPA buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5% NP-40, 0.5% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate (SDS)) supplemented with SIGMAFAST Protease Inhibitor Cocktail (Sigma). Lysates were sonicated at 4°C and centrifuged at 18 000 × *g* at 4°C for 10 min. Samples were heated to 95°C for 5 min, separated on 10% SDS polyacrylamide gels and transferred to nitrocellulose (Protran BA 85, Whatman) using a wet transfer apparatus (1 h, 100 V, 4°C). Membranes were blocked for 1 h with 5% skim milk in PBST buffer (phosphate buffered saline (PBS), 0.1% Tween-20) and incubated with a primary antibody against GFP (1:1000, Santa Cruz cat. no. sc-8334) or GAPDH (1:1000, Santa Cruz cat. no. sc-47724). Anti-rabbit (1:20 000, Sigma cat. no. A9169) and anti-mouse (1:2000, Millipore cat. no. 12-349) secondary antibodies were conjugated with horseradish peroxidase and detected using the Pierce ECL Plus Western Blotting Substrate (Thermo Scientific) detection kit.

CLIP-seq and RIP-seq

Crosslinking and immunoprecipitation (CLIP) experiments were performed on mouse skeletal muscles and hearts as well as transiently transfected C2C12 cells. RIP-seq experiments were performed on skeletal muscles. Step-by-step protocols and the MBNL Interactome Browser (MIB.amu.edu.pl) is described in Supplementary Material & Methods.

Microscopy

Confocal microscopy experiments were performed within 48 h after HeLa cell transfection. For live cell imaging, time-lapse sequences, FRAP, photoconversion of Dendra2 fluorescent protein and FLIM cells were seated on 96-well glass-bottom plates (Greiner Bio-One, cat no. 82051-531). Images were processed with Imaris software (intensity thresholds, gamma correction, scale bars) and presented as maximum intensity projection from all z slices.

Live-cell imaging was performed on Nikon A1Rsi confocal microscope with objectives Nikon Plan Apo VC 60x/1.4 Oil DIC N2 (FRAP, dual-channel imaging, FLIM) and Nikon Plan Apo 40x/0.95 DIC N2 (time-lapse acquisition). GFP and mCherry fluorescent proteins were excited with 488 nm Argon-Ion and 561 nm pumped-diode laser, respectively. For detection dichroic mirror 405/488/561 nm with spectral filters 525/50 nm (green channel) and 595/50 nm (red channel) were used. To avoid spectral bleed-through channels were scanned sequentially. For each cell several optical slices covering majority of the cell volume were acquired. Time-lapse sequences were acquired through 12-16 h with 2-5 min step between scans. Z-stacks were covering cells and space above and below to avoid movement of the cells to out of focus planes.

FRAP was performed on single focal plane using a sequence of 5 pre-bleach, 2 bleach and 294 post-bleach image acquisition (overall experiment time set to 5 min). A few GFP foci were bleached in each nucleus using 488 nm laser at full power and fluorescence intensities were measured by confocal PMT detector. Results were analyzed with usage easyFRAP software (48). For analysis, the intensities of bleached region of interests (ROIs) from single nucleus were averaged and the entire nucleus was set as reference ROI and the extracellular medium was used as background. The raw data were normalized by full scale method. The curve was fitting to single exponential equation to calculate the mobile fraction and t-half. Statistical significance was determined by Mann-Whitney U test; $*$ for $P < 0.05$, $**$ for $P < 0.01$ and $***$ for $P < 0.001$.

Photoconversion of the Dendra2 fluorescent protein was done using a 405 nm diode laser set to 12% power, and each ROI was bleached for 66 ms. Image acquisition was performed on single focal plane through 5 min with 10 s interval and 2 pre-bleach/conversion images. Results were normalized by full scale method in Excel software.

FLIM measurements were performed using the PicoQuant LSM Upgrade Kit. Fluorescence resonance energy transfer (FRET) donor molecule (GFP) was excited by the 485 nm pulsed-diode laser at a 40 MHz repetition rate. Photons were counted by single-photon avalanche diode (SPAD) detector with 520/35 nm spectral filter and 256 × 256 pixel FLIM images were acquired as long as average number of photon counts per pixel reached at least 100. FLIM analysis was performed with SymPhoTime64 software from PicoQuant. Statistical significance was determined by Student's *t*-test; $*$ for $P < 0.05$, $**$ for $P < 0.01$ and $***$ for $P < 0.001$.

Fluorescence *in situ* hybridization

FISH was conducted as described (49) with some modifications. Cells were fixed in 2% paraformaldehyde (PFA)/PBS at 4°C for 5 min and washed three times in cold PBS. Pre-hybridization was performed in 30% formamide and 2× saline-sodium citrate (SSC) buffer at 4°C for 10 min, followed by hybridization in buffer containing 30% formamide, 2× SSC, 0.02% BSA, 66 μg/ml yeast tRNA, 10% dextran sulfate, 10 U Rnasin (Promega) and 2 ng/μl DNA/locked nucleic acid (LNA) probes $(CAG)_6$-CA, labeled at the 5′-end with Cy3 and modified at positions 2, 5, 8, 13, 16 and 19 with LNA. Post-hybridization washing was done in 30% formamide and 2× SSC at 37°C for 30 min followed by 1× SSC at 37°C for the next 30 min. Microscopic slides were mounted using medium containing 2% propyl gallate (Sigma), 10% glycerol and 4′,6-diamidino-2-phenylindole (DAPI) and sealed with fingernail polish. All FISH images were acquired on Nikon A1Rsi microscope using the following excitation conditions: diode lasers 405, 488 and 561 nm, dichroic mirror 405/488/561, emission filters 450/50 for DAPI, 525/50 for GFP and 595/50 for Cy-3. All images were acquired with sequential scanning to avoid spectral bleed-through. A total of 18-25 optical sections were acquired. Foci volume was estimated by Imaris. Statistical significance was determined in GraphPad Prism software by Mann-Whitney U test; ∗ for $P < 0.05$, ∗∗ for $P < 0.01$ and ∗∗∗ for $P < 0.001$. The Spearman's rank correlation coefficient was calculated in Python (scipy.ststs).

Transcription *in vitro* and radiolabeling

DNA templates for *Atp2a1*, *Mbnl1*, *Mbnl2*, *Tnnt3*, *Calm3, Clcn1* and *Mfn* RNAs were obtained in two PCR reactions. First, longer products were amplified using genomic DNA and primer sets (see Supplementary Material & Methods) at the indicated annealing temperature (Ta) which were subsequently used as template for a second PCR reaction carried out at 55°C and utilizing primer sets (see Supplementary Material & Methods) with a 5′ promoter sequence for T7 RNA polymerase (Ambion). The only exception was *Insr*, which was amplified from a synthesized DNA template using specific primers. Double-stranded (CCAG)14 DNA was prepared as described (35) using a GGCCC(CAGG)14GGGCCTATAGTGAGTCGTATTA ssDNA and a T7 oligomer TAATACGACTCACTATAGG. The transcription reaction was performed in 50 μl composed of 10 μl of DNA template, 0.15 mM nucleoside triphosphates (NTPs) (Invitrogen) 0.45 mM guanosine (Sigma-Aldrich), 50 U T7 RNA Polymerase (Ambion), 1× T7 transcription buffer (Ambion), 40 U Rnasin Plus RNase Inhibitor (Promega). Purification of the transcript was conducted on a denaturing 6% polyacrylamide gel (19:1 acrylamide:bisacrylamide) followed by ethanol precipitation. For radiolabeling, 2-4 pmol of transcript was incubated with 2-4 pmol of [γ-^{32}P]ATP, 1 U Rnasin Plus, 10 U OptiKinase (Affymetrix), 1× reaction buffer (Affymetrix) and ddH_20 up to 10 μl, at 37°C for 1 h. Labeled RNA was subsequently run on a denaturing 8% polyacrylamide gel (19:1) in 0.5× Tris-borate-EDTA (TBE) buffer, at 100 V for 1 h, the RNA was visualized using FLA-5100 (FujiFilm) and the RNA was cut out followed by ethanol precipitation and resuspended in 20-40 μl ddH_2O. (CUG)20 and (CAG)20 were a gift (Włodzimierz Krzyżosiak, Polish Academy of Sciences).

Quantification of RNA-MBNL interaction and its inhibition by AONs *in vitro*

Filter binding assay was performed in 30 μl volume. To assess the MBNL1, MBNL2 and MBNL3 affinity to RNAs, 5′-labeled transcripts (0.05 nM) were incubated with the indicated concentrations of the proteins (ranging from 0 to 250 nM) in buffer B containing 250 mM NaCl, 15 mM KCl, 50 mM, Tris-HCl pH 8.0, 0.05% Tween-20, 1 mM $MgCl_2$ at 37°C for 25 min. DNA template preparation and *in vitro* transcription were described in Supplementary Material & Methods. To estimate the inhibitory property of AONs, 0.05 nM of labeled transcripts, incubation underwent a three step with 20 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 MBNLs were added to each sample and incubated at 37°C for 25 min. A total of 25 μl of samples were loaded onto filter binding apparatus with nitrocellulose (Protran BA 85, Whatman) and nylon (Hybond N+, Amersham) membranes previously wetted in buffer B. The signal from membranes was visualized on IP through FLA-5100 and quantified using Multi Gauge software (FujiFilm). K_d values were calculated in GraphPad Prism based on two experimental replicas, using the equation for one site-specific binding and standard error of the mean. Other statistical analyses were performed with GraphPad Prism, Python (scipy.stats) and Microsoft Excel software and significance was determined by the appropriate parametric or nonparametric statistical test.

RESULTS

Platform to evaluate the activities of MBNL paralogs and their splicing isoforms

The main aim of our project was to explain whether three protein paralogs with very similar primary structure have similar or distinguishable activity in the same conditions? The *MBNL1*, *MBNL2* and *MBNL3* genes have various expression profiles in different tissue types and developmental stages (11,50) and there are several isoforms of individual *MBNL* paralogs containing different combinations of a few alternatively spliced exons (Supplementary Figure S1A). Thus, to directly compare the activity of the three MBNL paralogs and their spliced isoforms, we generated and transiently overexpressed 10 different proteins in HeLa cells in which the expression of endogenous MBNL isoforms is relatively low (Supplementary Figure S1B). All experiments were conducted within 48 h post-transfection using four isoforms of MBNL1 (MB1-40, MB1-41, MB1-42 and MB1-43 kDa), four isoforms of MBNL2 (MB2-38, MB2- 39, MB2-40 and MB2-41 kDa) as well as two isoforms of MBNL3 (MB3-37 and MB3-39 kDa) (Figure 1A and Supplementary Table S1). All of them represent the most common protein variants detected in different fetal and adult tissues (50). The similarity of amino acid sequences in the core region of MBNL paralogs is 65-73% and the highest between MBNL1 and MBNL2 (Supplementary Figure S1C).

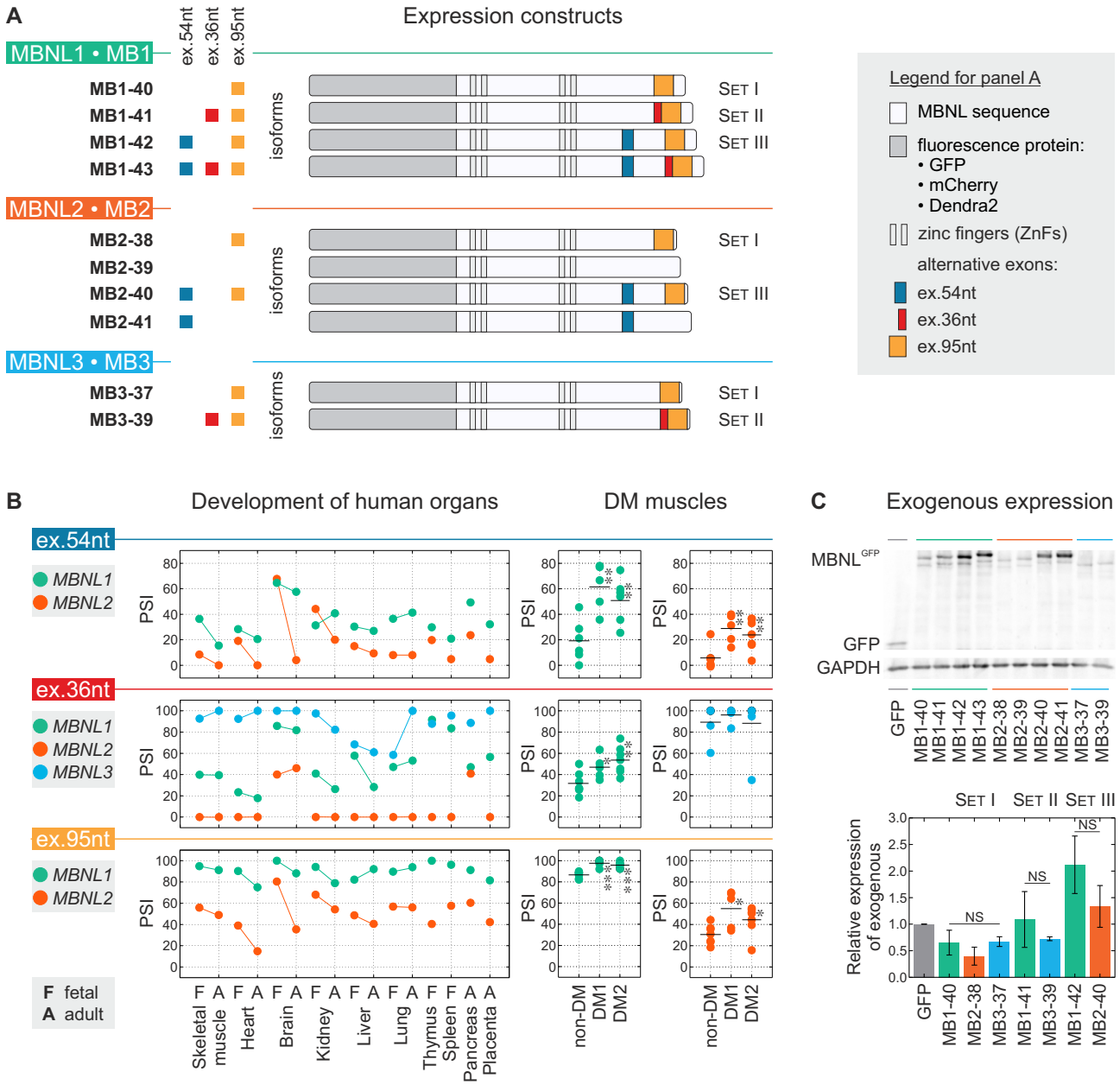


Figure 1. Distinct splicing and expression patterns for *MBNL* paralogs**.** (**A**) Scheme of 10 expression constructs containing sequences of fluorescence proteins (GFP, mCherry or Dendra2) and MBNL1, MBNL2 and MBNL3 paralogs possessing or lacking three alternative exons (ex.54nt, ex.36nt, ex.95nt). For details on MBNL exonic composition see Supplementary Table S1. Constructs for comparable isoforms having the same alternative exons are divided into SET I, SET II and SET III. (**B**) Splicing profiles for MBNL ex.54nt, ex.36nt and ex.95nt in human tissues analyzed by RT-PCR. Dots represent PSI values. The adult and fetal samples from particular tissue types are indicated by a bar. Splicing results obtained for non-DM (N = 7), DM1 (N = 5) and DM2 (N = 7) muscle samples. Statistical significance was determined by the Student's *t*-test (∗ for $P < 0.05$, ∗∗ for $P < 0.01$ and ∗∗∗ for $P < 0.001$). (**C**) Relative expression level of exogenous MBNL paralogs. Exogenous proteins were detected by anti-GFP antibody and normalized to GAPDH. Bars represent average expression level and standard deviations are from two independent biological experiments. Statistical significance was determined by the Student's *t*-test (NS for $P \geq 0.05$).

The variance in amino acid sequences of different isoforms of MBNL paralogs depends on the presence or absence of three highly conserved alternative exons which consist of 54, 36 and 95 nucleotides, herein called ex.54nt, ex.36nt and ex.95nt (previously numbered ex.5, ex.7 and ex.8, respectively) (2) (Figure 1A and Supplementary Figure S1A). The protein sequences encoded by MBNL1 and MBNL2 ex.54nt, which contains a nuclear localization signals, are 72% identical whereas no such sequence exists in MBNL3 (Supplementary Figure S1A) (16,17,19). The amino acid sequences encoded by ex.36nt and ex.95nt are 60-82% and 81-87% identical between the three paralogs, respectively, and form a C-terminus of the protein. Ex.36nt is believed to be responsible for homotypic MBNL1-MBNL1 dimerization (16,51).

RT-PCR analysis revealed that the splicing of ex.54nt, ex.36nt and ex.95nt undergoes significant changes during organ development (Figure 1B). As a measure of AS changes, we used the percent-spliced-in (PSI) parameter which indicates the mRNA fraction that includes the specific alternative sequence. Interestingly, all these exons are also significantly misspliced in both DM1 and DM2 in which the pool of functional MBNL protein is reduced due to sequestration on toxic RNAs containing CUG^{exp} or $CCUG^{exp}$, respectively (Figure 1B). Cumulatively, these data indicate that during tissue development, and in some pathological stages, the quantity and/or quality of *MBNL1*, *MBNL2* and *MBNL3* expression undergo significant regulation.

All tested MBNL isoforms were C-terminal fusions to fluorescent proteins (Figure 1A). The effectiveness of exogenous protein production in HeLa cells was assessed using western blotting with an anti-GFP antibody and normalization to GAPDH. We noticed that the presence of alternative ex.54nt always significantly elevated MBNL1 and MBNL2 protein levels (∼2.7 and ∼4.0-times, respectively). We observed marginal changes of the GFP-MBNL level for constructs carrying ex.36nt (∼1.3-times) and ex.95nt (∼0.8-times) (Figure 1C and Supplementary Figure S1D). We also compared the expression level of MBNL paralogs in three sibling splicing isoform sets differing in the presence or absence of one of three alternative exons. SET I consisted of MB1-40, MB2-38 and MB3-37 having only ex.95nt, SET II consisted of MB1-41 and MB3-39 having both ex.36nt and ex.95nt while SET III contained MB1-42 and MB2-40 having ex.54nt and ex.95nt (Figure 1A). Differences in the expression level of GFP fused MBNLs within these groups were very low (Figure 1C).

Cumulatively, these data indicate that this cellular model provides a reliable platform to evaluate the activities of MBNL paralogs and their alternatively spliced isoforms. Next, we focused on MBNL splicing activity in normal and pathological states.

MBNL1 possesses the strongest, while MBNL3 the weakest, splicing activity

To study the impact of MBNL paralogs on AS regulation for individual AS events and globally for the transcriptome, we first selected 38 AS events reported previously to be MBNL-sensitive and analyzed them in HeLa cells using RT-PCR (Supplementary Table S2). To standardize alternative exon enumeration, we applied the numbering system from FasterDB (52). We observed that 10 MBNL isoforms regulate selected individual AS events always in the same direction, either toward exon inclusion (herein termed exON) or exclusion (exOFF) (Supplementary Figure S3A). However, we noticed that taking all 38 AS events into consideration, the average strength of exOFF was significantly higher than exON for all MBNL isoforms (Figure 2A). We decided to confirm this observation based on the splicing changes of hundreds of AS events selected from other available data sets (6,22,53). MBNL silencing in C2C12 myoblasts and mouse *Mbnl* knockout muscles induced a stronger effect on exOFFs (Figure 2A; Supplementary Figure S3C and D). Even though $C(C)UG^{exp}$ RNAs may perturb the activity of several splicing factors, the same phenomenon was also observed in skeletal muscles of the HSA^{LR} mouse DM1 model as well as DM2 patients (Figure 2A and Supplementary Figure S3E).

Subsequently, we looked at differences between the activities of the three MBNL paralogs within comparable groups of SET I, SET II and SET III having the same alternative exons (Figure 1A). Surprisingly, we discerned that for combined splicing changes of all 38 AS events, MBNL1 had the strongest, whereas MBNL3 had the weakest, splicing activity (Figure 2B). MBNL1 showed ∼28% more splicing changes compared to MBNL3 and 30 of 38 tested AS events fit this rule while eight (21%) AS events did not, including *ATP2A1* ex.23 and *NASP* ex.7 that showed similar PSI values for each MBNL group (Supplementary Figure S2). To validate this observation, we coexpressed three SETs of paralogs with two splicing minigenes, mouse *Atp2a1* and human *TNNT2,* containing MBNL regulated exons. For the *Atp2a1* minigene, the level of ex.22 (orthologue of human *ATP2A1* ex.23) inclusion was always the highest for MBNL1, and the weakest for MBNL3, isoforms (Figure 2C). Similar results were obtained for the exclusion of *TNNT2* ex.4 although contrary to *Atp2a1*, MBNL1 and MBNL2 isoforms induced similar splicing changes. These results demonstrated that MBNL paralogs influence pre-mRNA splicing to varying degrees and MBNL1 has the strongest alternative processing whereas MBNL3 has the weakest. These results suggest that even relatively small differences in primary structures of MBNL paralogs influence their activities. MBNL proteins regulate hundreds of AS events (4,6,23) thus even moderate differences in their activity might have a cumulative and significant impact.

MBNL paralogs regulate same AS events differentially binding the same motifs

Potential explanations for the observed differences in AS regulation between the tested MBNL paralogs are target RNA binding properties-binding sites and affinity. To test this hypothesis, we first ascertained whether all three MBNL paralogs bind to the same or diverse RNA binding sites. To answer this question, global potential binding sites for the three MBNL paralogs were identified by combining all CLIP experiments followed by deep sequencing (CLIP-seq) results coming from our own (Supplementary Figure S5) as well as previously published datasets for MBNL1 (4,6,8), MBNL2 (4,20,23) and MBNL3 (4,18). The total number of mapped CLIP-seq unique reads amounted to about 1.0, 0.1 and 0.8 million for MBNL1, MBNL2 and MBNL3, respectively, and we developed the MBNL Interactome Browser (MIB) (available at MIB.amu.edu.pl) to analyze the localization of CLIP-seq reads specific for MBNL1 and MBNL3 within alternative exons or in neighboring introns (±250 nt) of 38 AS events (results for MBNL2 were underrepresented). Approximately half of the reads overlapped between the MBNL1 and MBNL3 CLIP-seq data (Figure 3A and Supplementary Table S2), which suggests that splicing regulation requires binding to the same RNA target sites. We also noticed that the consensus sequences for all MBNL proteins are very similar and include 1-2 YGCY/A motifs (Figure 3B).

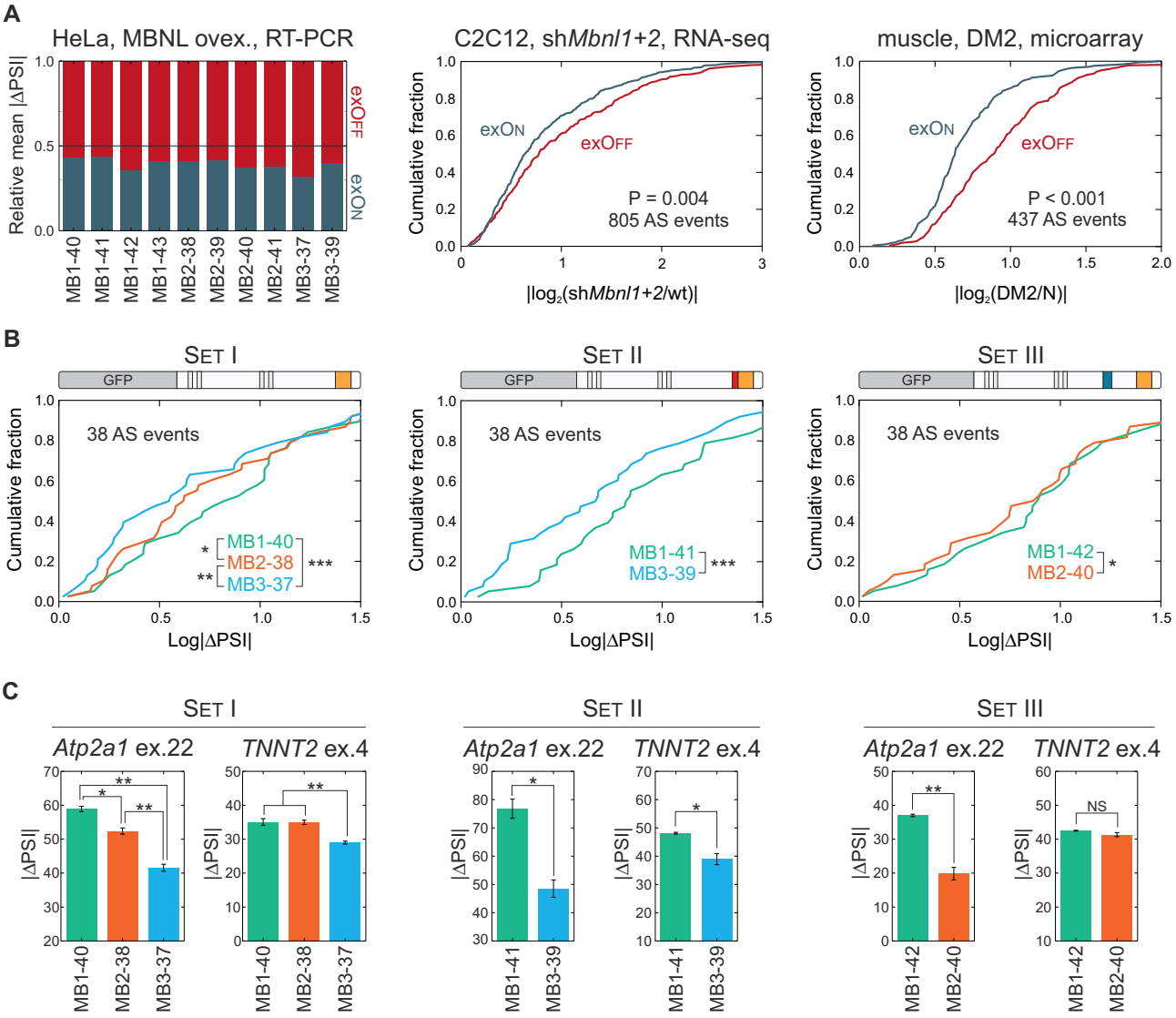


Figure 2. MBNL1 is the strongest, while MBNL3 is the weakest, AS regulator. (**A**) Comparison of the relative mean for exON/exOFF splicing changes (|ΔPSI|) of 38 endogenous AS events in HeLa cells transfected with 10 different MBNL isoforms (left chart). Global analysis of exON and exOFF changes in C2C12 with silencing of MBNL determined by RNA-seq (central chart) (6) and DM2 skeletal muscles analyzed by microarrays (right chart) (53). For all analyses, the fold change was used as a measure of splicing strength for the indicated number of AS events. The statistical significance was determined by the Mann-Whitney U test. All comparisons revealed higher average changes for exOFF than exON. For more examples see Supplementary Figure S3. (**B**) Combined analysis of splicing changes (expressed as |ΔPSI|) of 38 tested AS events induced by MBNL paralogs from SET I, SET II and SET III. Statistical significance was assessed by Wilcoxon signed rank test (∗ for P <0.05, ∗∗ for P < 0.01 and ∗∗∗ for P < 0.001). (**C**) Differences in the regulation of exogenous mouse *Atp2a1* ex.22 and human *TNNT2* ex.4 by MBNL paralogs. The represented mean and standard deviation come from two biological replicas. Statistical significance was determined by the Student's t-test (∗ for P < 0.05, ∗∗ for P < 0.01 and ∗∗∗ for P < 0.001). For MBNL dose dependent changes, see Supplementary Figure S4.

To check whether all MBNL paralogs efficiently bind to the same RNA molecules with the similar or different affinity we compared the binding affinity of MBNL proteins to short synthetic RNA fragments, which were known from previous studies to be MBNL1 targets (50). Using the *in vitro* assay we determined the dissociation constant (K_d) values for interaction between MBNL paralogs and 10 RNAs and found 3-fold differences in binding affinity. However, for some RNAs, including *Mbnl1* exon 3 and (CUG)20 repeats, all MBNL paralogs showed similar, <2-fold K_d fold change values (Figure 3C and Supplementary Figure S6C-E).

To confirm that all MBNL paralogs interact with exactly the same binding sites, we utilized *Atp2a1* intron 22 and *Mbnl1* exon 3 RNA fragments and identified the YGCY sequences, which are predicted to be MBNL recognition motifs. Having designed antisense oligonucleotides (AONs) complementary to selected YGCY motifs, we noticed that for RNA fragments bound to these specific AONs the K_d was at least 100-times greater for all MBNL paralogs (Figure 3D; Supplementary Figure S6B and C). Next, we prepared two sets of *Calm3* 3′UTR and *Mbnl2* intron 8/exon 9 RNA fragments with point substitutions into 1-3 YGCY motifs. Based on *in vitro* experiments we concluded that

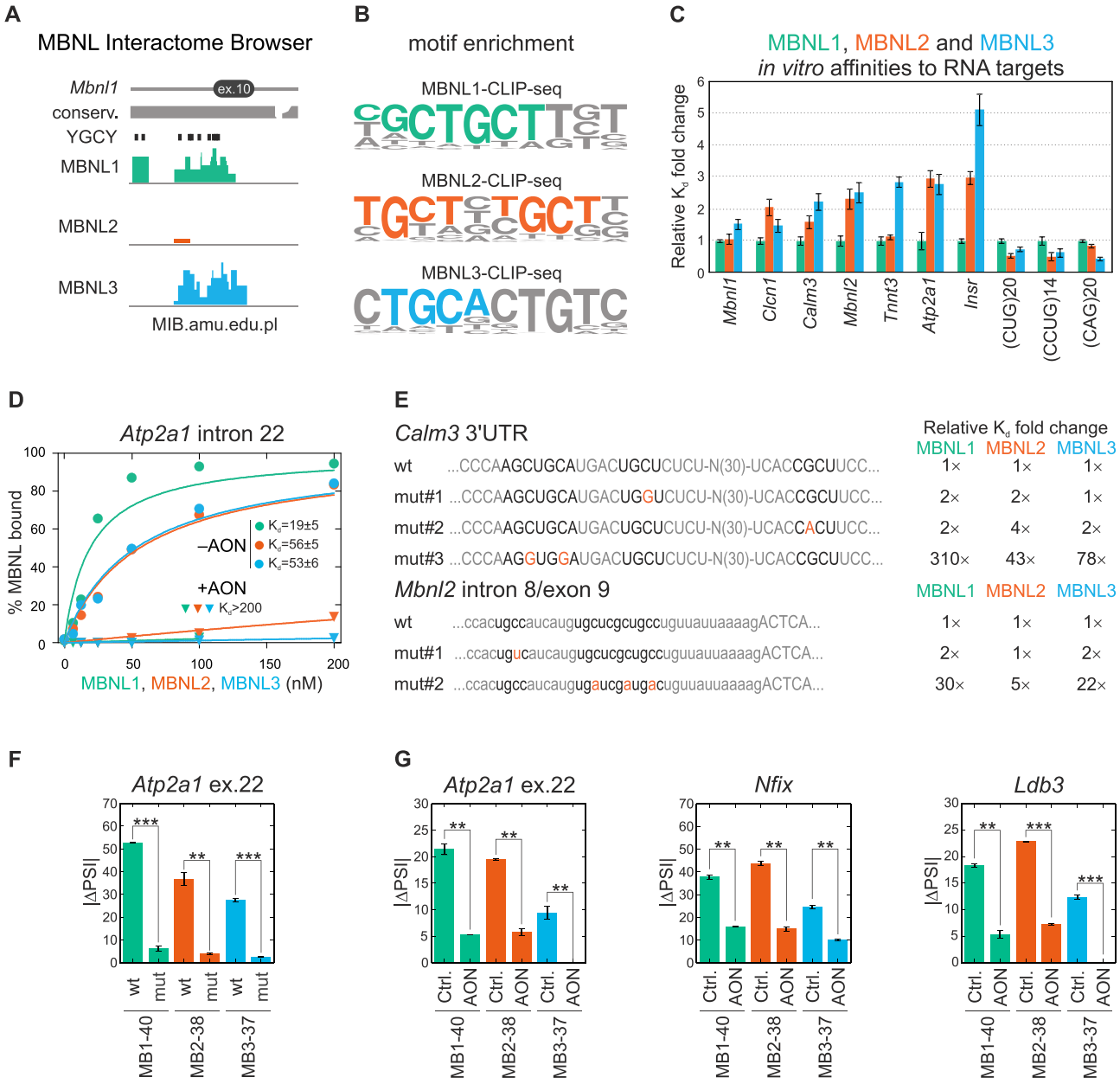


Figure 3. MBNL binding to RNA targets. (**A**) An example of overlapping MBNL1, MBNL2 and MBNL3 specific CLIP-seq clusters located in the region of intron 9 and alternative ex.10 (known also as ex.5; Supplementary Figure S1A) of mouse *Mbnl1* based on the MBNL Interactome Browser (MIB.amu.edu.pl). For additional AS examples, see Supplementary Table S2. The presence of overlaid reads was detected for 81, 21 and 71% of AS events for MBNL1, MBNL2 and MBNL3, respectively. All YGCY sequence motifs located in this region, as well as evolutionary conservation are also shown. Note that Multi MIB also contains crosslinking induced mutations (CIMs), MBNL1-RIP-seq and CLIP-seq results for CELF1 (8) and NOVA (58) proteins. (**B**) A 10-mer sequence motif enrichment analysis from MBNL1, MBNL2 and MBNL3 CLIP-seq data determined by HOMER software. (**C**) The comparison of binding affinity of MBNL paralogs for 10 RNA fragments, indicating the strongest binding properties of MBNL1 and about 3-fold weaker of the other paralogs. However, in three examined instances, the binding affinity of all paralogs was alike. The dissociation constants (K_d) for MBNL1 constituted the normalization values. (**D**) MBNL paralogs have slightly different affinities for the 145 nt RNA fragment of *Atp2a1* intron 22 *in vitro*. The K_d values significantly increase after blocking the MBNL1-specific binding site by 21nt-long antisense oligonucleotide (+AON) compared to control experiments (−AON). Dots represent the average percent of MBNL bound ± SD from two technical repetitions. For *Mbnl1* exon 3, see Supplementary Figure S6C. For more details, see Supplementary Material & Methods. (**E**) Point mutations in crucial YGCY motifs compromise MBNL1, MBNL2 and MBNL3 binding efficiency *in vitro*. (Left panel) Nucleotide sequences of wild-type (wt) and mutated (mut#1-3) *Calm3* 3′UTR fragment and *Mbnl2* intron 8/exon9 RNA fragment containing several YGCY motifs. (Right panel) Fold change of K_d values normalized to wt samples. (**F**) Removing the 111 nt sequence having the MBNL1-binding site in the *Atp2a1* minigene (mut) also vastly affects MBNL2 and MBNL3 activity. (**G**) Blocking the MBNL-specific binding site in intron 22 of *Atp2a1* minigene using AON reduces ex.22 inclusion for three MBNL paralogs in HeLa cells compared to the control AON (Ctrl.). Analogs results are represented for AONs designed against two other MBNL-binding sequences in *Nfix* and *Ldb3*. (F and G) Bars represent average PSI ± SD from two independent experiments and statistical significance was determined by the Student's t-test (∗ for $P < 0.05$, ∗∗ for $P < 0.01$ and ∗∗∗ for $P < 0.001$).

the binding affinity of each MBNL paralog is compromised by mutation in the same crucial binding motifs mut#3 and mut#2 for *Calm3* and *Mbnl2*, respectively (Figure 3E). On the other hand affinity was not affected by mutations in other YGCY motifs.

Next, we asked whether MBNL paralogs regulate AS events interacting with the same YGCY motifs in living cells. To answer this question, we deleted 111 base pair sequence encoding a previously defined MBNL1 binding site in the intron 22 of *Atp2a1* minigene and we cotransfected it with MBNL paralogs. As we expected, neither MBNL protein was able to efficiently regulate ex.22 splicing compared to the wild-type (wt) minigene (Figure 3F). In the next experiment, we cotransfected the wt *Atp2a1* minigene also with the same AONs as used *in vitro*. For three tested paralogs, identical and strong inhibitory effects of these AONs on ex.22 inclusion was observed (Figure 3G). Similar results were obtained for AONs designed against two other MBNL1-binding sequences from *Nfix* and *Ldb3* (Figure 3G).

These results indicate that a few determinants influence MBNL AS activity. MBNL paralogs recognize the same sequence motifs but the differences in splicing activity observed between paralogs are influenced by different affinity to RNA target, probably due to distinct sequence or structure sensitivity.

MBNL paralogs splicing activities correlate well with their subcellular localization

Subsequently, we addressed the question what is the the effective concentration of MBNL proteins in the nucleus. Firstly, we checked if there were any differences in subcellular localization of the studied proteins using quantitative confocal microscopy on HeLa cells transfected with three MBNL paralogs and their splicing isoforms. Surprisingly, we observed a significantly different distribution of the fluorescence signal between cytoplasm and nucleoplasm for cells transfected with different GFP or mCherry tagged MBNL constructs. The fluorescence intensity was measured from total volumes of ∼100 cells for each construct and the nuclear signal was normalized to the total cellular signal (Figure 4A and Supplementary Figure S7A). The comparison of paralogs from SET I and SET II revealed that MBNL1 predominates in nucleoplasm (∼72%), MBNL3 was primarily cytoplasmic (only 36% in nucleoplasm) and MBNL2 showed an intermediate localization pattern (∼55% in nucleoplasm). As we expected, for SET III, the presence of the amino acid sequence encoded by ex.54nt influenced the localization of MBNL proteins in the nucleus (16–17,19). We confirmed these observations with coexpression of MBNL isoform pairs fused with mCherry and GFP (Figure 4B and Supplementary Figure S7B). Interestingly, we also observed reduced nucleoplasmic distribution of both C-truncated (MB1-C) and N-truncated (MB1-N) MBNL1 proteins (Supplementary Figure S8).

Then, we investigated whether the different strength of AS changes reflected the expression level of MBNL paralogs. The Pearson correlation coefficient between the mean log|ΔPSI|, which reflected the strength of splicing changes, and the total MBNL level was moderate ($r = 0.59$). However, when the differences in localization pattern of the MBNL paralogs was assessed, a much stronger correlation was achieved for comparing log|ΔPSI| versus the MBNL nuclear concentration ($r = 0.91$) (Figure 4C).

These results indicate that MBNL nuclear concentration influences the MBNL-mediated regulation of AS. It is worthy to mention that instead of this phenomenon some AS events were equally regulated by MBNL paralogs (Supplementary Figure S2).

Alternative MBNL exons modulate splicing activity

Based on AS results, we also noted significant differences in the activity of splicing isoforms of individual MBNL isoforms (Figure 5A and Supplementary Figure S2). In the next step, we asked the question whether the presence of amino acid sequences encoded by ex.54nt, ex.36nt and ex.95nt influence MBNL splicing regulatory properties. We compared splicing changes of 38 AS events for the comparable pairs of MBNL paralogs differing in one alternative exon (Figure 5B). In spite of the significantly higher expression of MBNL1 and MBNL2 isoforms possessing ex.54nt and their exclusive nuclear localization, we observed a similar or even weaker effect on splicing regulation for the majority of AS events (∼67%) compared to isoforms lacking ex.54nt (Figure 5B). For all AS events with significant differences in splicing patterns affected by the presence or absence of ex.54nt, there was a relationship between the MBNL mechanism of action (exON or exOFF) and the strength of splicing changes. For AS events with a negative effect of ex.54nt, the exON events predominated (59%, $P = 0.058$) (Figure 5B). To confirm our observation, we coexpressed mouse *Atp2a1* and human *TNNT2* minigenes undergoing exON and exOFF, respectively, with comparable MBNL isoforms. In agreement with our previous observation, the presence of ex.54nt sequence induced a negative effect on exON of *Atp2a1* but a positive effect on exOFF of *TNNT2* (Figure 5C).

A similar analysis was performed for MBNL isoforms containing or lacking ex.36nt or ex.95nt. There were no significant differences between the activities of the analyzed proteins for the majority (67%) of AS events. However, these exons might have either a positive or negative effect on some specific AS events (Figure 5A and C). Interestingly, we observed that exOFF events predominated in AS events with a negative effect of ex.36nt (95%; $P = 0.002$) and a positive effect of ex.95nt (90%; $P = 0.010$) (Figure 5B). All together, these data indicate that the presence of sequences encoded by these alternative exons can significantly modulate MBNL splicing activity, but this regulation depends strongly on the targeted RNA. The majority of tested AS events was altered by at least one isoform (Figure 5A and Supplementary Figure S2) and thus various isoforms can either positively or negatively regulate different AS events dependent on the MBNL mode of action (exON and exOFF mechanism).

To further investigate the functions of MBNL1 isoforms, we conducted CLIP-seq experiments for overexpressed MB1-40, MB1-41 and MB1-43 (Figure 1A) that differed in cellular localization in C2C12 myoblasts (Supplementary Figure S5). Since MBNL proteins function in

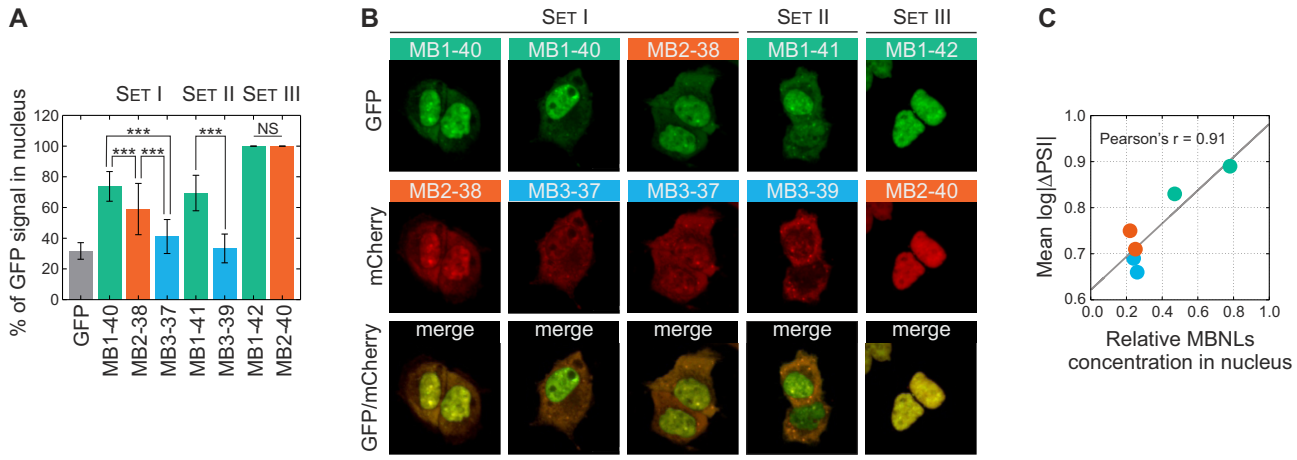


Figure 4. Distinct localization patterns for MBNL paralogs. (**A**) The percentage of nuclear signal for GFP-MBNL fusions belonging to Set I, Set II and Set III determined by quantitative confocal microscopy analysis of the GFP fluorescence signal. Bars represent the mean from the whole volume of about 50 cells. Statistical significance was assessed by the Student's *t*-test (NS for $P \geq 0.05$ and *** for $P < 0.001$). (**B**) Distinct nucleoplasmic and cytoplasmic distribution patterns for isoform pairs visualized by coexpression of MBNL proteins fused with GFP and mCherry (Set I and Set II). Isoforms having ex.54nt localize exclusively to the nucleus (Set III). (**C**) The Pearson correlation coefficient of relative nuclear concentration (normalized MBNLs expression level × relative nucleoplasmic distribution) with mean strength of 38 AS event changes (log|ΔPSI|) for six MBNL paralogs without alternative ex.54nt.

both the nuclear and cytoplasmic compartments (6), we expected that the frequency of CLIP-reads for MBNL1 isoforms without ex.54nt (MB1-40 and MB1-41), which localize in both cytoplasm and the nucleus, should be higher for exonic sequences (UTRs and protein coding regions) compared to the nuclear MB1-43 isoform containing ex.54nt. Indeed, looking globally we found more reads within mature mRNA elements for both MBNL1 isoforms with nuclear and cytoplasmic localization (Figure 5D).

MBNL paralogs are crowded in CUG^{exp} RNA foci

In DM1 and DM2, MBNL proteins are sequestered by transcripts containing expanded CUG^{exp} and CCUG^{exp} repeats, which aggregate in discrete nuclear foci *in vivo* (50,54). Therefore, we tested, whether different MBNL paralogs and their isoforms are capable of interacting with CUG^{exp} and CCUG^{exp} RNA and forming nuclear foci in a similar manner. Using the *in vitro* binding assay, we observed a very high affinity for all MBNL paralogs to non-pathogenic (CUG)20 and (CCUG)14 repeat transcripts (K_d in 4-15 nM range; the lowest among 10 tested RNA molecules) with just slight differences between paralogs (Figure 6A and Supplementary Figure S6E). Next, we coexpressed different GFP-MBNL fusions together with CUG^{exp} in HeLa cells. The combined analysis of ribonucleoprotein foci, visualized by FISH (to detect CUG^{exp}), and the signals from GFP-MBNL proteins revealed that all tested MBNL isoforms colocalized with repeat transcripts in nuclear foci (Supplementary Figure S9A). MBNL overexpression caused the average CUG^{exp} foci volume to increase by up to 2-fold, which we termed foci pumping, without increasing expression of this RNA and this strongly correlated with the volume of the fluorescence protein signal within foci (Figure 6B and C; Supplementary Figure S9B-D). These observations led us to elucidate the nature of protein-protein and RNA-protein interactions in nuclear foci.

We coexpressed different combinations of two fluorescence protein-MBNL fusions in the presence or absence of CUG^{exp} transcripts. To measure the proximity of fusion proteins in the cytoplasm, nucleoplasm or in CUG^{exp} foci, we applied FLIM and measured the efficiency of FRET between GFP (donor) and mCherry (acceptor). FRET occurs if the range between donor and acceptor is <10 nm and fused GFP-mCherry protein constituted a positive control for which 20% of FRET was detected (Figure 6D and Supplementary Figure S10A). For further experiments we selected cells only with an equal expression level (Supplementary Figure S10I). In HeLa cells, with or without the expression of CUG^{exp}, neither combination of MBNL isoforms exhibited efficient FRET in cytoplasmic or nucleoplasmic compartments (Figure 6D; Supplementary Figure S10B and H). Thus, this population of MBNL proteins does not undergo FRET detectable homotypic or heterotypic interactions. On the other hand, in nuclear CUG^{exp} foci FRET efficiency for each combination of MBNL protein, but not for a control MBNL1 and CELF1 pair, was very high and surpassed the positive GFP-mCherry control (average FRET efficiency 16-26%) (Figure 6; Supplementary Figure S10B-F and H). The most prominent difference between FRET efficiencies were observed for homotypic MBNL1 and homotypic MBNL3 interactions (Figure 6E). These results suggest that transcripts containing expanded CUG repeats induce an increase of local MBNL concentration leading to multimerization. To test this hypothesis, we utilized the MBNL protein truncated for the C-terminal domain (MB1-C) (Supplementary Figure S8A). Although, previous studies have shown that the C-terminal region is required for MBNL multimerization (16,26,51), we did not observe significant differences between FRET efficiency for the full-length and C-truncated MBNL1 isoforms (Figure 6D; Supplementary Figure S10G and H). This would imply that the proximity of MBNL proteins in CUG^{exp} foci is very high for all tested proteins, including MB1-C. Our

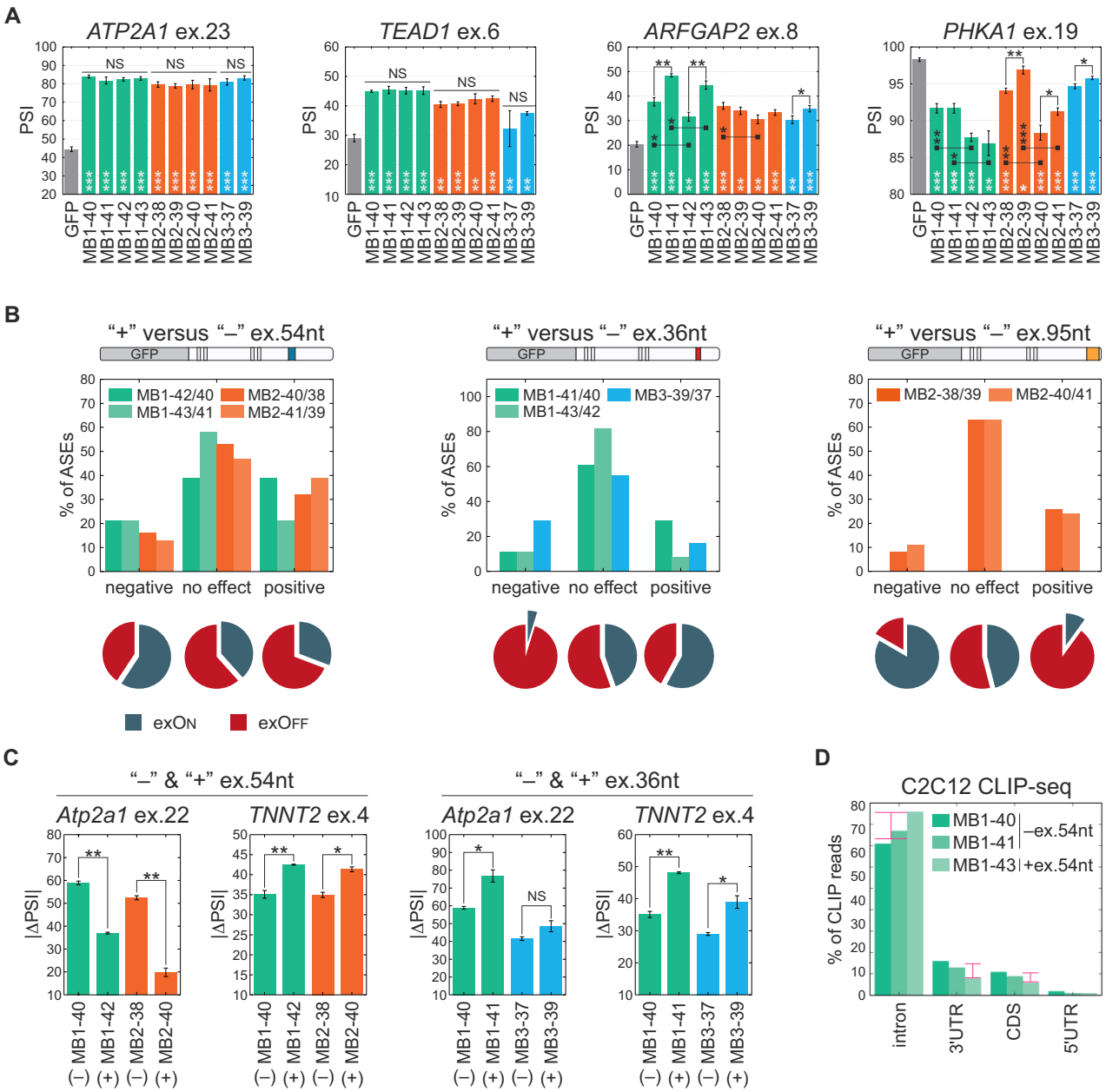


Figure 5. Specific MBNL alternative exons modulate splicing activity. (**A**) Compared splicing activity (expressed as PSI values) of all tested MBNL proteins for four selected AS events analyzed by RT-PCR. Bars represent average PSI ± SD from three independent experiments and statistical significance was determined by the Student's *t*-test (∗ for $P < 0.05$, ∗∗ for $P < 0.01$ and ∗∗∗ for $P < 0.001$). White stars refer to statistical differences between a particular MBNL isoform and GFP control. Black stars refer to statistical differences between comparable MBNL isoforms having or lacking a sequence encoded by alternative ex.54nt, ex.36nt and ex.95nt. Supplementary Table S2 contains results for all 38 AS events. (**B**) Combined analysis of the 38 tested AS events for the comparable MBNL isoforms differing in the presence or absence of sequences encoded by ex.54nt, ex.36nt or ex.95nt. The comparison was made for nine pairs of isoforms, four for ex.54nt, three for ex.36nt and two for ex.95nt. AS events were divided into categories showing stronger, weaker or no splicing changes for individual isoform pairs taking statistical significance into consideration ('no change', $P \geq 0.05$; weaker and stronger, $P < 0.05$; Student's *t*-test). Note that the mean strength of splicing changes for isoforms having ex.95nt was 2.6-time higher for exons under positive control compared to ex.36nt. The pie charts represent percentage of exONs and exOFFs for each category represent (more information in text) (**C**) The influence of ex.54nt and ex.36nt on the splicing regulation of exogenous mouse *Atp2a1* ex.22 and human *TNNT2* ex.4. Bars represent average PSI ± SD from three independent experiments and statistical significance was determined by the Student's t-test (∗ for $P < 0.05$, ∗∗ for $P < 0.01$ and ∗∗∗ for $P < 0.001$). (**D**) Differences in MBNL1 isoform binding site distribution determined by CLIP-seq analysis. Percentage of unique CLIP-seq reads mapped to different regions of transcripts, namely introns located only in pre-mRNA (nucleus) and coding sequences (CDS), 5′UTRs and 3′UTRs located in both pre-mRNA and mRNA (nucleus and cytoplasm) for three MBNL1 isoforms (MB1-40 and MB1-41 without ex.54nt and MB1-43 with ex.54nt showing exclusively nuclear localization). Pink arrows indicate differences between percentage of reads for MBNL1 isoforms with and without ex.54nt.

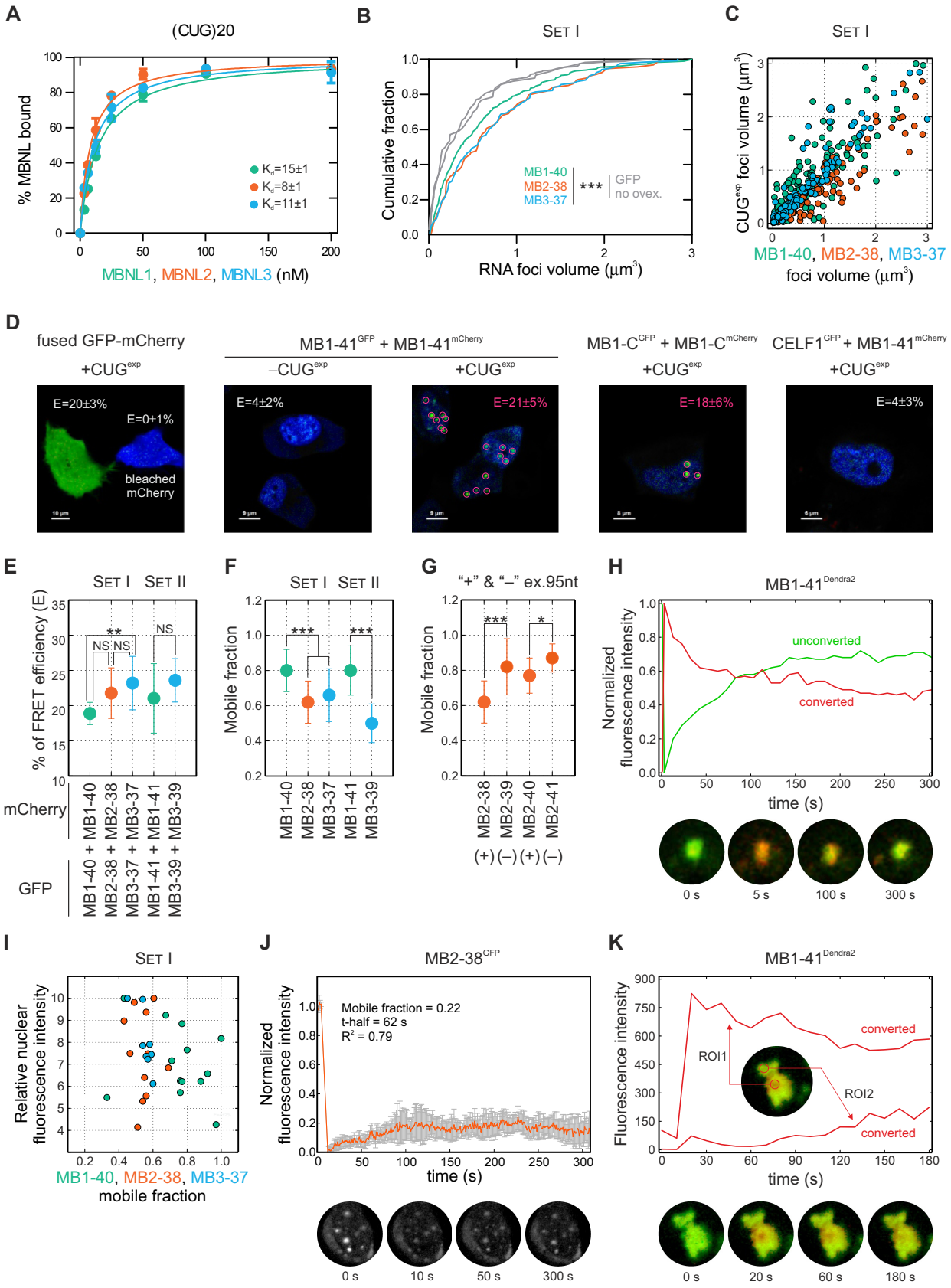


Figure 6. MBNL on CUG^{exp} RNA foci. (**A**) High affinity of recombinant MBNL1, MBNL2 and MBNL3 for (CUG)20 RNA *in vitro*. The K_d values are indicated. For (CCUG)14 and (CAG)20 transcripts, see Supplementary Figure S6E. (**B**) Overexpression of MBNL paralogs significantly increase CUG^{exp}

data indicate that MBNL proteins bind to CUG repeats with high affinity and CUG^{exp} constitutes a nucleation center for MBNL accumulation, which leads to increased RNA foci size. Although the nuclear localization of MBNL3 was the lowest among MBNL paralogs, MBNL3 formed even more compacted foci due to more effective CUG^{exp}-protein and/or protein-protein interactions.

Rapid exchange of MBNL proteins between the nucleoplasm and CUG^{exp} foci

To investigate whether MBNL isoforms differ in their mobility and accumulation in CUG^{exp} foci, we visualized the dynamic nature of MBNL-CUG^{exp} complexes by performing 3D time-lapse sequences over 16 h. RNA foci containing exogenous fluorescently labeled MBNL proteins were formed *de novo* and then observed for dispersion as well as structural and volumetric changes (Supplementary Videos S1 and S2). While the results were consistent with previous observations (38), longer CUG repeat expansions formed more stable foci and the half-life of foci formation was measured in either minutes or hours, depending on the CUG repeat tract length, which was (CUG)145 in (38) versus (CUG)960 in our study.

In both DM1 and mouse models of this disease, there is a threshold of CUG repeat length and/or expression level which influences the degree of MBNL sequestration (43). This suggests the importance of stoichiometry between the number of MBNL binding sites on CUG^{exp} RNA and the number of MBNL proteins in a single nucleus. Thus, we examined MBNL protein mobility in both saturated and unsaturated states in which there are more, or less, MBNL binding sites than MBNL proteins and also characterized differences in the mobility of MBNL proteins between nucleoplasm and foci. GFP-MBNL fusion proteins were coexpressed with CUG^{exp} RNA in HeLa cells and protein mobility assessed using FRAP. We first examined cells with a saturated MBNL protein level in which proteins were present in both nucleoplasm and foci (Supplementary Video S3). For FRAP, we calculated two parameters, the mobile fraction and the half-time of fluorescence recovery. The mobile fraction defines the percent of the GFP fluorescence signal which diffuses from the nucleoplasm to RNA foci after photobleaching (2 μm diameter of selected foci) and the half-time ($t_{\frac{1}{2}}$) is defined as the half maximal fluorescence recovery time. For each model we observed high heterogeneity of the mobile fraction and $t_{\frac{1}{2}}$ (Supplementary Figure S11A). For proteins representing SET I and SET II groups, the average mobile fractions for MBNL1 proteins were ∼39% higher compared to MBNL3 (Figure 6F) and this phenomenon was not affected by the nuclear concentration of GFP fusion proteins (Figure 6I). As anticipated, MBNL1 lacking the first ZnF tandems (MB1-N) (Supplementary Figure S8A) exhibited the lowest affinity for CUG repeats (55), had the highest mobility and the shortest $t_{\frac{1}{2}}$ (Supplementary Figure S11A) suggesting a very high rate of diffusion between CUG^{exp} foci and nucleoplasm. In contrast, MBNL3 proteins had the lowest mobile fraction. Moreover, a significant difference in the mobile fraction between MBNL2 proteins containing or lacking ex.95nt was observed and the inclusion of this exon reduced the efficiency of fluorescence recovery in foci (Figure 6G). These results show that MBNL3 proteins have the lowest, whereas MBNL1 (especially truncated MB1-Ns) proteins have the highest, ability to dissociate from CUG^{exp} RNA and exit RNA foci and some alternative MBNL exons modulate this mobility rate.

To confirm that MBNL proteins diffuse freely between the nucleoplasm and CUG^{exp} foci, we used a photoconvertible Dendra2 fluorescence protein fused to MBNL1 in cells expressing a high level of this protein. After laser-induced photoswitching of Dendra2 from unconverted green (507 nm) to converted red (573 nm), we observed a rapid decrease of the red signal intensity in favor of an elevation of the green signal (Figure 6H, Supplementary Figure S11B and Supplementary Video S4). This indicates that the photoconverted MBNL1 fusion protein migrating within foci was shifted away whereas the unconverted green version of the same protein was simultaneously approaching foci. The quantity of proteins exiting and entering the foci was sim-

←

foci volume ('foci pumping') measured by FISH for ∼130 foci. There is no difference in CUG^{exp} foci volume distribution between cells transfected with GFP and without protein overexpression (no ovex.). Statistical significance was assessed by Mann-Whitney U test (** for $P < 0.01$ and *** for $P < 0.001$). (**C**) Correlation between RNA foci volume with MBNL volume (measured by the GFP signal). The Spearman's ρ is between 0.87 and 0.95. FISH and GFP fluorescence measurements were performed using quantitative confocal microscopy, see Supplementary Figure S9. (**D**) Analysis of FRET efficiency (**E**) between GFP and mCherry either fused together (left picture) or with two MBNL proteins or with CELF1 in a control experiment (right picture) in the absence or presence of a CUG^{exp} transcript. The *E*-values which are average from 10-20 analyses for pairs of full length MB1-41 and truncated MB1-C in CUG^{exp} foci are as high as for a positive GFP-mCherry control in an entire cell. In the absence of CUG^{exp}, the *E*-value is slightly above the background and similar to the value observed for a CELF1 and MB1-41 pair. For more results, see Supplementary Figure S10. (**E**) MBNL paralogs differ in *E*. Statistical significance was determined by Student's *t*-test (NS for $P \geq 0.05$, ** for $P < 0.01$). (**F**) MBNL paralogs differ in their mobile fraction in FRAP experiments performed in cells saturated with MBNL protein. The results are mean from about 20 nuclei ± SD and statistical significance was determined by Mann-Whitney U test (* for $P < 0.05$ and *** for $P < 0.001$). See Supplementary Video S3. (**G**) The presence of a sequence encoded by alternative ex.95nt reduces MBNL2 mobility in FRAP experiments. For more examples, see Supplementary Figure S11A. (**H**) Photoswitching of Dendra2 fused with MB1-41 in cells with saturated levels of MBNL protein. Photoconverted (red, emission 573 nm) protein is shifting away whereas unconverted (green, emission 507 nm) protein is associating with CUG^{exp} foci (see Supplementary Video S4). The same analysis for MB1-40 was shown in Supplementary Figure S11B. (**I**) There is no correlation between the relative nuclear fluorescence intensity of GFP fused MBNL proteins and the mobile MBNL fraction. (**J**) FRAP experiment for MB2-38 in cells with the very low fluorescence signal of the fusion protein in nuclei. Note that the mobile fraction is three-times lower in the unsaturated compared to saturated state of MBNL (see Supplementary Figure S11A for comparison). (**K**) The diffusion of photoconverted Dendra2 fused with MBNL1 within individual CUG^{exp} focus in a cell with the low fluorescence signal in nucleoplasm (unsaturated MBNL state). Quantification was performed from two regions of interest (ROI1 and ROI2). The ROI1 is an area that was photoconverted by a laser and the red signal decreases during the course of time. On the other hand, in the distant ROI2 of the same CUG^{exp} focus in which only green signal was detected in time 0, the red signal from photoconverted Dendra2 increases for several seconds. In this situation, MBNL proteins are moving to other binding sites on CUG^{exp} RNA and are not dissociating from the foci.

ilar. Moreover, within a few minutes after laser induction, we monitored the red fluorescence signal enhancement at other foci in the same nucleus. Subsequently, we addressed the question of the rate of MBNL protein mobility in cells with a very low nucleoplasmic MBNL concentration. For this purpose, we performed FRAP in HeLa cells with weak nucleoplasmic fluorescence and relatively intense CUG^{exp} foci signals. As expected, there was almost no FRAP in these cells (Figure 6J). We hypothesized that MBNL proteins, present in the unsaturated state, are not able to escape from CUG^{exp} foci but circulate between available RNA binding sites within these structures. To test this hypothesis, we performed the experiments with a photoconvertible Dendra2 protein fused to MBNL1 in cells with a low nucleoplasmic signal and after photoswitching of Dendra2 we observed diffusion of the red fluorescence signal of the fusion protein within individual CUG^{exp} foci (Figure 6K). Cumulatively, these data indicated that when CUG^{exp} foci are saturated with MBNL, the proteins dissociate from the foci and are rapidly exchanged by nucleoplasmic MBNL proteins. In contrast, MBNL proteins circulate within RNA foci when unoccupied binding sites are available. Additionally, nuclear foci are mobile and dynamic structures, capable of slowly increasing and decreasing their volume and fusing with each other.

DISCUSSION

MBNL protein family consists of three paralogs that regulate RNA processing and localization (5,6,9–11) and undergo significant quantitative and qualitative changes during human tissue development. In spite of the strong sequence similarity between these paralogs, *Mbnl1*, *Mbnl2* and *Mbnl3* mouse knockouts develop strikingly different phenotypes. $Mbnl1^{\Delta E3/\Delta E3}$, $Mbnl2^{\Delta E2/\Delta E2}$ and $Mbnl3^{\Delta E2}$ exhibit muscle pathology and ocular cataracts, central nervous system disorders and impaired muscle regeneration, respectively (3,18,23). Moreover, depletion of only MBNL1 and MBNL2 cause the global metabolic RNA changes observed in DM1 adults (4,6,20). An explanation for these differences is the diverse MBNL expression pattern in specific tissues with MBNL1 predominately expressed in skeletal muscles whereas MBNL2 is expressed at a higher level in many neuronal cells (10). However, the differences in the activity of MBNL paralogs have never been directly tested. Thus, we addressed the basis of these MBNL activity differences, including the roles of alternatively spliced isoforms, by studying selected endogenous and exogenous MBNL transcripts. We focused on a comparison of MBNL1, MBNL2 and MBNL3 paralogs, as well as their splicing isoform activities, to determine the structural features which influence AS regulation (Figure 7A). All MBNL paralog activities were regulated by either promotion (exON) or repression (exOFF) of alternative exon inclusion. Notably, MBNL1 was the most active splicing factor and MBNL2 has intermediate activity whereas MBNL3 possessed the weakest influence on AS changes and individual AS events differed in their sensitivity to MBNL overexpression. We conclude that all three MBNL paralogs preferentially recognize the same RNA binding motifs on their RNA targets but often with different affinity. This difference is likely caused by the presence of linker sequences between the ZnF tandem motifs, which are quite different between the three paralogs, that have been shown to influence MBNL activity (16,26).

We also showed that the different splicing activity of MBNLs may depend, at least partially, on subcellular localization. MBNL paralogs differed significantly in nucleoplasmic and cytoplasmic distribution (Figure 7A). MBNL1 predominated in nucleoplasm whereas MBNL3 predominated in cytoplasm. It may have a significant impact on nuclear and cytoplasmic functions of MBNLs such as AS/polyadenylation and mRNA stability/trafficking. Previous studies have demonstrated that MBNLs truncated for the N-terminal sequence of protein containing the first tandem of ZnFs (MB1-N) and C-terminal (MB1-C) sequence containing KRPALE motif had a negative effect on nuclear localization (16,18,19). Both regions vary in the amino acid sequence between MBNL paralogs leading to relevant consequences. For instance, the difference in the substitution of a third amino acid in the KRPALE motif between all MBNLs was previously shown to effect MBNL1 nuclear localization (19). Moreover, each MBNL paralog expresses multiple protein isoforms (2,16,56) and our results indicate that that the inclusion of amino acid sequences encoded by the three tested alternative exons (ex.54nt, ex.36nt, ex.95nt) are either neutral or can increase or decrease splicing activity. However, the final effects on splicing depend primarily on the specific RNA targets and the mechanism of MBNL action (exON/exOFF). AS of these exons changes during normal human tissue development and in myotonic dystrophy (Figure 1B) and their aberrant regulation modulates spliceopathy evoked by MBNL sequestration in DM.

Currently, there is no consensus about the effect of the sequence encoded by MBNL1 and MBNL2 alternative ex.54nt on RNA splicing regulation (16,17,19). Our results suggest that this sequence reduces MBNL interactions with other factors participating in the exON mechanism since the inclusion of ex.54nt resulted in mainly negative effects on exON (e.g. *Atp2a1* ex.22). Previously ex.36nt was described to encode the sequence prominent for MBNL1-MBNL1 homotypic interactions (16,51). Here, we show that the sequence encoded by this exon, but also ex.95nt, also modulates MBNL splicing activity. The negative effect of ex.36nt and the positive effect of ex.95nt were almost exclusively specific for exOFFs. This observation suggests that the sequences encoded by both alternative exons are important for the interaction with other trans-acting factors and/or to enhance protein affinity for RNAs during the regulation of specific exon exclusion.

All MBNL paralogs have a very high affinity for expanded CUG and CCUG repeat *in vitro* and in cell models. We observed that overexpression of these RNAs significantly increased the volume of ribonuclear foci consistent with previously described results showing a reduction in RNA foci size due to the silencing of endogenous MBNL1 and MBNL2 (24,38,41). All tested isoforms form CUG^{exp} RNA foci that are heterogeneous in size although the average size of CUG^{exp} foci remained the same. RNA foci volume depends primarily on the amount of accumulated CUG repeat transcripts; however, they slightly differ in the density of MBNL proteins loaded on RNAs. MBNL1 is

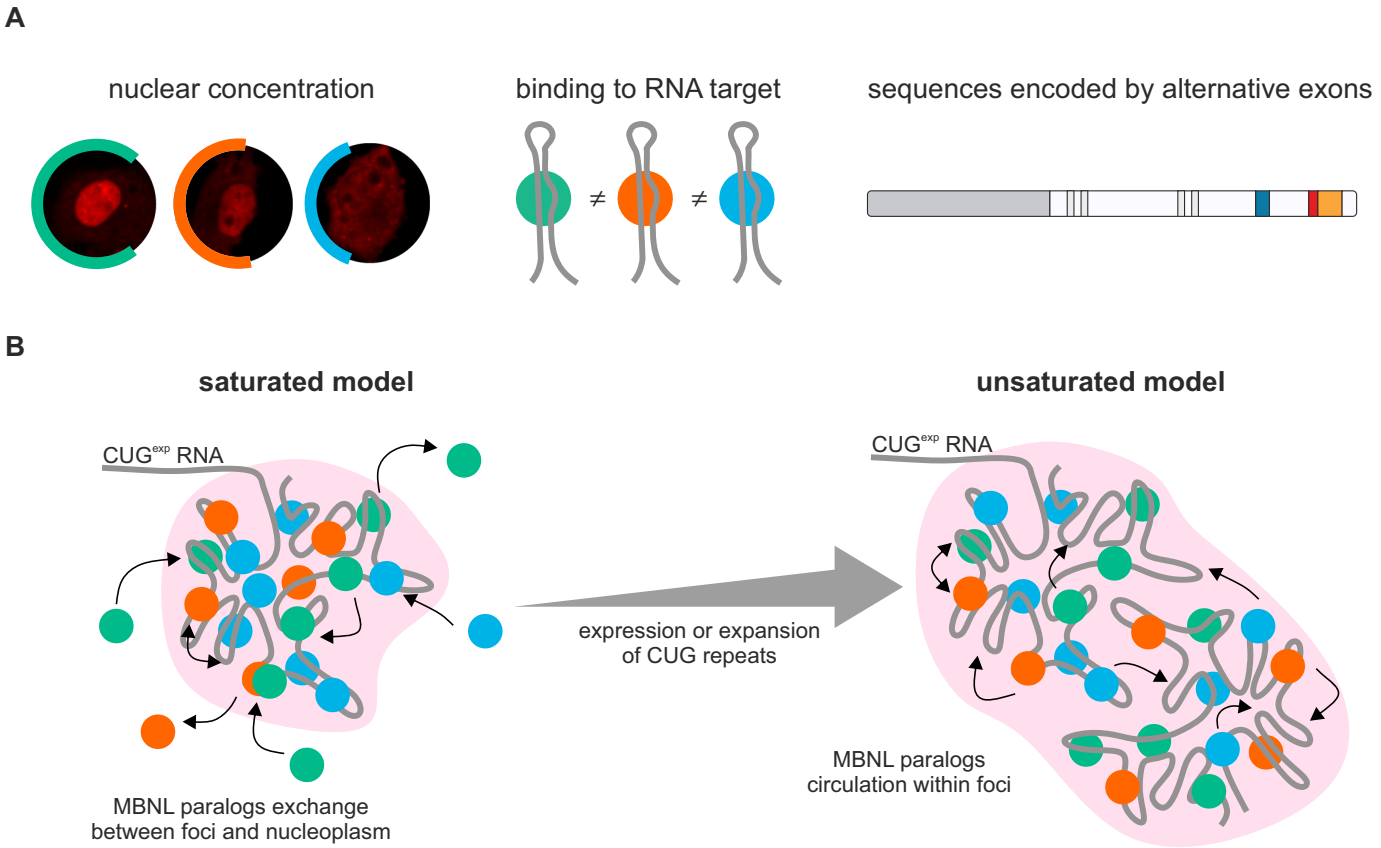


Figure 7. Intracellular determinants of MBNL activity and localization. (**A**) The three main determinants of MBNL activity determined in current study. (**B**) Deterministic nature of MBNL-CUG^{exp} foci formation is disrupted by the chaos within individual focus. Two distinct stages, differing in the number of MBNL-binding sites on CUG^{exp} were shown. MBNLs would be in saturated (left panel) or unsaturated state (right panel) in CUG^{exp} foci. The latter enables efficient MBNL sequestration. See the text for more details.

less densely packed with a lower FRET efficiency but it is also significantly more mobile and efficiently dissociates from, and associates with, CUG^{exp} RNA and the former feature is under control of the sequence encoded by alternative ex.95nt. In good agreement with previous data (38), these results suggest that MBNL-CUG^{exp} foci formation resembles a stochastic process (Figure 7B) although are some determinants, including the sequences encoded by alternatively spliced MBNL exons, that modulate the affinities of the three MBNL paralogs for CUG^{exp} RNA. Therefore, we conclude that foci formation is also deterministic and chaotic nature (deterministic chaos). Generally, MBNL paralogs bind to CUG repeats and dissociate from them when MBNL is in a saturated state (Figure 7B, left panel). In DM1 and DM2, the sequestration of proteins on C(C)UG^{exp} leads to a significant reduction in the free MBNL pool available for normal RNA splicing targets and this situation impacts global AS and APA regulation (57). Effective sequestration of MBNL activity depends on the length of the repeat tract which is modulated by somatic expansion during the lifespan of DM1 patients. When we examined cells designed to reproduce the impaired status of MBNL proteins in DM1, we noted that these proteins are unable to dissociate from RNA foci and instead continuously change intra-foci binding sites on CUG^{exp} transcripts. Since somatic expansion in tissues of DM1 patients alters the stoichiometry between MBNL proteins and their RNA binding sites leading to the unsaturated state, we conclude that phenomenon underlies MBNL sequestration and the progression of the DM1 disease phenotype (Figure 7B right panel).

How do these results impact our understanding of the roles of MBNL proteins in DM1 disease and therapeutic development? Current disease models suggest that age-related expansion of the *DMPK* CTG^{exp} leads to eventual titration and loss of MBNL splicing activity resulting in the appearance of more severe pathological manifestations, including muscle wasting. Ongoing clinical trials are evaluating the use of AON gapmers to target and degrade *DMPK* mutant allele transcripts (50). Our results shed light on the mechanism of action of these therapeutic agents as well as small molecule compounds that target MBNL-CUG^{exp} RNA interactions. MBNL proteins are mobile in RNA foci and freely dissociate from these structures so both gapmers and small molecule compounds may actively compete for RNA binding sites and modulate CUG^{exp} RNA structure to make these sites less accessible to circulating MBNL proteins. Since MBNL3 is more prone to be tightly packed on CUG^{exp} RNA, but is primarily expressed only during embryogenesis and tissue regeneration, overexpression of this MBNL paralog should be considered for the future gene therapy development for myotonic dystrophy.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank M.S. Swanson (University of Florida) for critical comments and manuscript corrections, J.A. Berglund (University of Florida) for critical comments, T.A. Cooper (Baylor College of Medicine) for *TNNT2* and DT960 minigenes, W.J. Krzyzosiak (Polish Academy of Sciences) for (CUG)20, (CAG)20 molecules and (CAGG)14 template, I. Makalowska (Adam Mickiewicz University) for MIB hosting and cooperation agreement, C.A. Thornton (University of Rochester Medical Center) for A2764 antibody and RNA samples from human muscles, M. Figiel (Polish Academy of Sciences) for mouse muscles, P. Wojtaszek (Adam Mickiewicz University) for cooperation agreement, M. Taylor for manuscript proofreading.
Author Contributions: Conception and design: Ł.J.S. and K.S. Acquisition of data: Ł.J.S. (all without *in vitro* assays, cellular experiments with AONs and CLIP-seq in C2C12), M. Michalak (confocal microscopy without FISH), K.T. (*in vitro* assays), P.C. (cellular experiments with AONs, AS events in human tissues), A.W-S. (FISH), M. Matłoka (*in vitro* assays) and P.K. (CLIP-seq in C2C12). Analysis and interpretation of data: Ł.J.S. (all), M. Michalak (confocal microscopy without FISH), M.K. (RIP-seq, CLIP-seq), K.T. (*in vitro* assays), P.C. (cellular experiments with AONs) and K.S. (all-supervision). Drafting or revising the article: Ł.J.S and K.S.

FUNDING

Foundation for Polish Science TEAM program cofinanced by the European Union within the European Regional Development Fund [TEAM/2011-7/10 to K.S.]; Polish National Science Centre [2011/01/B/NZ1/01603, 2014/15/B/NZ2/02453 to K.S., 2014/12/T/NZ2/00516 to Ł.J.S.]; National Multidisciplinary Laboratory of Functional Nanomaterials NanoFun [POIG.02.02.00-00-025/09]. 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 for the years 2014-2019.
Conflict of interest statement. None declared.

REFERENCES

1. Scotti,M.M. and Swanson,M.S. (2016) RNA mis-splicing in disease. *Nat. Rev. Genet.*, **17**, 19–32.
2. Pascual,M., Vicente,M., Monferrer,L. and Artero,R. (2006) The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation*, **74**, 65–80.
3. 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*, **302**, 1978–1980.
4. Batra,R., Charizanis,K., Manchanda,M., Mohan,A., Li,M., Finn,D.J., Goodwin,M., Zhang,C., Sobczak,K., Thornton,C.A. *et al.* (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell*, **56**, 311–322.
5. Wang,E.T., Ward,A.J., Cherone,J.M., Giudice,J., Wang,T.T., Treacy,D.J., Lambert,N.J., Freese,P., Saxena,T., Cooper,T.A. *et al.* (2015) Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.*, **25**, 858–871.
6. 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.
7. 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.
8. 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.
9. Taliaferro,J.M., Vidaki,M., Oliveira,R., Olson,S., Zhan,L., Saxena,T., Wang,E.T., Graveley,B.R., Gertler,F.B., Swanson,M.S. *et al.* (2016) Distal alternative last exons localize mRNAs to neural projections. *Mol. Cell*, **61**, 821–833.
10. 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.
11. 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 Expr. Patterns*, **3**, 459–462.
12. 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.
13. He,F., Dang,W., Abe,C., Tsuda,K., Inoue,M., Watanabe,S., Kobayashi,N., Kigawa,T., Matsuda,T., Yabuki,T. *et al.* (2009) Solution structure of the RNA binding domain in the human muscleblind-like protein 2. *Protein Sci.*, **18**, 80–91.
14. 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.
15. Lambert,N., Robertson,A., Jangi,M., McGeary,S., Sharp,P.A. and Burge,C.B. (2014) RNA bind-n-seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell*, **54**, 887–900.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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.
21. Jog,S., Paul,S., Dansithong,W., Tring,S., Comai,L. and Reddy,S. (2012) RNA splicing is responsive to MBNL1 dose. *PLoS One*, **7**, e48825.
22. 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.
23. 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.

24. 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.
25. Sen,S., Talukdar,I., Liu,Y., Tam,J., Reddy,S. and Webster,N. (2010) Muscleblind-like 1 (Mbnl1) promotes insulin receptor exon 11 inclusion via binding to a downstream evolutionarily conserved intronic enhancer. *J. Biol. Chem.*, **285**, 25426–25437.
26. 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.
27. Carpentier,C., Ghanem,D., Fernandez-Gomez,F.J., Jumeau,F., Philippe,J.V., Freyermuth,F., Labudeck,A., Eddarkaoui,S., Dhaenens,C.M., Holt,I. *et al.* (2014) Tau exon 2 responsive elements deregulated in myotonic dystrophy type I are proximal to exon 2 and synergistically regulated by MBNL1 and MBNL2. *Biochim. Biophys. Acta*, **1842**, 654–664.
28. 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.
29. Mankodi,A., Urbinati,C., Yuan,Q., Moxley,R., Sansone,V., Krym,M., Henderson,D., Schalling,M., Swanson,M. and Thornton,C. (2001) Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.*, **10**, 2165–2170.
30. 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.
31. Fu,Y.H., Pizzuti,A., Fenwick,R.G. Jr, King,J., Rajnarayan,S., Dunne,P.W., Dubel,J., Nasser,G.A., Ashizawa,T., de Jong,P. *et al.* (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science*, **255**, 1256–1258.
32. 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.
33. 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.
34. Napierała,M. and Krzyzosiak,W. (1997) CUG repeats present in myotonin kinase RNA form metastable "slippery" hairpins. *J. Biol. Chem.*, **272**, 31079–31085.
35. Sobczak,K., de Mezer,M., Michlewski,G., Krol,J. and Krzyzosiak,W. (2003) RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.*, **31**, 5469–5482.
36. Sobczak,K., Michlewski,G., de Mezer,M., Kierzek,E., Krol,J., Olejniczak,M., Kierzek,R. and Krzyzosiak,W. (2010) Structural diversity of triplet repeat RNAs. *J. Biol. Chem.*, **285**, 12755–12764.
37. Jiang,H., Mankodi,A., Swanson,M., Moxley,R. and Thornton,C. (2004) Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.*, **13**, 3079–3088.
38. Querido,E., Gallardo,F., Beaudoin,M., Ménard,C. and Chartrand,P. (2011) Stochastic and reversible aggregation of mRNA with expanded CUG-triplet repeats. *J. Cell Sci.*, **124**, 1703–1714.
39. Ho,T., Savkur,R., Poulos,M., Mancini,M., Swanson,M. and Cooper,T. (2005) Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J. Cell Sci.*, **118**, 2923–2933.
40. Pettersson,O.J., Aagaard,L., Andrejeva,D., Thomsen,R., Jensen,T.G. and Damgaard,C.K. (2014) DDX6 regulates sequestered nuclear CUG-expanded DMPK-mRNA in dystrophia myotonica type 1. *Nucleic Acids Res.*, **42**, 7186–7200.
41. Smith,K., Byron,M., Johnson,C., Xing,Y. and Lawrence,J. (2007) Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type I is blocked at entry into SC-35 domains. *J. Cell Biol.*, **178**, 951–964.
42. 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.
43. 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*, **289**, 1769–1773.
44. Chen,C., Sobczak,K., Hoskins,J., Southall,N., Marugan,J., Zheng,W., Thornton,C. and Austin,C. (2012) Two high-throughput screening assays for aberrant RNA-protein interactions in myotonic dystrophy type 1. *Anal. Bioanal. Chem.*, **402**, 1889–1898.
45. Wojciechowska,M., Taylor,K., Sobczak,K., Napierala,M. and Krzyzosiak,W.J. (2014) Small molecule kinase inhibitors alleviate different molecular features of myotonic dystrophy type 1. *RNA Biol.*, **11**, 742–754.
46. Philips,A.V. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science*, **280**, 737–741.
47. Hino,S.-I., Kondo,S., Sekiya,H., Saito,A., Kanemoto,S., Murakami,T., Chihara,K., Aoki,Y., Nakamori,M., Takahashi,M. *et al.* (2007) Molecular mechanisms responsible for aberrant splicing of SERCA1 in myotonic dystrophy type 1. *Hum. Mol. Genet.*, **16**, 2834–2843.
48. Rapsomaniki,M.A., Kotsantis,P., Symeonidou,I.E., Giakoumakis,N.N., Taraviras,S. and Lygerou,Z. (2012) easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. *Bioinformatics*, **28**, 1800–1801.
49. Wojtkowiak-Szlachcic,A., Taylor,K., Stepniak-Konieczna,E., Sznajder,L.J., Mykowska,A., Sroka,J., Thornton,C.A. and Sobczak,K. (2015) Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res.*, **43**, 3318–3331.
50. 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.
51. 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.
52. Mallinjoud,P., Villemin,J.P., Mortada,H., Polay Espinoza,M., Desmet,F.O., Samaan,S., Chautard,E., Tranchevent,L.C. and Auboeuf,D. (2014) Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin. *Genome Res.*, **24**, 511–521.
53. 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.
54. Pettersson,O.J., Aagaard,L., Jensen,T.G. and Damgaard,C.K. (2015) Molecular mechanisms in DM1— a focus on foci. *Nucleic Acids Res.*, **43**, 2433–2441.
55. 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.
56. Kanadia,R., Shin,J., Yuan,Y., Beattie,S., Wheeler,T., Thornton,C. and Swanson,M. (2006) Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 11748–11753.
57. Chau,A. and Kalsotra,A. (2015) Developmental insights into the pathology of and therapeutic strategies for DM1: back to the basics. *Dev. Dyn.*, **244**, 377–390.
58. Licatalosi,D., Mele,A., Fak,J., Ule,J., Kayikci,M., Chi,S., Clark,T., Schweitzer,A., Blume,J., Wang,X. *et al.* (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, **456**, 464–469.

Mechanistic Determinants of MBNL Activity

Łukasz J. Sznajder[1*], Michał Michalak[2], Katarzyna Taylor[1], Piotr Cywoniuk[1], Michał Kabza[3], Agnieszka Wojtkowiak-Szlachcic[1], Magdalena Matloka[1], Patryk Konieczny[1], Krzysztof Sobczak[1*]

[1] Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland.

[2] Department of Molecular and Cellular Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland.

[3] Department of Bioinformatics, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland.

* To whom correspondence should be addressed. Tel: +4861 829 5958; Fax: +4861 829 5949;
Email: ksobczak@amu.edu.pl.
Correspondence may also be addressed to Łukasz J. Sznajder. Email: lukasz.j.sznajder@gmail.com or ljsznajder@ufl.edu.

Supplementary Figures

(in this file)

Supplementary Figure S1. Exonic structure and expression pattern of MBNL paralogs.

Supplementary Figure S2. The effect of 10 different MBNL isoforms on endogenous AS events.

Supplementary Figure S3. Higher impact of MBNL proteins on alternative exon exclusion (exOFF) than on exon inclusion (exON).

Supplementary Figure S4. Activity of different MBNL proteins from SET I and SET II on *Atp2a1* ex.22 and *TNNT2* ex.4 alternative splicing.

Supplementary Figure S5. CLIP-seq procedure for endogenous and exogenous MBNL1 study.

Supplementary Figure S6. Comparison of affinity of three MBNL paralogs to selected RNA fragments.

Supplementary Figure S7. Distinct localization of MBNL paralogs in nucleoplasm and cytoplasm.

Supplementary Figure S8. Truncated MBNL1s.

Supplementary Figure S9. MBNL proteins influence the CUG^{exp} foci volume.

Supplementary Figure S10. Comparative analysis of FRET efficiency (E) between GFP and mCherry fused with different MBNL proteins.

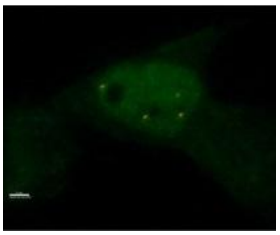
Supplementary Figure S11. Differences in mobility of MBNL isoforms between CUG^{exp} foci and nucleoplasm based on fluorescence recovery after photobleaching (FRAP) experiments.

Supplementary Tables

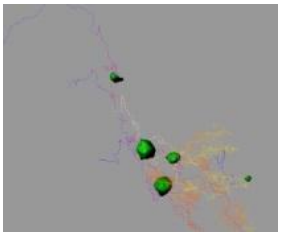
Supplementary Table S1 (.xlsx file). Detailed information about MBNL1, MBNL2 and MBNL3 genetic constructs.

Supplementary Table S2 (.xlsx file). Detailed information about tested AS events.

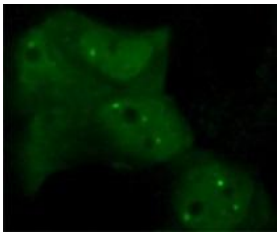
Supplementary Videos



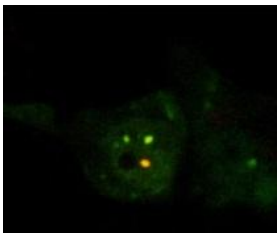
Supplementary Video S1 (.avi file). 16 hours time-lapse sequence of living HeLa cell expressing exogenous MBNL1 protein fused with GFP and CUG^{exp} (part 1).



Supplementary Video S2 (.avi file). 16 hours time-lapse sequence of living HeLa cell expressing exogenous MBNL1 protein fused with GFP and CUG^{exp} (part 2).



Supplementary Video S3 (.avi file). Fluorescence recovery after photobleaching (FRAP) of exogenous MBNL1 protein fused with GFP in CUG^{exp} foci.



Supplementary Video S4 (.avi file). Photoswitching of Dendra2 fluorescence protein fused with MBNL1 in CUG^{exp} foci.

Supplementary Material & Methods

(in this file)

Supplementary Figures

Supplementary Figure S1

Supplementary Figure S1

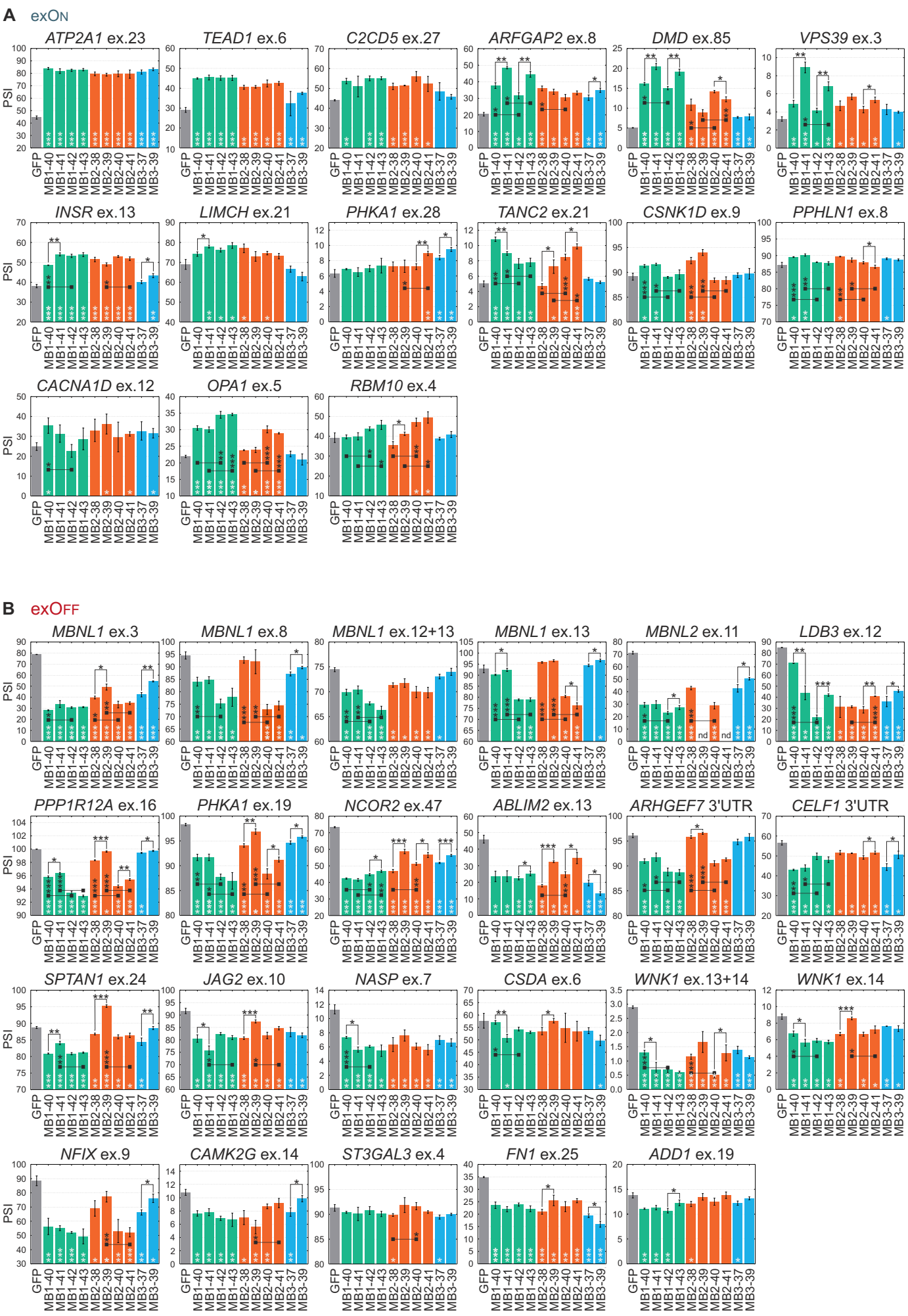
Exonic structure and expression pattern of MBNL paralogs.

(A) The exonic structure of MBNL paralogs. Exon enumeration derived from previous studies (1) and FasterDB (2). For more information see Supplementary Table S1.
(B) Gene expression level of *MBNL* paralogs and *CELF1* in 73 cell lines derived from FasterDB (2).
(C) Amino acid identity of MB1-43 (ex.54nt, ex.36nt and ex.95nt), MB2-40 (ex.54nt and ex.95nt) and MB3-39 (ex.36nt and ex.95nt) proteins. Zinc fingers, ex.54nt, ex.36 and ex.95 are highlighted on gray, blue, red and orange respectively. Various amino acids are highlighted.
(D) The level of exogenous MBNL proteins. Exogenous proteins were detected by anti-GFP antibody and normalized to GAPDH. Bars represent average expression level ± SD from two independent biological experiments.

References

(1) Pascual M, Vicente M, Monferrer L, Artero R. The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. Differentiation. 2006;74:65-80. doi:10.1111/j.1432-0436.2006.00060.x
(2) Mallinjoud P, Villemin JP, Mortada H, Polay Espinoza M, Desmet FO, Samaan S, et al. Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin. Genome Res. 2014;24:511-521. doi: 10.1101/gr.162933.113.

Supplementary Figure S2



Supplementary Figure S2

The effect of 10 different MBNL isoforms on endogenous AS events.

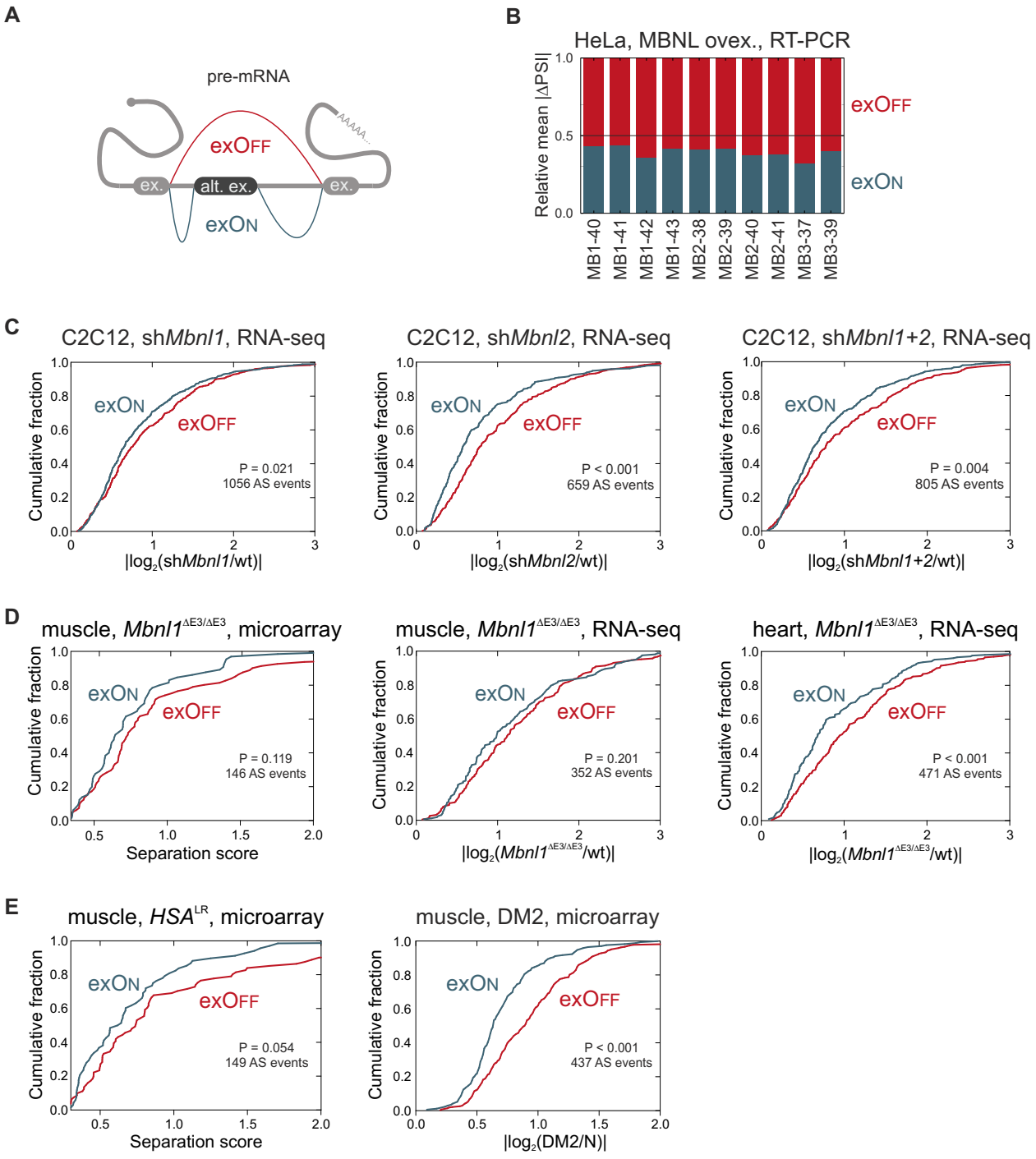
HeLa cells were transfected with 10 MBNL constructs. Splicing pattern of 38 AS events were analyzed by RT-PCR. To standardize alternative exon enumeration we applied numbering from FasterDB (1). For more information see Supplementary Table S2. As a measure of AS changes we used the PSI parameter.

(A) Splicing pattern of exons positively regulated by MBNL proteins (exON) and **(B)** exons negatively regulated by MBNLs (exOFF). Bars represent average PSI ± SD from three independent experiments and statistical significance was determined by the Student's t-test (∗ $P < 0.05$, ∗∗ $P < 0.01$, ∗∗∗ $P < 0.001$). White stars refer to statistical difference between particular MBNL isoforms and GFP control. Black stars refer to statistical differences between comparable MBNL isoforms having or not sequence encoded by alternative ex.54nt (matched by horizontal lines on bars), ex.36nt and ex.95nt (matched above bars).

References

(1) Mallinjoud P, Villemin JP, Mortada H, Polay Espinoza M, Desmet FO, Samaan S, et al. Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin. Genome Res. 2014;24:511-521. doi: 10.1101/gr.162933.113.

Supplementary Figure S3



Supplementary Figure S3

Higher impact of MBNL proteins on alternative exon exclusion (exOFF) than on exon inclusion (exON).

(**A**) Schematic representation of exOFF and exON.

(**B**) Comparison of relative mean for exON/exOFF splicing changes (|ΔPSI|) of 38 endogenous AS events in HeLa cells transfected with 10 different MBNL isoforms. Global analysis of exON and exOFF changes in (**C**) C2C12 with silencing of *Mbnl* paralogs determined by RNA-seq (1), (**D**) heart and skeletal muscles of *Mbnl1*-knockout mouse (*Mbnl1*$^{\Delta E3/\Delta E3}$) determined by RNA-seq and microarray (1,2), (**E**) skeletal muscles of DM1 mouse model (*HSA*^{LR}) and DM2 patients analyzed by microarrays (2,3). For all analyses the fold change was used as a measure of splicing strength for the indicated number of AS events. The statistical significance was determined by Mann-Whitney U test. All comparisons revealed higher average changes for exOFF than exON.

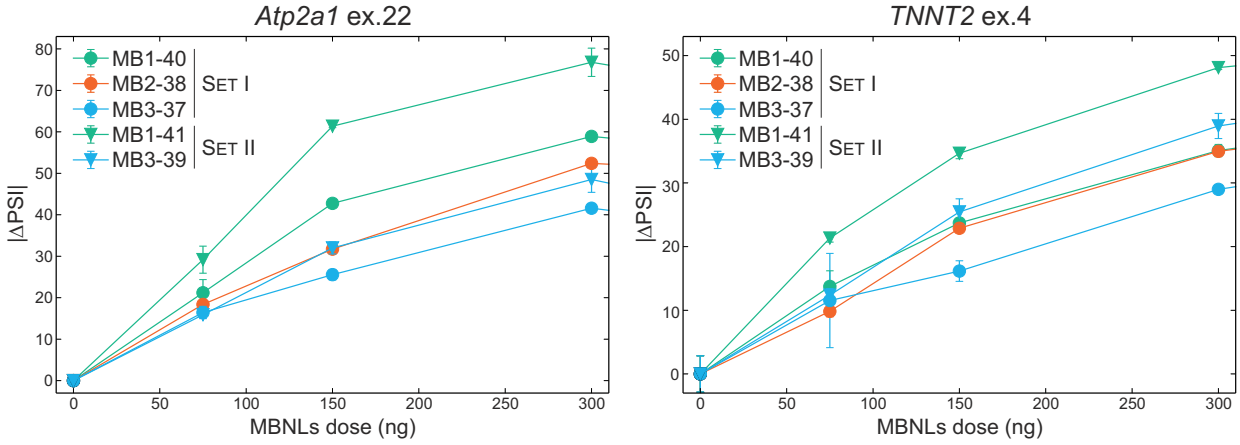
References

(1) Wang E, Cody N, Jog S, Biancolella M, Wang T, Treacy D, et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell. 2012;150:710-724. doi: 10.1016/j.cell.2012.06.041.

(2) Du H, Cline M, Osborne R, Tuttle D, Clark T, Donohue J, et al. Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. Nat Struct Mol Biol. 2010;17:187-193. doi: 10.1038/nsmb.1720.

(3) Nakamori M, Sobczak K, Puwanant A, Welle S, Eichinger K, Pandya S, et al. Splicing biomarkers of disease severity in myotonic dystrophy. Ann Neurol. 2013;74:862-872. doi: 10.1002/ana.23992.

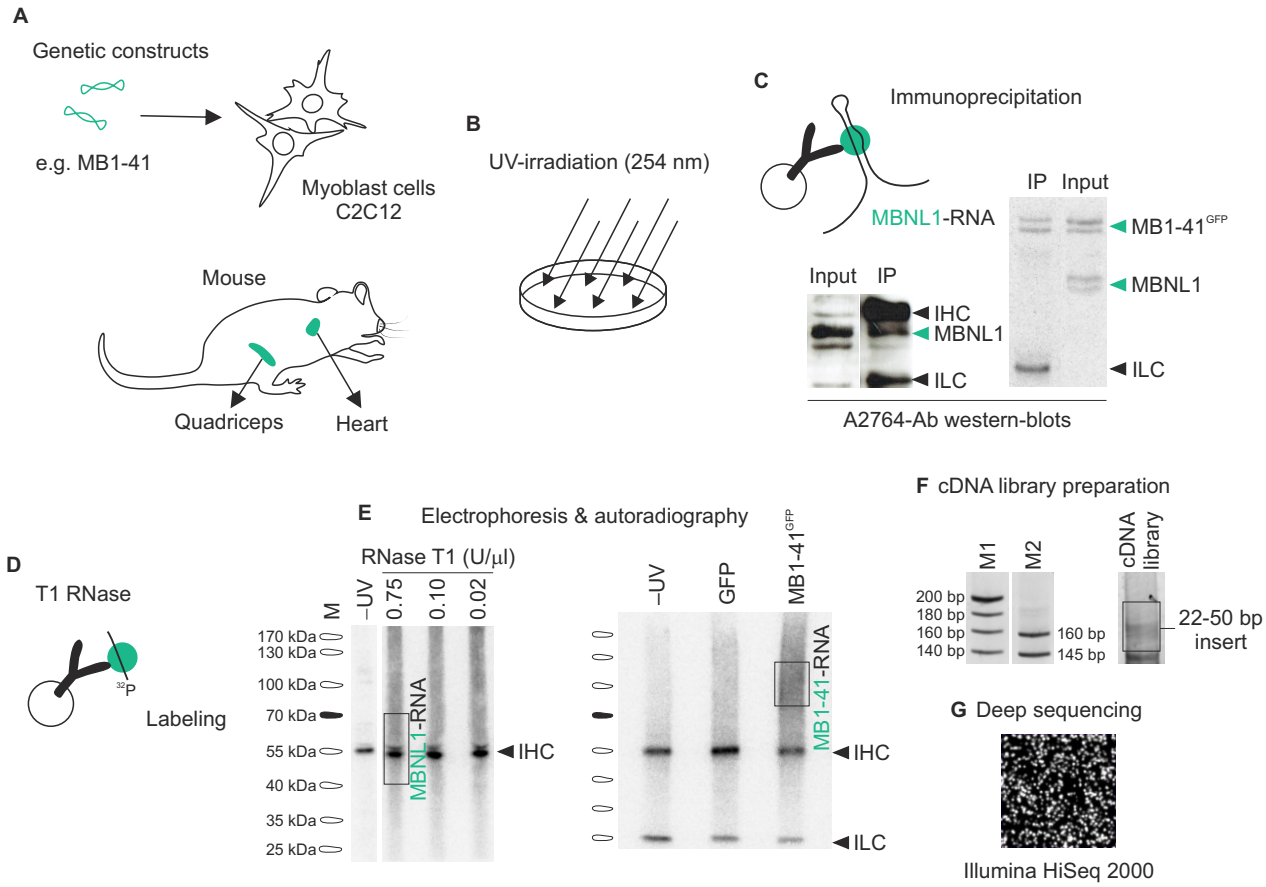
Supplementary Figure S4



Activity of different MBNL proteins from SET I and SET II on *Atp2a1* ex.22 and *TNNT2* ex.4 alternative splicing.

Bars represent average PSI ± SD from two independent experiments.

Supplementary Figure S5



CLIP-seq procedure for endogenous and exogenous MBNL1 study.

(A) C2C12 cells were transfected with exogenous MB1-40, MB1-41 and MB1-43 (Figure 1A). Heart and quadriceps were harvested from C57BL/6J mouse. Muscles were cut into thin slices.

(B) C2C12 and muscle slices were UV-irradiated.

(C) For endogenous MBNL1 and exogenous MBNL1 immunoprecipitation highly specific A2764 and anti-FLAG antibodies were used, respectively (IHC, ILC – Ig heavy/light chains).

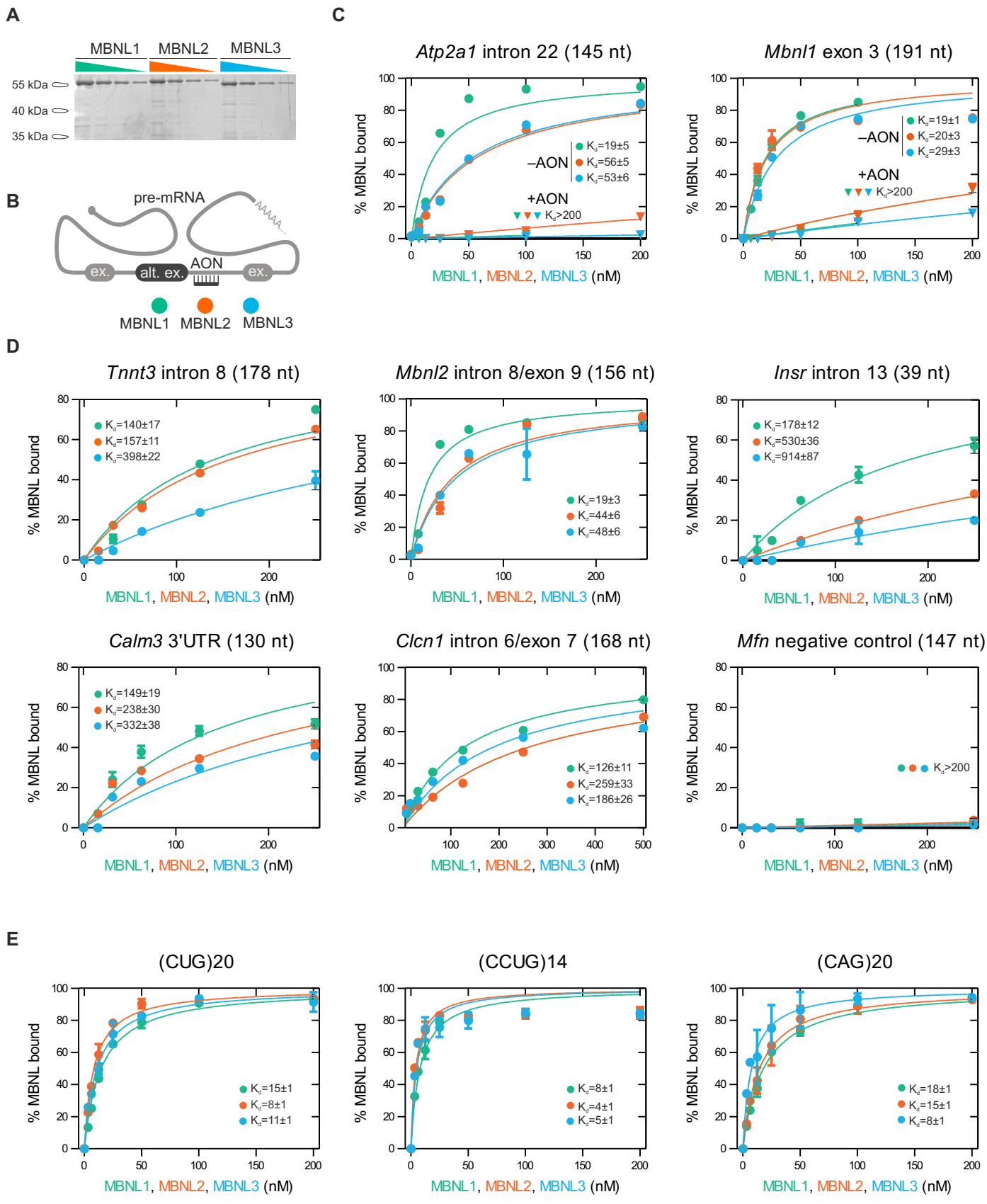
(D) Fixed complexes were treated with T1 RNase and RNA were radioactively labeled on 5' end.

(E) Complexes were separated on denaturing polyacrylamide gel and electrotransferred on membrane. After autoradiography RNA fragments were isolated from selected membrane part. M – ladder.

(F) cDNA libraries were prepared and **(G)** sequenced on Illumina HiSeq 2000.

For more information see Supplementary Material & Methods.

Supplementary Figure S6



Supplementary Figure S6

Comparison of affinity of three MBNL paralogs to selected RNA fragments.

(A) MBNL1, MBNL2 and MBNL3 recombinant proteins quantification on polyacrylamide gel.

(B) Experimental strategy of antisense oligonucleotide (AON) usage for masking the MBNL-binding sites.

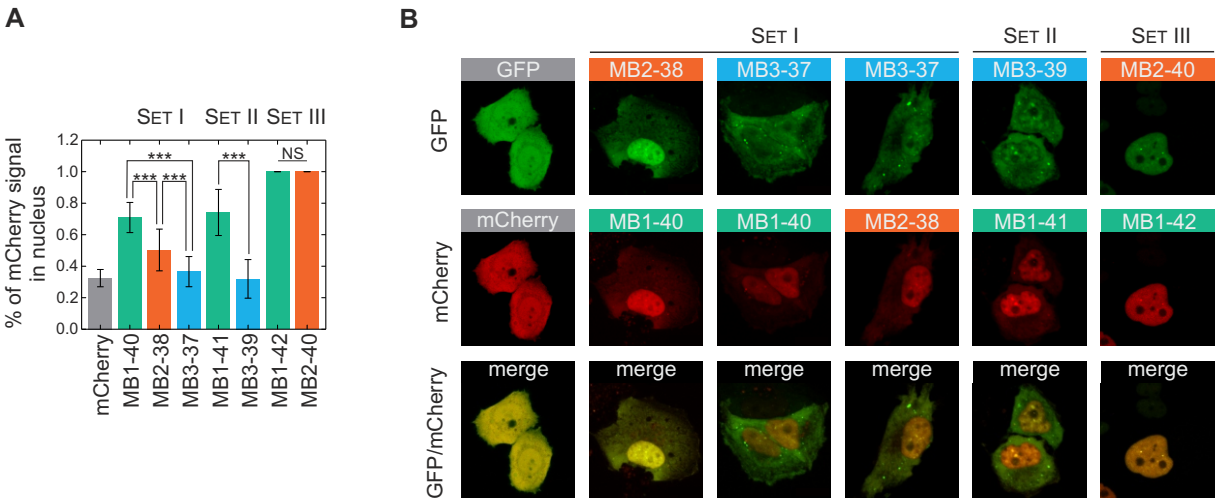
(C–E) Results of *in vitro* filter binding assays. For more details see Supplementary Material & Methods.

(C) Dissociation constants (K_d) of two selected RNA fragments, *Atp2a1* intron 22 and *Mbnl1* exon 3, significantly increase after blocking the MBNL1-specific binding site by short antisense oligonucleotide (+AON) compare to control experiments (−AON). For RNA and AON sequences see Supplementary Material & Methods.

(D) Different K_d for interaction between MBNL paralogs and selected RNA fragments known to be a target for MBNLs. Fragment of *Mfn* transcript was used as a negative control.

(E) MBNL paralogs have very high affinity to transcripts consisting of CUG, CCUG and CAG repeats.

Supplementary Figure S7

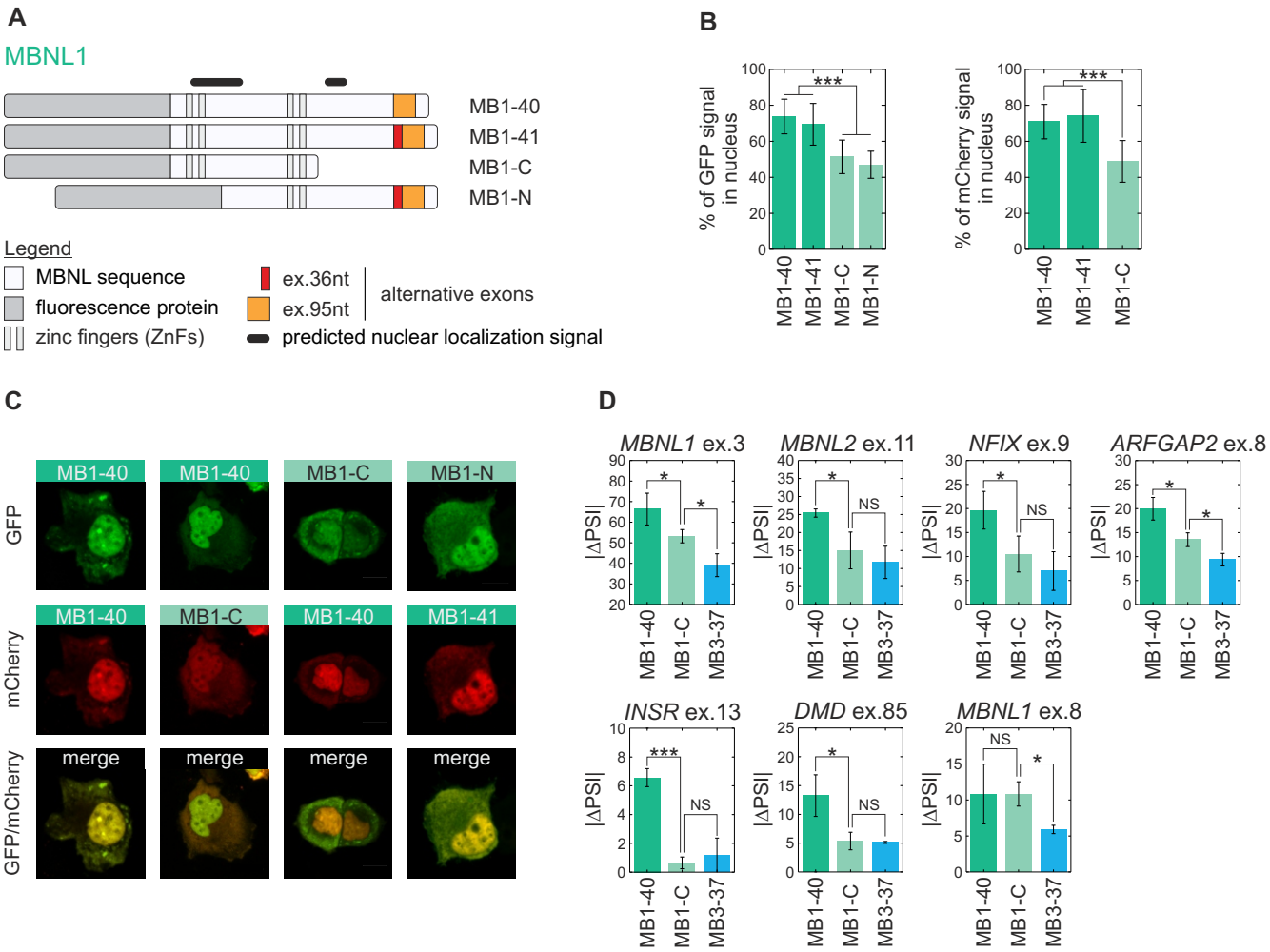


Distinct localization of MBNL paralogs in nucleoplasm and cytoplasm.

(A) The percent of nucleoplasmic signal for mCherry fused MBNL paralogs belonging to SET I, SET II and SET III were determined by quantitative confocal microscopy analysis of mCherry fluorescence signal. Bars represent mean from entire volume of about 50 cells. Statistical significance was assessed by the Student's t-test (NS for $P \geq 0.05$ and *** for $P < 0.001$).

(B) Distinct nucleoplasmic and cytoplasmic distribution pattern of isoform pairs visualized by coexpression of MBNL proteins fused with GFP and mCherry (SET I and SET II). Isoforms having ex.54nt localize exclusively in the nucleus (SET III). Free GFP and mCherry proteins have an equal distribution pattern.

Supplementary Figure S8



Truncated MBNL1s.

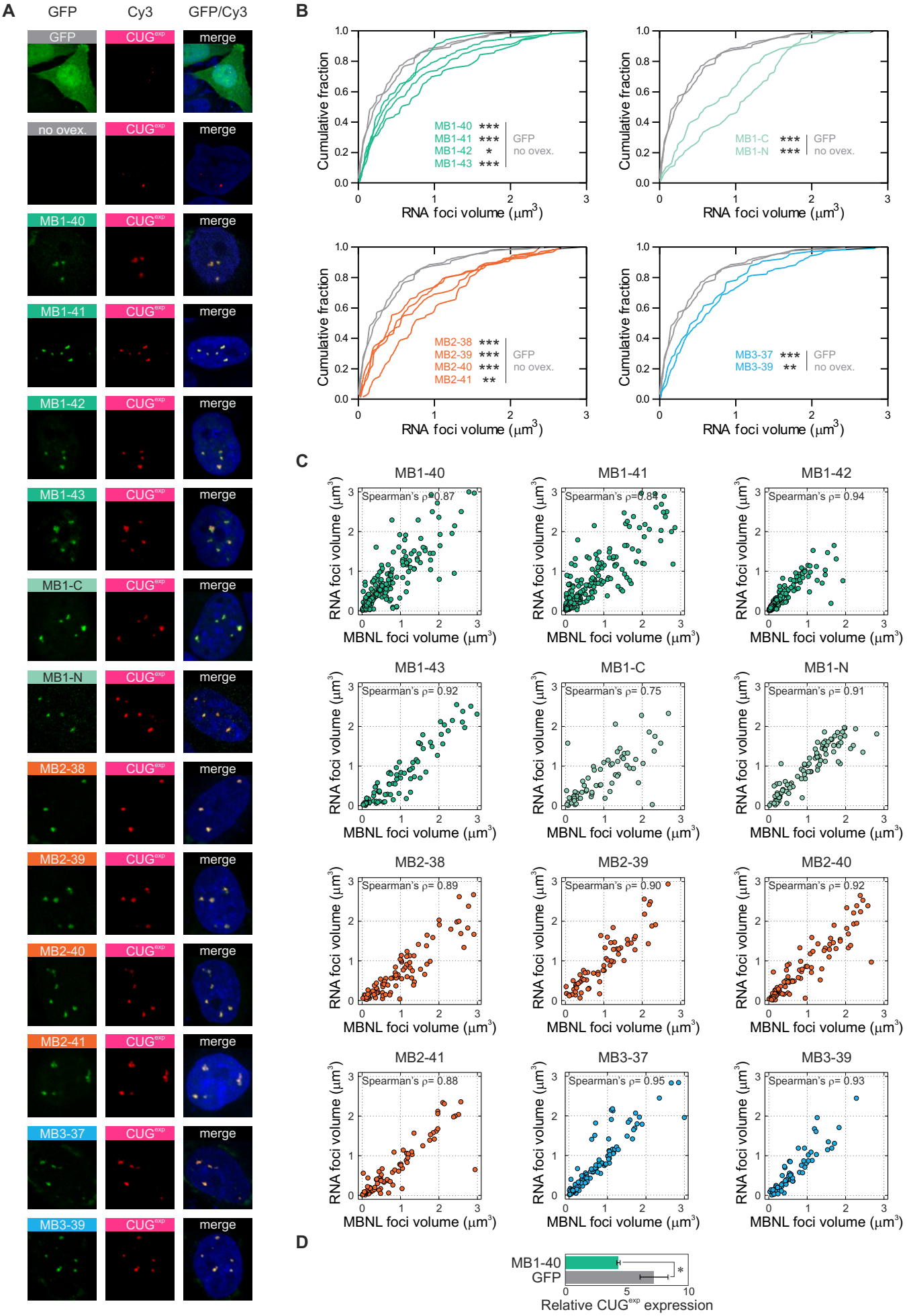
(A) Scheme of expression constructs containing sequences of fluorescence protein (GFP, mCherry) and two full length (MB1-40, MB1-41) and two truncated MBNL1 proteins without C-terminal domain (MB1-C) and without first ZnF tandem domain located on N-terminus (MB1-N). Both, MB1-C and MB1-N lacked predicted nuclear localization signals.

(B) Reduced nucleoplasmic distribution of both MB1-C and MB1-N compared to full length MBNL1 isoforms. Cellular distribution was determined by a quantitative confocal microscopy analysis of GFP and mCherry fluorescence signal. Bars represent means and standard deviations derived from analysis of about 100 cells. Statistical significance was assessed by the Student's t-test (NS for $P \geq 0.05$ and *** for $P < 0.001$).

(C) Distinct nucleoplasmic and cytoplasmic distribution pattern visualized by coexpression of MBNL isoforms fused with GFP and mCherry.

(D) For seven AS events the MB1-C reveals similar splicing activity to MB3-37 than MB1-40. Selected AS events in previous analysis had significantly different |ΔPSI| values for MBNL1 and MBNL3. Represented means and standard deviations were derived from three biological replicas. Statistical significance was assessed by the Student's t-test (NS for $P \geq 0.05$ and *** for $P<0.001$).

Supplementary Figure S9



Supplementary Figure S9

MBNL proteins influence the CUG^{exp} foci volume.

Combined analysis of signals from RNA foci visualized by fluorescent *in situ* hybridization (FISH, Cy3 labeled probe) and GFP signal.

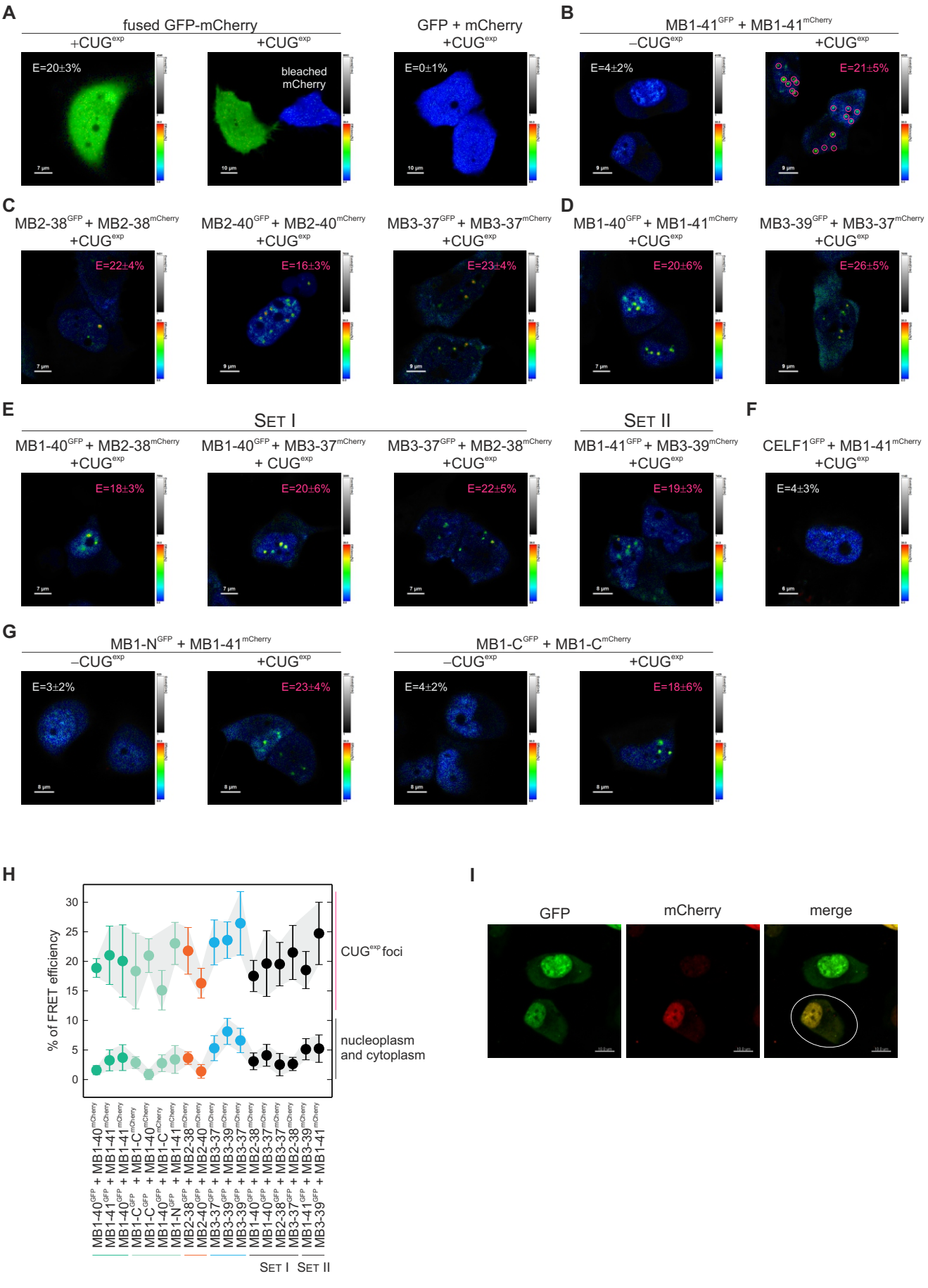
(A) All tested MBNL isoforms colocalize with CUG^{exp} in discrete foci.

(B) Overexpression of MBNL paralogs significantly increase CUG^{exp} foci volume measured by FISH for ~130 foci. There is no difference in CUG^{exp} foci volume distribution between cells transfected with GFP and without protein overexpression (no ovex.). Statistical significance was assessed by Mann-Whitney U test ($*$ for $P < 0.05$, $**$ for $P < 0.01$ and $***$ for $P < 0.001$).

(C) Correlation between RNA foci volume with the volume of MBNL isoforms (measured as GFP signal). The Spearman's ρ is between 0.75 to 0.95. FISH and GFP fluorescence measurements were performed using quantitative confocal microscopy.

(D) Exogenous CUG^{exp} transcript expression/stability is not elevated after MBNL1 overexpression. HeLa cells were cotransfected with CUG^{exp} encoding transcript with MB1-40 or GFP. Bars represent average expression level ± SD measured by semi-quantitative RT-PCR from two independent experiments. Statistical significance was determined by the Student's t-test ($*$ for $P < 0.05$).

Supplementary Figure S10



Supplementary Figure S10

Comparative analysis of FRET efficiency (E) between GFP and mCherry fused with different MBNL proteins.

(**A**) GFP-mCherry fusion protein with nine amino acid linker show ~20% E and after mCherry photobleaching there is no FRET. No FRET is also observed in cells cotransfected with GFP and mCherry constructs.

(**B**) E value for the pair of full length MB1-41 in CUG^{exp} foci is as high as in positive GFP-mCherry control. In absence of CUG^{exp} E value measured in nucleoplasm or cytoplasm is slightly above background. High E is also observed in foci for (**C**) homologous and (**D**) heterologous pairs of MBNL paralogs as well as (**E**) pairs of MBNL paralogs in SET I and SET II.

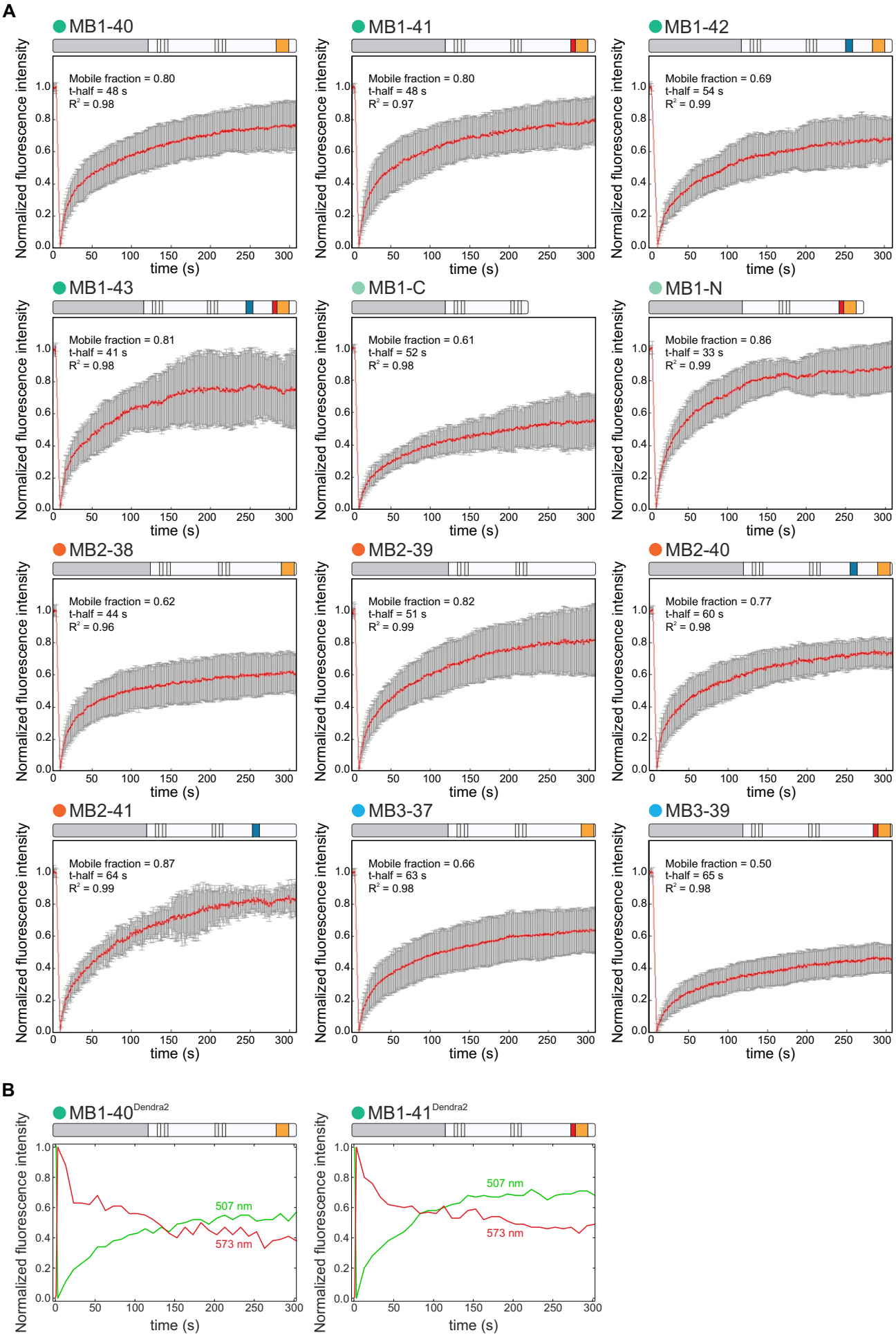
(**F**) There is no FRET in foci of cells cotransfected with MB1-41 and CELF1 which do not colocalize with CUG^{exp}. For this pair of proteins observed E value is as low as in cells without CUG^{exp}.

(**G**) E values for truncated MBNL1 (MB1-C and MB1-N) that are lacking in sequences important for multimerization or without two ZnFs are similar to those observed for full length MBNL proteins.

(**H**) Quantification of data from all experiments. Represented means and standard deviations were derived from 10-20 FLIM measurements.

(**I**) Coexpression of MBNL proteins fused with GFP and mCherry visualized by confocal microscopy. Only selected cells with equal fluorescence emission of both GFP and mCherry labeled MBNL isoforms were used for FRET experiments.

Supplementary Figure S11



Supplementary Figure S11

Differences in mobility of MBNL isoforms between CUG^{exp} foci and nucleoplasm based on fluorescence recovery after photobleaching (FRAP) experiments.

(**A**) All tested MBNL isoforms are mobile, but their mobile fractions are different (0.50-0.87). T-half values are also various (33-65 s). FRAP experiments were performed for ~23 cells for each model. We analyzed a few foci in each nucleus. Raw data were analyzed with easyFRAP software (1). We applied full scale normalization and single exponential fitting equation ($R \geq 0.96$). See Supplementary Video 3.

(**B**) Photoswitching of Dendra2 fused with MBNL1 isoforms. Photoconverted (red, emission 573 nm) protein is passing out whereas unconverted (green, emission 507 nm) protein flow into the foci. Represented results were derived from five independent experiments. See Supplementary Video 4.

References

(1) Rapsomaniki MA, Kotsantis P, Symeonidou IE, Giakoumakis NN, Taraviras S, Lygerou Z. easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. Bioinformatics. 2012;28:1800-1801. doi: 10.1093/bioinformatics/bts241.

Supplementary

Material & Methods

(A) List of primers for genetic constructs generation.

Primers for MBNL2 constructs generation.

1. GCGAT**GAATTC**GGCTTTGAACGTTGCCCCAGTCA (**EcoRI**, forward for MBNL2)
2. CGTTTCTGC**GGATCC**TTTCAGAATTAT (**BamHI**, reverse for MB2-38, MB2-40)
3. ACTACATCTT**GGATCC**GCATGCAGTT (**BamHI**, reverse for MB2-39, MB2-41)

Primers for MBNL3 constructs generation.

1. GCGAT**GAATTC**GACGGCTGTCAATGTTGCCCTGAT (**EcoRI**, forward for MBNL3)
2. TAACTCTGC**GGATCC**GAATTTCAGCTGAT (**BamHI**, reverse for MBNL3)

Primers for mCherry constructs generation.

1. GCGAT**GCTAGC**ATGGTGAGCAAGGGCGAGGA (**NheI**, forward for MBNL1, MBNL2 and MBNL3)
2. ATCGC**CTCGAG**CCTTGTACAGCTCGTCCATGC (**XhoI**, reverse for MBNL1)
3. ATCGC**AAGCTT**TCTTGTACAGCTCGTCCATGC (**HindIII**, reverse for MBNL2 and MBNL3)
4. GCGAT**AAGCTT**TGGTGAGCAAGGGCGAGGAG (**HindIII**, forward for EGFP-mCherry)
5. ATCGC**GAATTC**CTTGTACAGCTCGTCCATGC (**EcoRI**, reverse for EGFP-mCherry)

Primers for Dendra2 constructs generation.

1. GCGAT**GCTAGC**ATGAACACCCCGGGAATTAA (**NheI**, forward for MBNL1)
2. ATCGC**CTCGAG**CCTGGCTGGGCAGGGGGCTG (**XhoI**, reverse for MBNL1)

Primers for recombinant MBNL2 generation.

1. TCGAT**GGATCC**TCGGCTTTGAACGTTGCCC (**BamHI**, forward)
2. TTTCTG**GAATTC**CCATGACTGTGGCCGCGG (**EcoRI**, reverse)

Primers for recombinant MBNL3 generation.

1. TCGAT**GGATCC**TCGACGGCTGTCAATGTTGC (**BamHI**, forward)
2. TTTCTG**GAATTC**CCATGGCAGAGGCAGCTGAA (**EcoRI**, reverse)

Conditions for all Pfx50 (Invitrogen) PCRs

94°C/2 min; 35 cycles {94°C/15 s; 60°C/15 s; 68°C/45 s} 68°C/5 min.

(B) RNA fragments and oligonucleotides used in filter binding assays.

***Atpta1* intron 22** (mouse)

Primer set for long product amplification (Ta = 58°C)

Forward: GGTGGGATTTTCTCCCAACCT

Reverse: GCTGTATCTCCTTGCGCATC

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGGCACCCCTGCGGTCACT

Reverse: AGCAGACTGGACGGCCA

Figure 3C and Supplementary Figure S6 *in vitro* transcript (145 nt)

gggcaccccugcggucacuugcgcgcgcucgcagcucucccuggagccacu**gcugcuguugccacugcccgc**uuccacaugagggccaugccguuacugugggcacuggcagcugugaggcuguguccuggccguccagucugcu

Antisense oligonucleotide (AON)

gcgggcagtggcaacagcagc

***Mbnl1* exon 3** (mouse)

Primer set for long product amplification (Ta = 60°C)

Forward: GATGGCTGGCTGCAATATGCC

Reverse: CCTAGGGCAATGGCAGATACTC

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGTTGGTACTAAGAAGTGCCT

Reverse: CCTAGGGCAATGGCAGATACTC

Figure 3C and Supplementary Figure S6 *in vitro* transcript (191 nt)

GGGUUGGUACUAAGAAGUGCCUUUCCUGACGUCUCUGCUGCUUGGAACCGCUUCUAGAGCAGCCUC**UGCUUUUGCCUUGCUUGCUGC**CAGCUAGACUGACGACAGCACAUCCGCCCUCCACCUCUAGCCCAGACACCCCAUUUCUACUUCUAAUCAGGAGAAAAGCUCUGAGUAUCUGCCAUUGCCCUAGG

Antisense oligonucleotide (AON)

gcagcaagcaaggcaaaagca

***Mbnl2* intron 8/exon 9** (ex.54nt) (mouse)

Primer set for long product amplification (Ta = 60°C)

Forward: TGTGTATGCAAGCTGCTGTT

Reverse: ACAAATCTCGGCTGGTTCTCT

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGGTTCCCATGATGCACCTCTGC

Reverse: CAGGTCTACAGTTGCTTCGA

Figure 3C and Supplementary Figure S6 *in vitro* transcript (156 nt)

ggguucccaugaugcaccucugcuugcuguuuaccuguauguuaauucgcuugaaucccauuggcccacugcc
aucaugugcucgcugccuguuauuaaaagACUCAGUCGACUGCCAAAGCACUGAAGCG
ACCUCUCGAAGCAACUGUAGACCUG

Figure 3E (**wt**) *in vitro* transcript (92 nt)

gggucguuugaaucccauuggcccacugccaucaugugcucgcugccuguuauuaaaagACUCAGUCG
ACUGUCAAAGCACUGAAGCGACCa

Figure 3E **mut#1** *in vitro* transcript (92 nt)

gggucguuugaaucccauuggcccacugucaucaugugcucgcugccuguuauuaaaagACUCAGUCG
ACUGUCAAAGCACUGAAGCGACCa

Figure 3E **mut#2** *in vitro* transcript (92 nt)

gggucguuugaaucccauuggcccacugccaucaugugaucgaugacuguuauuaaaagACUCAGUCG
ACUGUCAAAGCACUGAAGCGACC

Tnnt3 **intron 8** (mouse)

Primer set for long product amplification (Ta = 60°C)

Forward: GAAGGTAGAGACCGGCTTCG

Reverse: GTCACTGAGGAACCATGCCA

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGCGTGCCTGCCTCTGCTGTT

Reverse: TTCCTCGTCCTCCTCCCGCTCC

Figure 3C and Supplementary Figure S6 *in vitro* transcript (178 nt)

ggacagcaaguauccaugaaggggaggagaccuaugggaagguguaggagcugcccugccacccucaggcagg
gggcucugugcgugccugccucugcuguuccauguggccacugcugcuguguggccucugcacuucugcagc
ugcguggcuuuuuucuugcaugugcgcuugugc

Calm3 **3´UTR** (mouse)

Primer set for long product amplification (Ta = 60°C)

Forward: GACTGACGGACTCAACCCTG

Reverse: CTGAGCAGGGACAGGAATGG

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGGACTGAGAATCTGATAAAGC

Reverse: GGAGACTGGGAATGGAAGGG

Figure 3C and Supplementary Figure S6 *in vitro* transcript (130 nt)

GGGACUGAGAAUCUGAUAAAGCACCAAAAGAUUUGUCCCAAGCUGCAU
GACUGCUCUCUCUCCUUCCUCCUGACUGUCCCUCCACGCCCUCACCGCU
UCCUUUGGCCUUCCCCUUCCAUUCCCAGUCUCC

Figure 3E (**wt**) *in vitro* transcript (91 nt)

ggGCACCAAAAGAUUUGUCCCAAGCUGCAUGACUGCUCUCUCUCCUUCC
UCCUGACUGUCCCUCCACGCCCUCACCGCUUCCUUUGGCCCC

Figure 3E (**mut#1**) *in vitro* transcript (91 nt)

ggGCACCAAAAGAUUUGUCCCAAGCUGCAUGACUGGUCUCUCUCCUUCC
UCCUGACUGUCCCUCCACGCCCUCACCGCUUCCUUUGGCCCC

Figure 3E (**mut#2**) *in vitro* transcript (91 nt)

ggGCACCAAAAGAUUUGUCCCAAGCUGCAUGACUGCUCUCUCUCCUUCC
UCCUGACUGUCCCUCCACGCCCUCACCACUUCCUUUGGCCCC

Figure 3E (**mut#3**) *in vitro* transcript (91 nt)

ggGCACCAAAAGAUUUGUCCCAAGGUGGAUGACUGCUCUCUCUCCUUCC
UCCUGACUGUCCCUCCACGCCCUCACCGCUUCCUUUGGCCCC

***Clcn1* intron 6/exon 7** (mouse)

Primer set for long product amplification (Ta = 58°C)

Forward: CCTTTCCATGTTTCCTCCTGTG

Reverse: ACCAAGGTAGGGAGGAAGTG

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGCCCCCGTTCTTCTGTG

Reverse: AAAGTAGCATCCCACGCCCA

Figure 3C and Supplementary Figure S6 *in vitro* transcript (168 nt)

ggcccccgttcttctgtgcttcctgacacccatccacctggtttacataccacctgtctgtccccctctgccacctgcctcgccc
gtcgtgcttctctgttgcagACCGTGCCTGGGCAGCTTGATCTCCTGGTGCCAGCCTGT
GCAGTGGGCGTGGGATGCTACTTT

***Insr* intron 13** (mouse)

DNA template

GCTCTCCAGCACAGCTGCCCGCCGCATGCAAAAAGCCCTATAGTGAGTCG
TATTA

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGG

Reverse: GCTCTCCAGCACAGCTG

Figure 3C and Supplementary Figure S6 *in vitro* transcript (39 nt)

gggcuuuuugcaugcggcgggcagcugugcuggagagca

negative control *Mfn* (mouse)

Primer set for long product amplification (Ta = 58°C)

Forward: TGTAGGGCTAGGACTTGGGG

Reverse: GCAGTCAGGGCTCAAAGGTA

Primer set for promoter containing template generation:

Forward: TAATACGACTCACTATAGGGCACCTATCAGTGACCTTTTTC

Reverse: GCAATGTTGGTTAGTAACCC

Figure 3C and Supplementary Figure S6 *in vitro* transcript (147 nt)

gggcaccuaucagugaccuuuuuccuucuaccaagucacuuguggggcuggguacuacuuaucuccuccagcug accuguuuauucuguuuccaugguagcagguguaaggcaugcaggauguuaguaggguuacuaaccaacauu gc

Conditions for all GoTaq Flexi DNA Polymerase (Promega) PCRs

94°C/2 min; 30-32 cycles {94°C/20 s; 55°C/20 s; 72°C/20 s} 72°C/5 min.

(C) Antisense oligonucleotides (AONs).

AONs used in cellular experiments were synthesized by RiboTask.

(AON) 2´-OMePS-*Atp2a1*	GCGGGCAGUGGCAACAGCAGC
(AON) 2´-OMePS-*Nfix*	CCAGCAAGCACAGGCAGCGGG
(AON) 2´-OMePS-*Ldb3*	CAGAAGCAGGCAGCAGCGGGG
(Ctrl.) 2´-OMePS-control	GGAGAAGCAAAAAGCAAGAGA

(D) Step-by-step protocol for Mbnl1-CLIP-seq on mouse tissues.

Procedure was developed based on the original HITS-CLIP protocol (1, 2).

1. Mouse tissue preparation and UV-crosslinking.

- Dissect skeletal muscle (quadriceps) and/or heart from 3-mounth old C57BL/6J buck mouse. Immediately freeze the samples in liquid nitrogen.
- Cut tissue into 60 μm slices and put into cryostat at -23°C. Tissue slices plate on 10 cm culture dishes with 25 ml frozen buffer: 1X HBSS (Gibco, cat no. 14185) and 10 mM HEPES pH 7.3 (BioShop, cat no. HEP003).
- Place the culture dish with frozen buffer directly on UVP CL-1000 device. Crosslink RNA-protein complexes within tissue slices by UV irradiation with 400 mJ/cm^2 at 254 nm.
- Culture dishes with crosslinked biological material should be stored at -80°C.

2. Magnetic bead preparation.

- Resuspend Dynabeads Protein G (Invitrogen, cat no. 10003D) and transfer 50 μl of solution to a fresh 1.5 ml tube.
- Wash beads three times with 100 mM Na-citrate pH 5.0. To separate beads keep tube 1 minute on a magnet (Dynal MPC).
- Resuspend beads in 50 μl 100 mM Na-citrate with 2.5 μl A2764 anti-Mbnl1 antibody (gift from Charles Thornton).
- Mix gently tube at 4°C, O/N.

3. Tissue lysis.

- Thaw buffer with tissue slices at 4°C and transfer to 50 ml conical tube.
- Centrifuge at 400 × g, 10 min, 4°C (Beckman, Allegra X-30R, rotor CO650).
- Discard supernatant, resuspend tissue pellet in 700 μl ice-cold PBS ant transfer to a clean 1.5 ml tube.
- Centrifuge at 400 × g, 10 min, 4°C (rotor F2402H)
- Discard supernatant, resuspend pellet in 400 μl ice-cold lysis & immunoprecipitation buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM EDTA pH 8.0, 0.1% NP-40, 0.05% Tween-20, 1 mM PMSF*, 2 mM benzamidine*). * add fresh.
- Homogenize resuspended pellet in glass Dounce homogenizer using loose (20 moves) and tight (35 moves) pestles respectively. Transfer homogenous suspension to a clean 1.5 ml tube.
- Centrifuge at 18,000 × g, 10 min, 4°C.

- Transfer supernatant to a clean 1.5 tube, add 10 µl of DNase RQ1 (Promega, cat no. M6101) and 10 µl of RNasin (Promega cat no. N2511).

4. Immunoprecypitation.

- Coated beads with A2764 antibody (step 2) wash three times with buffer: 0.1 mM Na-citrate pH 5.0/0.05 % Tween-20. Transfer beads to a clean 1.5 ml tube.
- Add prepared homogenous suspension (step 3) to washed coated beads.
- Rotate 15 rpm on Stuart rotator SB3 (BioCote) for 2 h at 4°C.
- Short spin the tube (Eppendorf MiniSpin) to remove liquid from the cup. Separate beads from the magnet (1 min) and discard supernatant.
- Wash beads five times with 700 µl of PTM buffer (PBS/0.05% Tween-20/1 mM $MgCl_2$). Each time use Stuart rotator, short spin and then separate on a magnet.
- Transfer beads with RNA-Mbnl1 complexes to a clean 1.5 ml tube.

5. T1 RNase digestion.

- Resuspend beads in 50 µl of PTMCa buffer (PBS/0.05% Tween-20/1 mM $MgCl_2$/1 mM $CaCl_2$*) with 0.1 U/µl RNase T1 (Sigma, cat no. R1003). Add 5 µl of DNase RQ1. * RQ1 requires Ca^{2+}.
- Incubate for 20 min. at 25°C in Thermomixer (Eppendorf) with 500 rpm shaking.
- Separate the beads on a magnet, remove buffer, wash the beads with PTE buffer (PBS/0.05% Tween-20/5 mM EDTA pH 8.0), 1 min rotate on Stuart rotator, short spin on centrifuge.

6. BAP dephosphorylating of 3´-RNA edns.

- Prepare 50 µl of dephosphorylation solution: 36 µl H_2O, 5 µl Dephosphorylation Buffer (100 mM Tris-HCl pH 8.0), 4 µl BAP (150 U/µl, Invitrogen cat no. 18011-015), 2 µl Rnasin.
- Add dephosphorylation solution to beads (step 5). Incubate 20 min at 25°C in a Thermomixer with 500 rpm shaking.
- Separate beads on a magnet, remove buffer.
- Wash beads three times with PTE buffer, transfer them to a clear 1.5 ml tube.

7. Radioactive labeling of 5´-RNA ends.

- Prepare the 50 µl reaction solution for radioactive labeling: 32.5 µl H_2O, 5 µl 10× OptiKinase Reaction Buffer (500 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 50 mM DTT), 2 µl Rnasin , 1 µl [γ-^{32}P]ATP (Hartman Analytics, 5000 Ci/mmol, 10 mCi/ml), 2 µl Opti Kinase (10 U/µl, USB 78334Y). The volume of beads is about 7.5 µl.
- Separate beads on a magnet, remove buffer (step 6).

- Add radioactive labeling reaction solution, incubate 20 min at 25°C in Thermomixer with 500 rpm shaking.
- Add 5 µl of cold ATP (10 mM Invitrogen cat no. 18330-019), incubate additional 5 min.
- Separate beads on a magnet, wash beads two times with PTE buffer, transfer beads to a tube.

8. RNA-protein complexes denaturation.

- Separate beads on a magnet, remove radioactive solution, short spin on centrifuge, again separate beads on a magnet, remove the rest of radioactive solution, and short spin to sit beads on the bottom of the tube.
- Add 31 µl of Sample Loading Buffer (1× NuPAGE MOPS-SDS buffer (cat no. NP0001), 20% glycerol, 2% β-mercaptoethanol, 2% SDS, 1 mM EDTA pH 8.0, 0.002% Briliant Blue), pipet 6-8 times.
- Heat sample at 95°C for 1 min.
- Cool down sample on ice, short spin, separate on a magnet, transfer Sample Loading Buffer that contain denatured RNA-Mbnl1 complexes to a clean tube.

9. Electrophoresis and elektrotransfer.

- Prepare 800 ml 1× MOPS-SDS buffer: 40 ml 20× NuPAGE MOPS-SDS, 760 ml H2O.
- Run the sample on gradient 4-12% NuPAGE gel (cat no. NP0321) in XCell SureLock MiniCell. Use 200 ml of 1× MOPS running buffer with 0.5 ml NuPAGE Antioxidant (NP0005) in upper chamber and 600 ml in lower chamber. Load 3 µl of the PageRuler pre-stained protein size marker (Fermentas, cat no.SM0671).
- Run the gel for 15 min 50V and then ~1h at 200 V.
- Transfer the protein-RNA complexes from the gel to a Whatmann OPTITRAN BA-S 85 nitrocellulose membrane (Sigma cat no. 8418878) using the XCell II Blot Module wet transfer apparatus. Transfer 1 h at 30 V, 300 ml buffer:15 ml NuPAGE Transfer Buffer (20×, cat no. NP0006), 255 ml H_2O, 30 ml methanol, 0.3 ml Antioxidant.
- After the transfer, two times rinse the membrane in PBS buffer, then wrap it in saran wrap and expose it to the screen O/N.

10. Proteinase K digestion.

- Isolate the Mbnl1-RNA complexes of the appropriate size using the autoradiograph as a mask (40-70 kDa). Cut this piece of membrane into several small slices and place them into the 2.0 ml tube.
- Prepare 600 µl Proteinase K (Sigma, cat no. P2308) solution (2 mg/ml). Mix 120 µl Proteinase K stock solution (10 mg/ml) with 180 µl H_2O and 300 µl 2x buffer (2 mM $CaCl_2$, 1% SDS,

200 mM Tris-HCl pH 7.5, 100 mM NaCl). Incubate at 37°C for 30 min at 500 rpm (Thermomixer) to remove RNases.

- Add EDTA pH 8.0 to final 5 mM concentration. Incubate for 3 min.
- Add 300 µl of such prepared 2 mg/ml Proteinase K solution to the nitrocellulose slices. Incubate at 25°C for 30 min at 1000 rpm (Thermomixer).

10. RNA isolation.

- Transfer the solution (without the nitrocellulose) to a new 1.5 ml tube and add 1 ml of Tri Reagent (Sigma, cat no. 93289).
- Add 0.2 ml chloroform and mix.
- Store for 15 min at RT.
- Centrifuge 20,000 × g for 15 min at 4°C.
- Transfer aqueous phase to a clean tube.
- Add 0.5 ml isopropanol and 1 µl glycogen (USB, cat no. 16445) and mix.
- Store at -20°C O/N.
- Centrifuge 20,000 × g for 8 min at 4°C.
- Wash the pellet twice with 1 ml 75% ethanol.
- Centrifuge 7,500 × g for 5 min at 4°C.
- Wash the pellet twice with 0.1 ml 75% ethanol.
- Centrifuge 7,500 × g for 5 min at 4°C.
- Air dry the pellet.
- Dissolve the RNA pellet in 5.5 µl.

11. Library preparation.

Proceed with TruSeq Small RNA Preparation Guide with the modifications below:

- Reduced the volume of PCR to 25 µl.
- Applied 20 cycles.
- From the 6% PAA gel select part from 145 to about 200 bp.

(E) Step-by-step Protocol for MBNL1-CLIP-seq in C2C12 myoblasts.

Procedure was developed based on original iCLIP protocol (3).

1. C2C12 cell transfection.

- Split C2C12 cells on 10 cm culture dish 24 hrs before transfection.
- Transfect cells with using X-tremeGENE HP DNA Transfection Reagent (Roche). Prepare 800 ul of serum- and antibiotic-free DMEM with 16 μg of DNA. Add 48 μl of transfection reagent (1:3 ratio). Incubate 30 min at RT.
- Add the DNA/ transfection reagent mix to 80% confluent C2C12 cells.
- Visualize cells after 24-30 h.

2. UV cross-linking of tissue culture cells.

- Remove the media and add 1 ml ice-cold PBS to cells grown in a 10 cm plate.
- Remove lid and place the dish on a tray with ice. UV cross-link cells with 150 mJ/cm^2 at 254 nm UV Crosslinker (UVP CL-1000).
- Harvest the cells by scraping with a cell lifter.
- Transfer the cell suspension to a 1.5 ml tube. Spin at top speed for 10 sec at 4°C to pellet cells then remove the supernatant.
- Freeze the cell pellet at -80°C until use.

3. Magnetic beads preparation.

- Add 5 μl of protein G Dynabeads per experiment to a clean 1.5 ml tube.
- Wash beads two times with lysis buffer: 50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate.
- Resuspend beads in 100 μl of lysis buffer with 1 μl (μg) of FLAG-M2 antibody (Sigma, cat no. F1804).
- Rotate tubes at RT for 60 min.
- Wash three times with 500 μl of lysis buffer and leave in the last wash until ready to proceed to the next step.

4. Cell lysis.

- Resuspend the cell pellet in 0.5 ml of lysis buffer supplemented with fresh 1 mM PMSF and 2 mM benzamidine.
- Spin at 4°C and 15,000 × g for 15 min to clear the lysate. Carefully collect the supernatant (leave about 50 μl lysate with the pellet).

- Add 1 µl TURBO DNase (Ambion, cat no. AM2238) and 12.5 µl Rnasin (for 0.5 ml).

5. Immunoprecipitation.

- Remove the wash buffer from the beads then add the cell lysate.
- Rotate the samples for 2 h at 4°C.
- Discard the supernatant and wash the beads two times with 500 µl of 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.
- Wash three times with 900 µl wash buffer: 20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20.

6. Next steps perform according to protocol of Mbnl1 CLIP-seq on mouse tissues.

(F) Step-by-step protocol for Mbnl1-RIP-seq.

1. **Mouse tissue preparation.**
   - Dissect skeletal muscle (quadriceps) from 3-mounth old C57BL/6J buck mouse. Immediately freeze the sample at liquid nitrogen.

2. **Magnetic bead preparation.**
   - Resuspend Dynabeads Protein G (Invitrogen, cat no. 10003D) and transfer 75 µl of solution to a fresh 1.5 ml tube.
   - Wash beads three times with 100 mM Na-citrate pH 5.0. To separate beads keep the tube 1 minute on a magnet (Dynal MPC).
   - Resuspend beads in 75 µl of 100 mM Na-citrate with 3.75 µl A2764 anti-Mbnl1 antibody (gift from Charles Thornton, University of Rochester).
   - Mix gently at 4°C, O/N.

3. **Tissue lysis.**
   - Muscle tissue triturate with liquid nitrogen in chilled mortar.
   - Tissue powder transfer to glass Dounce homogenizer. Add 0.5 ml ice-cold lysis & immunoprecypitation buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM EDTA pH 8.0, 0.1% NP-40, 0.05% Tween-20, 1 mM PMSF*, 2 mM benzamidine*). * add fresh. Homogenize tissue using loose (20 moves) and tight (35 moves) pestles respectively.
   - Transfer the solution to a clean 1.5 ml tube and centrifuge 20,000 × g for 10 min at 4°C.
   - Transfer supernatant to a clear 1.5 ml tube and again centrifuge 20,000 × g for 5 min at 4°C.
   - Add add 10 µl of DNase RQ1 (Promega, cat no. M6101) and 10 µl of RNasin (Promega cat no. N2511).

4. **Immunoprecypitation.**
   - Perform according to protocol of Mbnl1 CLIP-seq on mouse tissues.
   - Separate equal amount of beads to three clear 1.5 tubes during last washing step.

5. **T1 RNase digestion.**
   - Perform according to protocol of Mbnl1 CLIP-seq on mouse tissues with the modification below.
   - Use three different concentrations of T1 RNase: 0.002, 0.01, 0.5 U/µl.

6. RNA isolation.

Add 1 ml of TRI Reagent (Sigma, cat no. 93289) to each tube. Isolate RNA according to protocol of Mbnl1 CLIP-seq on mouse tissues.

7. Library preparation

Proceed with TruSeq Small RNA Preparation Guide with the modifications below.

- Reduced the volume of PCR to 25 μl.
- Applied 25 cycles.
- From the 6% PAA gel select part from 145 to about 300 bp.

(G) CLIP-seq and RIP-seq data analysis.

FASTX Toolkit (link) was used for quality filtering and the removal of redundant sequences in analyzed datasets. Only one copy of each unique read sequence was kept. Filtered read sequences were aligned to the mouse reference genome (mm10, UCSC) using Bowtie (4), allowing for maximum 5 mappings per read. Overlapping reads were merged into clusters and only clusters consisting of at least 3 reads and spanning at least 30 bp were kept. Clusters were annotated by comparing them to Ensembl 73 annotations (5) using custom Python scripts.

In order to generate Mbnl Interactome Browser (MIB), we pooled our sequencing data with public Mbnl1-CLIP-seq dataset (6-8). From MIB, we excluded C2C12 samples with MBNL1 overexpression. MIB was supplemented in CLIP-seq data for Mbnl2 (8-10), Mbnl3 (8, 11) and other non-Mbnl splicing factors, Celf1 (7) and Nova (12).

References

(1) Ule J, Jensen K, Mele A, Darnell R. CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods. 2005;37:376-386. doi:10.1016/j.ymeth.2005.07.018

(2) Jensen KB, Darnell RB. CLIP: crosslinking and immunoprecipitation of *in vivo* RNA targets of RNA-binding proteins. Methods Mol Biol. 2008;488:85-98. doi: 10.1007/978-1-60327-475-3_6.

(3) Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, et al. iCLIP-transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. J Vis Exp. 2011(50). doi: 10.3791/2638.

(4) Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25. doi: 10.1186/gb-2009-10-3-r25.

(5) Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, et al. Ensembl 2014. Nucleic Acids Res. 2014;42(Database issue):D749-755. doi: 10.1093/nar/gkt1196.

(6) Wang E, Cody N, Jog S, Biancolella M, Wang T, Treacy D, et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell. 2012;150:710-724. doi: 10.1016/j.cell.2012.06.041.

(7) Masuda A, Andersen H, Doktor T, Okamoto T, Ito M, Andresen B, et al. CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. Sci Rep. 2012;2:209. doi: 10.1038/srep00209.

(8) Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. Mol Cell. 2014;56:311-322. doi: 10.1016/j.molcel.2014.08.027.

(9) Charizanis K, Lee K-Y, Batra R, Goodwin M, Zhang C, Yuan Y, et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron. 2012;75:437-450. doi: 10.1016/j.neuron.2012.05.029.

(10) Lee K-Y, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, et al. Compound loss of muscleblind-like function in myotonic dystrophy. EMBO Mol Med. 2013;5:1887-1900. doi: 10.1002/emmm.201303275.

(11) Poulos M, Batra R, Li M, Yuan Y, Zhang C, Darnell R, et al. Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. Hum Mol Genet. 2013;22:3547-3558. doi: 10.1093/hmg/ddt209.

(12) Licatalosi D, Mele A, Fak J, Ule J, Kayikci M, Chi S, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature. 2008;456:464-469. doi: 10.1038/nature07488.

CO-AUTHOR STATEMENT

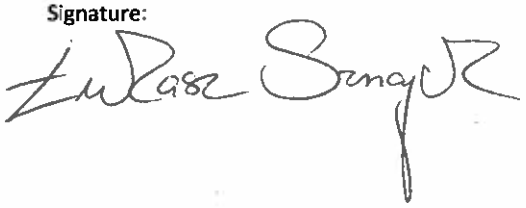
I declare that I am aware that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 27.09.2018, Gainesville, FL

Name: Łukasz J. Sznajder

Signature:



CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 27.09.2018, Poznań

Name: Michał Michalak

Signature: 

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 20.09. 2018, Poznań

Name: Katarzyna Taylor

Signature: Taylor

CO-AUTHOR STATEMENT

I declare that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of my PhD dissertation (PC). I performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. I analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 15.10.2018, Poznań

Name: Piotr Cywoniuk

Signature: Piotr Cywoniuk

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper ,,**Mechanistic determinants of MBNL activity**'' Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 27.09.2018, Poznań

Name: Michał Kabza

Signature: Michał Kabza

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 20.09.2018, Poznań

Name: Agnieszka Wojtkowiak-Szlachcic

Signature: Wojtkowiak-Szlachcic Agnieszka

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper ,,**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 27.09.2018, Paris

Name: Magdalena Matłoka

Signature: Matłoka Magdalene

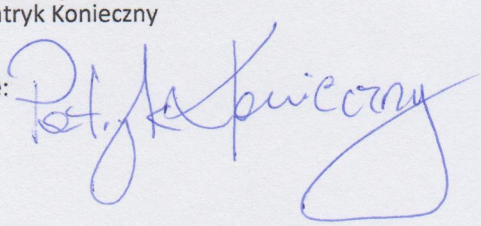
CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 20.09.2018, Poznań

Name: Patryk Konieczny

Signature: 

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**Mechanistic determinants of MBNL activity**" Łukasz J. Sznajder, Michał Michalak, Katarzyna Taylor, Piotr Cywoniuk, Michał Kabza, Agnieszka Wojtkowiak-Szlachcic, Magdalena Matłoka, Patryk Konieczny & Krzysztof Sobczak in ***Nucleic Acids Research*** 2016, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (PC) who performed PCR analysis of alternative splicing of MBNL paralogs in human tissues (Fig. 1B), designed and prepared constructs of Atp2a1, Atp2a1-mutant, Nfix and Ldb3 minigenes, designed AONs against particular transcript regions and performed cellular splicing analysis for binding site verifications via RNA isolation and RT-PCR. PC analyzed obtained data (Fig. 3F, G) and prepared a figure legend of Fig. 3F and 3G and materials and method section "Minigenes" (concerning part), "Cell cultures and transfection" (concerning part) and "RNA isolation and RT-PCR".

ŁJSz generated constructs of MBNL paralogs and their isoforms (without GFP-MB1-40/41/42/43/-N), performed western blot (Fig. 1A, C) and studied expression level of MBNL paralogs in different tissues in silico (Fig. S1). ŁJSz performed analyses of splicing changes of 38 endogenous alternative splicing (AS) events upon MBNL isoform exogenous expression in HeLa cells via RNA isolation and reverse-transcription and PCR (RT-PCR) (Fig. 2A left chart, B, C, Fig. 5A-C, S2, S8D), performed cellular splicing analysis of Atp2a1 and TNNT2 minigenes via RT-PCR (Fig.2C, S4), performed Mbnl1 UV cross-linking and immunoprecipitation as well as RNA immunoprecipitation with high-throughput sequencing (CLIP-seq, RIP-seq, respectively) on mouse tissues (Fig. 3A, B, S3B, C, D, S5), analyzed data from CLIP-seq in C2C12, microarrays and RNA-seq (Fig.2a, S3c, S3E), performed fluorescence microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig.4, S7, S8A-C).
MM (Michał Michalak) with ŁJSz performed confocal microscopy analysis of mCherry/GFP-MBNL cellular localization (Fig. 4, S7, S8A-C), fluorescence resonance energy transfer (FRET), Photoswitching and fluorescence recovery after photobleaching (FRAP), processed and analyzed the data (Fig. 6D-K, S10, S11, Suppl. Videos S1-4).
KT (Katarzyna Taylor) designed, prepared and analyzed all *in vitro* experiments including *in vitro* transcription and radiolabeling of RNA fragments of transcripts and their mutants, filter binding assays (FBA) and all recombinant proteins purification (Fig. 3C-E, 6A, S6).
MK performed CLIP-seq and RIP-seq data analyses in silico and with ŁJSz designed Mbnl Interactome browser (MIB).
AW-Sz with ŁJSz performed fluorescence in situ-hybridization (FISH) assay, analyzed the data (Fig. 6B, C, S9).
MM (Magdalena Matłoka) prepared GST-MBNL2, 3-His constructs and took part in protein purification and FBA under KT supervision.
PK performed CLIP-seq in C2C12 cells (Fig. 5D) and prepared the GFP-MBL-40/41/42/43/-N constructs.
ŁJSz and **KS** designed the study, supervised the experiments, analyzed the data as well as wrote the manuscript. **KT** helped to prepare all parts of the manuscript.

Date: 20.09. 2018, Poznań

Name: Krzysztof Sobczak

Signature:

Nucleic Acids Research, 2018 **1**
doi: 10.1093/nar/gky565

MBNL splicing activity depends on RNA binding site structural context

Katarzyna Taylor[1], Łukasz J. Sznajder[2], Piotr Cywoniuk[1], James D. Thomas[2,3], Maurice S. Swanson[2] and Krzysztof Sobczak[1,*]

[1]Laboratory of Gene Therapy, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland, [2]Center for NeuroGenetics and the Genetics Institute, Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, 2033 Mowry Road, Gainesville, FL 32610, USA and [3]Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

Received February 22, 2018; Revised June 07, 2018; Editorial Decision June 12, 2018; Accepted June 16, 2018

ABSTRACT

Muscleblind-like (MBNL) proteins are conserved RNA-binding factors involved in alternative splicing (AS) regulation during development. While AS is controlled by distribution of MBNL paralogs and isoforms, the affinity of these proteins for specific RNA-binding regions and their location within transcripts, it is currently unclear how RNA structure impacts MBNL-mediated AS regulation. Here, we defined the RNA structural determinants affecting MBNL-dependent AS activity using both cellular and biochemical assays. While enhanced inclusion of MBNL-regulated alternative exons is controlled by the arrangement and number of MBNL binding sites within unstructured RNA, when these sites are embedded in a RNA hairpin MBNL binds preferentially to one side of stem region. Surprisingly, binding of MBNL proteins to RNA targets did not entirely correlate with AS efficiency. Moreover, comparison of MBNL proteins revealed structure-dependent competitive behavior between the paralogs. Our results showed that the structure of targeted RNAs is a prevalent component of the mechanism of alternative splicing regulation by MBNLs.

INTRODUCTION

Alternative splicing (AS) is a complex process in pre-mRNA maturation, enabling the production of multiple protein isoforms distinguished by distinct functions from a single transcript. As much as 95% of pre-mRNAs in mammalian cells undergo AS governed by a unique set of RNA-binding proteins (RBPs) and their corresponding RNA binding sites (1). The pre-mRNA itself has a prominent impact on the activity of RBP on several levels, including RNA sequence specificity, distribution of RBP-binding motifs within target RNAs, the structural organization of RNA motifs and their position with respect to other RBP-binding motifs as well as the sequence composition of the local RBP binding environment (2–4). Along with ubiquitous splicing regulatory proteins, tissue-specific splicing factors with various RNA binding domains and RNA substrate preferences play a crucial role in developmental processes. One of these *trans*-acting factors is the muscleblind-like (MBNL) protein family composed of three paralogs, namely MBNL1, MBNL2 and MBNL3. Despite a primary function of MBNL proteins in the AS of hundreds of pre-mRNAs (5–7), MBNL proteins also regulate subcellular localization (7,8), mRNA stability (7,9,10) and alternative polyadenylation (11).

MBNL paralogs differ considerably in terms of function and expression level during embryonic and postnatal development (12–18). While MBNL1 is the major MBNL family member in adult skeletal muscles and governs myoblast differentiation, MBNL3 is mostly present at early stages of development as well as in the adult placenta, spleen and lung (15,19). MBNL1 is partially compensated by upregulation of MBNL2 but not by MBNL3 (7,20). In contrast, MBNL3 suppresses or delays muscle differentiation when it is constitutively expressed or knocked out in myoblasts, respectively (13,21). MBNL3 also plays a crucial role in a process of muscle regeneration (13,21,22). Mbnl1 knockout mice show retention of fetal splice isoforms in adult organs (23), which recapitulates the phenotype characteristic of myotonic dystrophy type 1 (DM1). DM1 is a multi-systemic disease associated with nuclear sequestration of MBNL proteins in nuclear RNA foci by CUG expansion RNAs transcribed from microsatellite CTG expansions in the 3′ untranslated region (3′UTR) of the *dystrophia myotonica protein kinase* gene (*DMPK*) (24,25). Another form of DM, DM2, is caused by structurally related CCTG expansion within intron 1 of the *cellular nucleic-acid binding protein* (*CNBP*) gene (26), and

*To whom correspondence should be addressed. Tel: +4861 829 5958; Fax: +4861 829 5949; Email: ksobczak@amu.edu.pl

MBNL sequestration by CCUG expansion RNAs also occurs in the DM2 nuclei.

MBNL splicing factor activity generates an AS profile dependent on the pre-mRNA binding site location. For example, alternative cassette exon (CE) skipping is promoted by MBNL binding to the alternative exon or the upstream intron, whereas inclusion is favored by MBNL binding to the CE 3′-end or to the downstream intron (7,16). Current evidence indicates the strength of alternative exon exclusion and inclusion, which defines the expression profile of adult-specific mRNA isoforms during postnatal development (27–30), depends on the binding affinity, activity and relative concentration of MBNL paralogs (31,32). The degree to which the structural organization of MBNL-binding regulatory elements within RNA targets contributes to splicing regulation is currently unclear.

MBNL paralogs contain four zinc fingers (ZF1-4) (of the type $CX_7CX_{4-6}CX_3H$) serving as RNA-binding domains which specifically recognize an RNA motif consisting of a GpC dinucleotide flanked by pyrimidines, YGCY (Y – U or C) (33). Moreover, a GpC flanking adenosine instead of a pyrimidine is an additional motif for MBNL3 binding (A/YGCY/A) (22,30). However, high affinity binding requires at least 9–10 nucleotides outside the YGCY motif with which ZFs also interact (34). The ZFs are arranged as two tandems not functionally equivalent for all splicing events (35,36) and which differ in specificity of YGCY recognition (37). These ZFs are separated by an ∼80 residue linker, which could facilitate binding to separated RNA motifs in *cis* or in *trans* (38,39).

Considering the structural organization of MBNL-binding RNA motifs, MBNL1 does not interact with fully double-stranded RNAs or CUG hairpins stabilized with pseudouridine (40–42). Recent whole transcriptome-based analysis highlighted the necessity of single stranded pyrimidines for MBNL binding, whereas two GpC steps may be either single or double stranded (28). The presence of several pyrimidines over purines in a neighborhood of YGCY motif might be relevant to MBNL binding preference due to their impact on destabilizing RNA structure (33,38). MBNL1 binding to a sequence within the RNA helical region may shift the RNA structure equilibrium towards partial unwinding (34,43). Although an increasing body of evidence indicates an essential role of RNA structure comprising YGCY motifs in MBNL1-binding, the RNA structural determinants have not yet been defined.

We previously showed that the three MBNL paralogs bind to similar RNA motifs with slightly different affinity and control the AS of the same alternative exons with different strength (30). The differences between MBNL paralogs in recognition of the canonical, near-optimal binding motifs and the structural organization of RNA targets may underlie the importance of expressing MBNL paralogs. Here, we examine the RNA structural features preferred by MBNL proteins and their impact on modulation of AS.

Our results demonstrate that MBNL-mediated alternative CE inclusion is modulated by the structural organization of RNA regulatory elements to a greater extent than by the number of MBNL-binding motifs. Strikingly, CE inclusion is promoted by UGCU motifs within unstructured regions or when these motifs are located on one side of an RNA hairpin structure, whereas their distribution on the opposite side of the hairpin weakens splicing activity. Additionally, the distance between these motifs, regardless of the RNA secondary structure, is less significant for MBNL-RNA complex formation compared to splicing regulation. We show that the splicing activity of all MBNL paralogs is fine-tuned by the same RNA structural features. Intriguingly, MBNL3 acts as a competitor of MBNL1 for specific RNA substrates under certain conditions and a similar relationship is observed between different splicing isoforms of MBNL1. Overall, our data suggest that MBNL-dependent splicing changes underlying both physiological and pathological development are fine-tuned by RNA substrate-dependent, tissue-specific and the relative expression levels of MBNL paralogs and their isoforms.

MATERIALS AND METHODS

Preparation of radiolabeled RNA fragments

The templates for transcription reactions for RNA fragments derived from pre-mRNAs (Mbnl1_1, Mbnl1_2, Atp2A1_1, Atp2A1_2, Pphln1_1, Pphln1_2, Calm3, Calm3 sequence mutants, neg_Ctrl, Mbnl2, Mbnl2 sequence mutants) were obtained in two rounds of PCR. First, longer products were amplified from mouse or human genomic DNA using specific _F/_R primer sets at certain T_a. Obtained PCR products constituted a template for the second round of PCR with specific _TF/_TR primer sets having a promoter for polymerase T7 and conducted at T_a = 55°C. The temples for short RNA fragments were either purchased (Future Synthesis) or prepared in a primer extension reaction using two ss DNA sequences (_TF and _TR) complementary to each other at 20 continuous nucleotides at the 3′ or 5′-end, respectively, or through the hybridization of one strand of DNA sequence with a short ss T7 promoter sequence. Prior to the 5 min denaturation step at 90°C, the DNA sequences were gel purified and ethanol precipitated. The templates for longer RNA sequences were prepared by PCR using one strand of DNA sequence and specific primers _TF/TR. The transcription reaction and 5′-end radiolabeling were performed as previously described (44). The sequences of all primers and T_a are listed in Supplementary Table S1.

Biochemical and *in silico* analysis of RNA structure

The analysis was conducted as previously described (45) with concentrations of S1, T1 and T2 endonucleases indicated in Supplementary Figures S1a, S3a, S8a and S9a. The analyzed RNA sequences were also run through the following secondary structure prediction software programs: RNAfold (46), Mfold (47), RNAstructure (48), using biochemically defined constrains or default settings. Each program predicted the same structures as the most energetically favorable.

Hybrid minigene preparation

Hybrid minigenes based on the *Atp2a1* minigene were prepared as previously described (44). The sequences of interest which played a role of MBNL-responsive regulatory

elements were either derived from particular fragments of genes or were artificially designed and amplified in PCR using set of primers (_cF/_cR) or in primer extension reaction using two ss DNA templates (_cF/_cR) complementary to each other at ∼20-nt-long continuous nucleotides at the 3′- or 5′-end, respectively and caring restriction sites for *Not*I and *Sal*I enzymes. The sequences of primers and ss templates are listed in Supplementary Table S1.

Recombinant GST and (His)6-tagged MBNL1, MBNL2 and MBNL3

GST fusions with MBNL1, MBNL2 and MBNL3 proteins encoded by exons 1–4 of the *MBNL1*, *MBNL2* and *MBNL3* genes with a (His)6-tag at the C-terminus were prepared and purified as previously described (30,49,50). The protein concentration was measured using both the Bradford assay and Sypro Ruby staining using 10% SDS-PAGE.

Quantification of rMBNL–RNA interaction

A biochemical assay based on double-membrane filtration analysis was performed as previously described (44). The K_d of the rMBNL–RNA complexes was calculated based on the signal of free RNA and MBNL-RNA complexes in Graph-Pad using *one site specific binding curve* equation, $Y = B_{max}*X/(K_d + X)$.

Cell culture and transfection

The growth conditions of the human HeLa and Human Skeletal Myoblast (HSkM) cell lines were previously described (44), whereas those for Mouse Embryonic Fibroblasts (MEFs) and monkey Cos7 were the same as used for HeLa cells. Prior to transfection, the cells were plated in 12-well plates and transfected at 50–60% confluence with Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocol. Genes were knocked down with siRNAs against MBNL1 (51) or MBNL2 (52) (Future Synthesis and RiboTask™, respectively), or MBNL3 (22) with the ON-TARGET plus SMARTpool (Dharmacon), or siCtrl (44) (Future Synthesis) at 50 nM concentration each. Co-transfection was conducted with 200 ng of minigene and 500 ng of eGFP-MBNL, mCherry or eGFP expressing vectors (or as indicated in the figures). eGFP-MBNL constructs were previously described (30,45). Competition assay was conducted with 200 ng of minigene, 250 ng of MBNL1 (different isoforms) and increasing concentrations of MBNL3 or MBNL1 (different isoforms) as indicated in the figures, and substituted with eGFP. HSkMs or other cells were harvested 24 or 72 h after transfection with siRNA, respectively, or 48 h for the remaining experiments.

Splicing and expression analyses of precursor and mature mRNA

Total RNA isolation, cDNA synthesis and RT-PCR were conducted as previously described (44). The primer sets and T_a are listed in Supplementary Table S2.

eGFP fluorescence detection

HeLa cells were lysed with RIPA buffer [150 mM NaCl, 50 mM Tris–HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5% NP-40, 0.5% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS)] supplemented with SIGMAFAST Protease Inhibitor Cocktail (Sigma). Lysates were incubated on ice for 30 min and vortexed every 10 min prior to centrifugation at 18 000 × *g* at 4°C for 10 min. Samples without heating were separated on 10% SDS polyacrylamide gels, and eGFP fluoresce was quantified on the Amersham Typhoon RGB laser scanner.

Immunoblotting

HeLa cells were lysed with RIPA buffer supplemented with SIGMAFAST Protease Inhibitor Cocktail (Sigma). Lysates were sonicated at 4°C and centrifuged at 18 000 × *g* at 4°C for 10 min. Samples were heated to 95°C for 5 min, separated on 10% SDS polyacrylamide gels and transferred to nitrocellulose (Protran BA 85, Whatman) using a wet transfer apparatus (1 h, 100 V, 4°C). Membranes were blocked for 1 h with 5% skim milk in PBST buffer (phosphate-buffered saline (PBS), 0.1% Tween-20) and incubated with a primary antibody against eGFP (1:1000, Santa Cruz cat. no. sc-8334) or human GAPDH (1:1000, Santa Cruz cat. no. sc-47724). Anti-rabbit (1:20 000, Sigma cat. no. A9169) and anti-mouse (1:2000, Millipore cat. no. 12-349) secondary antibodies were conjugated with horseradish peroxidase and detected using the Pierce ECL Plus Western Blotting Substrate detection kit (Thermo Scientific).

Statistic

Group data obtained in biochemical assays and *in cellula* are expressed as the mean ± standard deviation (SD) or as specified in the figures. Statistical significance of RT-PCR and the double membrane-filtration assay was determined by a two-tailed Student's *t*-test using Microsoft Excel (ns—non-significant; $*P < 0.05$; $**P < 0.01$ and $***P < 0.001$). Statistical analysis was calculated using 2–4 biological replicates (*n*).

RESULTS

Different structures of the same RNA binding site affect MBNL1 binding and splicing activity

The significance of the structural organization of RNA regulatory elements for MBNL1-binding affinity has been noted (53,54), although the specific RNA structural determinants promoting or abrogating MBNL-RNA interactions remain unclear. To gain further insight, we first asked whether alterations of RNA secondary structure encompassing MBNL-binding motifs modify the MBNL1 affinity and its ability to regulate AS. For this purpose, we selected three MBNL-sensitive pre-mRNA regulatory elements from Mbnl1 exon1 (E1) (45), Atp2a1 intron 22 (I22) and Pphln1 intron 6 (I6) (44). Each of these AS substrates contains a few conserved MBNL-recognized 5′-YGCY-3′ motifs (33,55). Next, we transcribed *in vitro* 100 to 190-nt-long fragments of selected transcripts each in two variants

containing the same MBNL-binding motifs but differing in the length of natural 5′ and/or 3′ ends (e.g. Pphln1_v1 and Pphln1_v2 having shorter and longer YGCY-flanking sequences, respectively). Then, for each transcript the RNA secondary structure was determined experimentally with the use of enzymatic probes (RNase T1 and T2 and nuclease S1) recognizing and cleaving single-stranded regions of 5′-end labelled RNAs (56). The experimental data was combined with the results of the RNA structure predictions using software programs (46–48). The secondary structures of RNA variants differed substantially, especially within the regions containing YGCY motifs (Figure 1A and Supplementary Figure S1). Next, we determined the association of recombinant MBNL1 (rMBNL1) to studied RNAs using a biochemical assay based on a double membrane filtration method and observed significant difference in dissociation constant (K_d) values between variants of the same transcript, which suggests the high sensitivity of MBNL1 to RNA secondary structural changes (Figure 1B, C, Supplementary Figure S1 and S2). The lowest K_d value, and the highest affinity, was uniformly characteristic of RNAs containing YGCY motifs within thermodynamically less stable structures.

We next sought to determine whether observed differences in structure-influenced affinities underlie MBNL-driven splicing changes *in cellula*. Thus, we substituted a natural MBNL-sensitive regulatory element within I22 of *Atp2a1* minigene (44) with six of the above-described RNA fragments (hybrid *Atp2a1* minigenes) (Figure 1D). The design of the minigenes preserves experimentally determined RNA secondary structures in cells by a slightly extended, natural and thermodynamically stable region within Atp2a1 I22 (Supplementary Figure S2b). Following co-transfection of HeLa cells with hybrid minigenes and an eGFP-control or MBNL1 expression vector, we measured the splicing activity of the endogenous MBNL pool or overexpressed MBNL1 (MBNL1 OE) as E22 percent spliced-in (PSI). A higher E22 inclusion level was elicited by RNA variants exhibiting a lower K_d and having less stable RNA secondary structures encompassing MBNL-binding motifs (Figure 1C, E and Supplementary Figure S1).

Although the Mbnl1 E1, Pphln1 E7 and Atp2a1 E22 RNAs contain multiple YGCY motifs, from 4 for Pphln1 to 11 for Mbnl1, the number of motifs was not the only determinant of MBNL1 binding affinity or splicing activity (Figure 1C, E, Supplementary Figures S1 and S2). Thus, we next addressed the question of the relationship between YGCY motif number and secondary structure context.

MBNL1 binding and splicing activity is regulated by both the number of UGCU motifs and RNA structural context

We designed several short artificial RNA sequences containing two to four 5′-UGCU-3′ motifs that either folded into semi-stable double-stranded (ds) RNA hairpin structures or remained single-stranded (ss) RNAs (Figure 2A) and subsequently cloned these sequences into the *Atp2a1*-Δ minigene followed by splicing analysis in cells. RT-PCR analysis showed that in the presence of endogenous MBNL proteins, the efficiency of E22 inclusion increased with increasing UGCU number when this motif was in either dsRNA or ssRNA regions (Figure 2B). While 3xUGCU and 4xUGCU, but not 2xYGCY, ssRNA motifs showed significantly higher activity. MBNL1 OE also induced E22 inclusion for all hybrid minigenes showing a saturated effect of E22 for the strongest regulatory elements (e.g. 4ss) (Figure 2C). We confirmed the linkage between observed splicing variations and MBNL1 activity following MBNL1 OE in *Mbnl1*; *Mbnl2* double knockout (DKO) mouse embryonic fibroblasts (MEFs) (Figure 2E). MBNL1 binding affinities were in agreement with the *in cellula* splicing results for endogenous pool of MBNL proteins (Figure 2D).

Our results demonstrated that the number and structural arrangement of RNA motifs recognized by MBNL proteins both impact MBNL-mediated splicing activity. Increasing the number of binding motifs, if analyzed in the same structural context, correlates with stronger affinity and splicing activity of MBNL1. Two UGCU motifs within a pyrimidine-rich ssRNA serve as a functional splicing regulatory element that responds solely to a high MBNL1 concentration.

Distance between UGCU motifs modulates MBNL-mediated splicing regulation but not binding affinity

Next, we investigated the consequences of UGCU motif separation within unstructured RNAs on MBNL1 splicing activity. We designed three sets of RNA regulatory elements containing two, three or four UGCU motifs organized as either overlapping [(–1); …UGC*U*GCU…], sequential [(0); …UGC*UU*GCU…] or separated by five U residues [(5); …UGCU*UUUUU*UGCU…] (Figure 3A). We also experimentally confirmed the single-strand structure of these RNAs in solution (Supplementary Figure S3). Binding assays indicated that rMBNL1 slightly prefers to bind to contiguous UGCU motifs compared to interspaced motifs [e.g. 4ss(0) *versus* 4ss(5)] (Figure 3B). Subsequently, hybrid minigenes were prepared as previously described to investigate the regulatory properties of these RNAs in cells with endogenous MBNL levels. Interestingly, and in contrast to the results of the binding assays, all RNAs with overlapping UGCU motifs had no or reduced splicing activity [e.g. 4ss(–1)] (Figure 3C). On the other hand, RNAs with sequential UGCU motifs showed the highest splicing activity within sets of RNAs containing the same number of UGCU motifs [e.g. 4ss(0)] (Figure 3C). UGCU motif separation [e.g. 4ss2(5)2, 4ss1(5)3, 4ss(5)] always led to reduction of E22 inclusion compared to sequential UGCU motifs (Figure 3C).

The overlapping and sequential arrangements of the UGCU motifs described above correspond to the arrangements of $(CUG)_n$ and $(CCUG)_n$ repeats occurring in DMPK and CNBP transcripts associated with DM1 and DM2 expansion mutations, respectively. Previously, we showed all MBNL paralogs have a high affinity, ∼10 nM, for RNA hairpin structures formed by a non-pathological number of repeats, $(CUG)_n$ and $(CCUG)_n$ (30). The question then arose as to whether the splicing regulatory properties of the repeats differ, as we observed for overlapping and sequential motif arrangements. We tested $(CUG)_{17}$ and $(CCUG)_{14}$, -containing regulatory elements within *Atp2a1* hybrid minigenes in *Mbnl1; Mbnl2* DKO MEFs to di-

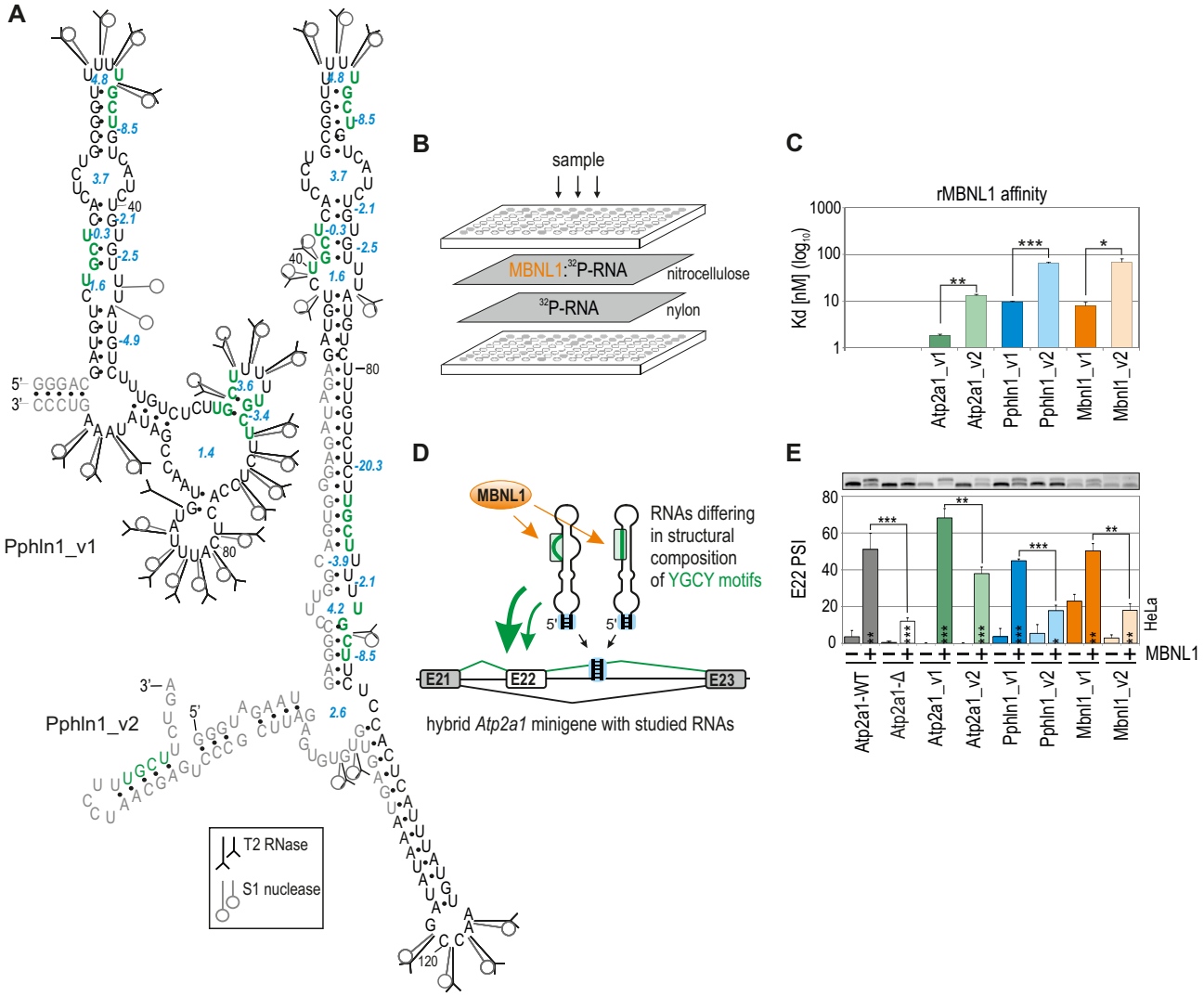


Figure 1. Different structures of the same RNA binding site affect MBNL1 binding and splicing activity. (**A**) Proposed secondary structures of variants of the Pphln1 transcript fragment with previously defined functional MBNL1-binding motifs marked in green and in bold (44). The optimal thermodynamic stability of YGCY-containing RNA regions is expressed in Gibbs free energy (ΔG) in kcal/mol for the reaction at 37°C using RNAfold software (46) and is marked in blue. T2 and S1 nucleases-specific cleavage sites and cleavage intensities are specified for each probe according to symbols shown in the legend. Unique nucleotides, not shared between RNA variants, are marked in grey. See Figure S1 for RNA structures of other transcript variants. (**B**) A simplified scheme of a biochemical assay based on double membrane filtration. The radioactivity on the upper nitrocellulose film (representing ^{32}P-RNA bound to rMBNL1) and that on a lower nylon film (representing free ^{32}P-RNA) was quantified for multiple protein concentrations and MBNL1 affinity (K_d value) for each transcript was calculated (see Figure S2 for more details). (**C**) The quantification of the biochemical assay showing different rMBNL1 binding affinities to two structural variants of three natural transcript fragments. The results are mean K_d values; $n = 2$ for each protein concentration (in the range of 0–250 nM of rMBNL1). (**D**) A model of a hybrid *Atp2a1* minigene with the *Atp2a1* gene fragment between exons 21 and 23 and with a natural MBNL-binding site within intron 22 (I22) in pre-mRNA (Atp2a1-WT) substituted with different YGCY motif-containing RNA fragments. These RNA fragments serve as artificial regulatory elements to test MBNL activity depending on RNA structural organization and YGCY arrangement. A stable helical region formed by Atp2a1 I22, restriction sites used for cloning and a 5-bp-long artificial sequence located at the base of an insert allowing the preservation of its experimentally and *in silico*-defined RNA secondary structure are marked with a blue box (see also Figure S2b). Alternative and constitutive exons are marked with white and grey boxes, respectively; E, exon. (**E**) Different efficiencies of alternative E22 inclusion into the mRNA of hybrid *Atp2a1* minigenes depending on the studied RNA regulatory elements upon eGFP-control or MBNL1 overexpression (OE) in HeLa cells; $n = 3$.

rectly link the splicing response to MBNL1 activity. For MBNL1 OE, we detected higher increase in E22 splicing for Atp2a1-$(CCUG)_{14}$ compared to Atp2a1-$(CUG)_{17}$, which confirmed splicing enhancement for sequential motifs regardless of their number (Figure 3D and Supplementary Figure S4a).

Additionally, the impact of motif arrangement was greater than motif quantity on MBNL-mediated AS [e.g. 3ss(0) versus 4ss(5)] (Figure 3C). MBNL1 OE mostly enhanced the regulatory properties of weaker regulatory elements, indicating the susceptibility of these elements to a higher concentration of MBNL1 (Supplementary Figure S4b and c).

In summary, we have shown that the distance between motifs modulates MBNL1-driven splicing regula-

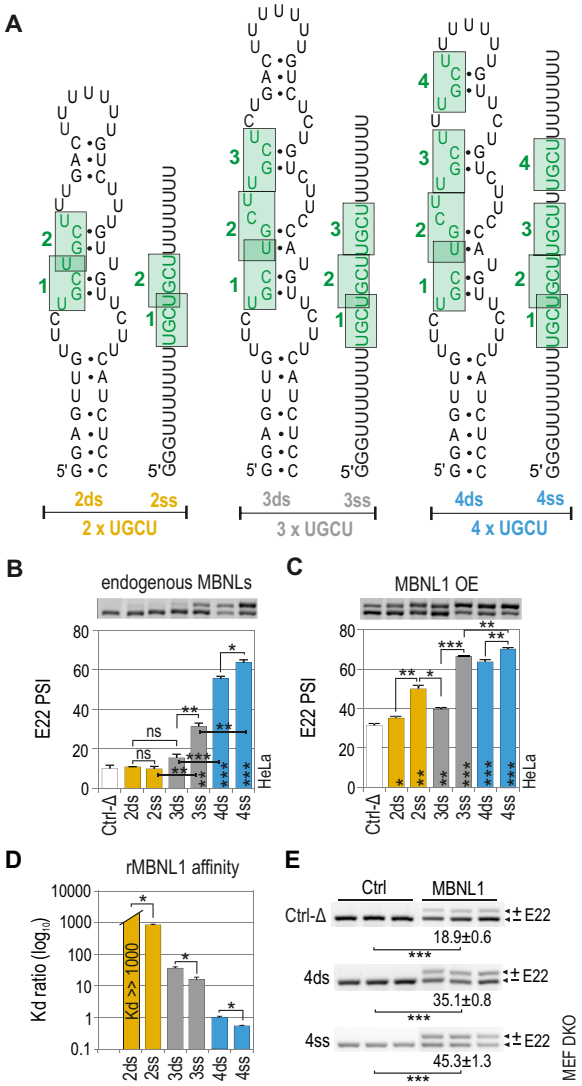


Figure 2. MBNL1 binding and splicing activity is regulated by both the number of UGCU motifs and structural context. (**A**) RNA secondary structures of semi-stable (ds; double-stranded) and unstructured (ss; single-stranded) artificial regulatory elements which are base paired at the 5′ and 3′ ends and internally (only ds RNA) to force the formation of one thermodynamically optimal secondary structure confirmed by us in several RNA structure prediction software programs (46–48). MBNL-binding motifs are numbered from 5′-end and marked in green. The RNA names correspond to the number of UGCU motifs (2, 3 or 4) and the structure type (ds or ss). The splicing response of pre-mRNAs of hybrid minigenes with ds or ss RNAs incorporated in I22 upon (**B**) endogenous level of MBNLs or (**C**) MBNL1 overexpression (OE) in HeLa cells. Vertical asterisks denote the statistical significance of results in comparison to a control experiment (Ctrl-Δ; transfection with *Atp1a1*-Δ minigene). Note a positive effect of the increasing UGCU motif number on alternative splicing - the PSI of E22 for 2ss < 3ss < 4ss; MBNL1 dose-dependency of weaker regulatory elements - the PSI of E22 for 2ds < 2ss > 3ds; and the structural organization of an RNA regulatory element surpassing the number of UGCU motifs - PSI of E22 for 4ds = 3ss; 2ss > 3ds; $n = 2$. (**D**) Quantification of the biochemical assay showing relative rMBNL1 binding affinity for RNA fragments normalized to K_d for the 4ds RNA molecule; $n = 2$ for each protein concentration (in the range of 0–200 nM of rMBNL1). (**E**) RT-PCR showing the splicing response of hybrid *Atp2a1* minigenes representing ss and ds groups of RNAs with or without MBNL1 OE in *Mbnl1*; *Mbnl2* DKO MEFs. The asterisks denote the statistical significance of results compared to cells treated with control eGFP construct (Ctrl). Ctrl-Δ, transfection with *Atp1a1*-Δ minigene; $n = 3$.

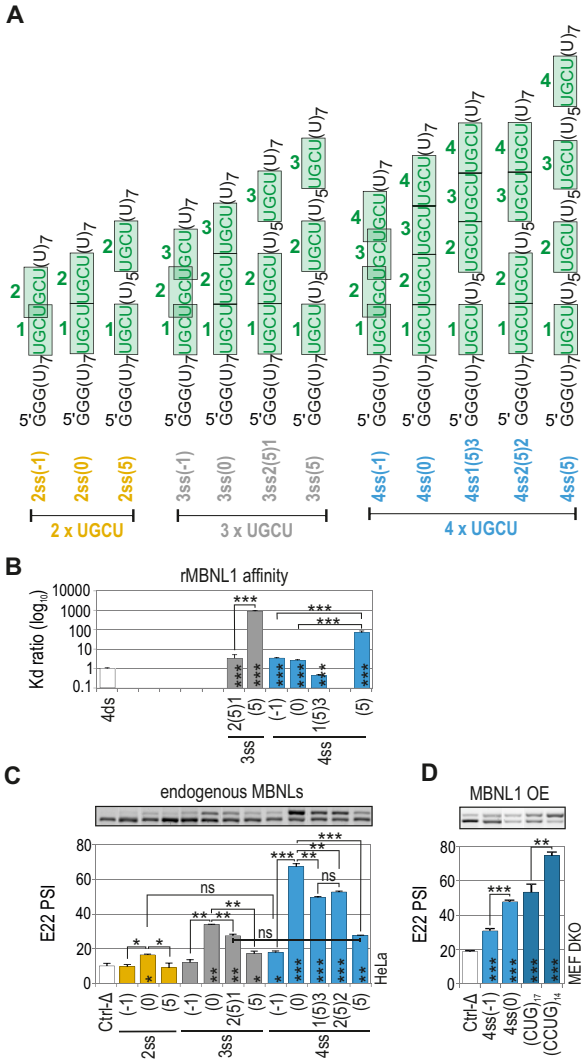


Figure 3. Distance between UGCU motifs modulates MBNL-mediated splicing regulation but not binding affinity. (**A**) The structures of ss RNA regulatory elements with distinctly interspaced UGCU motifs marked in green. The RNA names correspond to the distance between consecutive UGCU motifs, which is marked in brackets surrounded by the number of adjacent motifs. (**B**) As in Figure 2D but for a group of ss RNAs with interspaced UGCU motifs; $n = 2$–4 for each protein concentration (in the range of 0–400 nM of rMBNL1). (**C**) As in Figure 2C but for a group of ss RNAs with interspaced UGCU motifs; $n = 2$. Note no or very weak regulatory properties of RNAs with overlapping motifs - 2ss(–1), 3ss(–1) and 4ss(–1). RNAs with sequential motifs - 4ss(0), 3ss(0) and 2ss(0), exceeded all other tested regulatory elements within the UGCU number sets. UGCU separation had an adverse effect on E22 inclusion - the PSI of E22 for 4ss(0) > 4ss3(5)1 = 4ss2(5)2 > 4ss(5). Note the redundant impact of UGCU number - PSI of E22 for 3ss2(5)1 = 4ss(5). Vertical asterisks denote the statistical significance of results in comparison to a control experiment (Ctrl-Δ; transfection with *Atp1a1*-Δ minigene). (**D**) The splicing response of hybrid *Atp2a1* minigenes with regulatory elements distinguished by either overlapping or sequential arrangement of distinct number of UGCU motifs. The graph shows RT-PCR results for DKO MEFs co-transfected with hybrid minigenes and the MBNL1 expression vector. Vertical asterisks denote the statistical significance of results in comparison to Ctrl-Δ construct; $n = 3$.

tion, which is not a function of differences in affinity for targeted RNAs.

Effect of RNA structure context on MBNL splicing activity

Another important question was the importance of YGCY motif arrangement within the semi-stable RNA structures that predominate at natural MBNL binding sites. We designed several regulatory elements of energetically optimal secondary structures comprising four UGCU motifs that were either: (i) contiguous or interspaced; (ii) located on one side or opposite sides of the RNA helix; (iii) interspaced with, or surrounded by, a stabilizing structure composed of six-base-pairs (Figure 4A).

The rMBNL1 binding affinity for most of RNAs was only slightly reduced in comparison to a reference RNA with contiguous UGCU motifs [*e.g.*, 4ds1(12)2] but strongly decreased for motifs located opposite to each other [e.g. 4ds2(15)2_{OP}] and interspaced with, or surrounded by, six-base-pairs (e.g. 4ds3′S) (Figure 4B, C and Supplementary Figure S5a). Analogous regulatory elements with two or three UGCU motifs showed either no MBNL1–RNA complex formation or depicted the same rMBNL1 preferences to UGCU arrangement as it was shown for 4xUGCU-containing RNAs, respectively (Supplementary Figures S6 and S7).

The analysis of MBNL splicing activity for distinct arrangements of UGCU motifs under cellular conditions gave concordant results only for certain minigenes. MBNL splicing activity strongly decreased for UGCU motifs located on opposite sides of the RNA helix [e.g. 4ds1(29)3_{OP}, 4ds2(15)2_{OP}] (Figure 4D) and when separated or surrounded by stable helical regions [e.g. 4ds2(S)2] (Figure 4E). The same adverse effect was observed for overlapped or interspaced UGCU motifs, although the affinity of MBNL1 for these RNA examples remained high [e.g. 4ds2(6)2] (Figure 4B and D). Intrigued by these results, we next looked at designed RNA examples by correlating the defined values of MBNL1 affinity and the strength of E22 inclusion *in cellula* for the endogenous pool of MBNLs. As expected, the obtained correlation between these two parameters was moderate ($r = –0.56$), but many regulatory elements with low K_d were distinguished by a wide range of E22 inclusion (Figure 4F).

Interestingly, for all RNA regulatory elements, MBNL1 OE proportionally increased E22 inclusion, which confirmed that secondary structural features are sensitive to higher MBNL1 concentrations (Supplementary Figure S5b).

Overall, these results demonstrate that adjacent UGCU motifs located on one side of a hairpin structure are strong splicing regulatory elements while overlapping UGCU motifs impact MBNL splicing, but not binding activity. Furthermore, helix stabilization proximal to UGCU motifs markedly inhibits MBNL–RNA interactions. Since we defined the impact of artificial UGCU motifs within single-stranded and semi-stable RNA structures and the inter-motif distance, reciprocal location and RNA structure stability on MBNL splicing activity, we next examined these RNA features using natural RNA targets of MBNL proteins.

Natural RNA structural determinants that modulate MBNL activity

To study natural RNA structures that influence MBNL activity, we selected two RNA regions: (i) a subregion of the 3′UTR of the Calm3 transcript that undergoes MBNL-dependent selection of alternative polyadenylation sites; (ii) a subregion derived from Mbnl2 E5 that is downregulated by MBNL1 upon binding to intron 4 (Figure 5A, Supplementary Figures S8 and S9) (57). The secondary structure of an ∼100-nt-long sequence of Calm3 RNA was semi-stable ($\Delta G = -12.8$ kcal/mol) with four internal loops and a long terminal loop (Figure 5A and Supplementary Figure S8). By combining a biochemical assay and mutagenesis, we previously defined two major (Cmt1,2) and two subsidiary (Cmt3,4) motifs in Calm3 RNA bound with high affinity by three rMBNL paralogs (30). We confirmed the functionality of these binding sites in AS regulation by incorporating WT and mutant variants of Calm3 in hybrid minigenes and co-expressing them along with MBNL1, MBNL2 and MBNL3. The percent of E22 inclusion decreased significantly upon mutations of the Cmt1 and 2 major motifs in contrast to mutations of the Cmt3 and 4 subsidiary motifs (Figure 5B). We noted certain determinants, including overlapping motifs and motifs located on opposite sides of the RNA helix, which most likely contribute to the reduced responsiveness of this RNA regulatory element. Moreover, it was not responsive to the endogenous MBNL level (Supplementary Figure S8c).

To assess the impact of the structural organization of the Calm3 regulatory element on MBNL splicing activity, we introduced point mutations solely into a hairpin side located opposite to MBNL-binding motifs preserving the primary sequence of the motifs and adjacent nucleotides (Figure 5A). Conceivably, the conversion of a semi-stable hairpin to a more single-stranded structure slightly increased the E22 inclusion rate for all MBNL paralogs (C1,2D; Figure 5B) exhibiting minor sensitivity to MBNL endogenous levels (Supplementary Figure S8c). In contrast, stabilizing the RNA secondary structure of subsidiary motifs (C4S, C3S) slightly reduced E22 inclusion, overcoming the effect of their point mutations (Cmt4, Cmt3), which was consistent with rMBNL affinity (Figure 5B, C, Supplementary Figure S8c and d). These results suggested that MBNL access to the major RNA motifs was constrained by the increased structural stability of flanking sequences. Base-pairing of the major motifs induced the strongest inhibitory effect (C1,2S). We also selected an ∼100-nt-long sequence of Mbnl2 I4 containing four YGCY motifs and confirmed, given certain structural alterations, the negative and positive impact of structural parameters on MBNL1 splicing activity (Supplementary Figure S9).

Thus, the binding and splicing activity of MBNL1, MBNL2 and MBNL3 are strongly modulated by the same RNA structural features in the same sequence contexts. We previously showed that depending on the AS event, the strength of splicing activities of MBNL paralogs and their AS isoforms may differ (30). Hence, we decided to investigate the reciprocal interactions of paralogs in the context of AS regulation and organization of RNA regulatory elements.

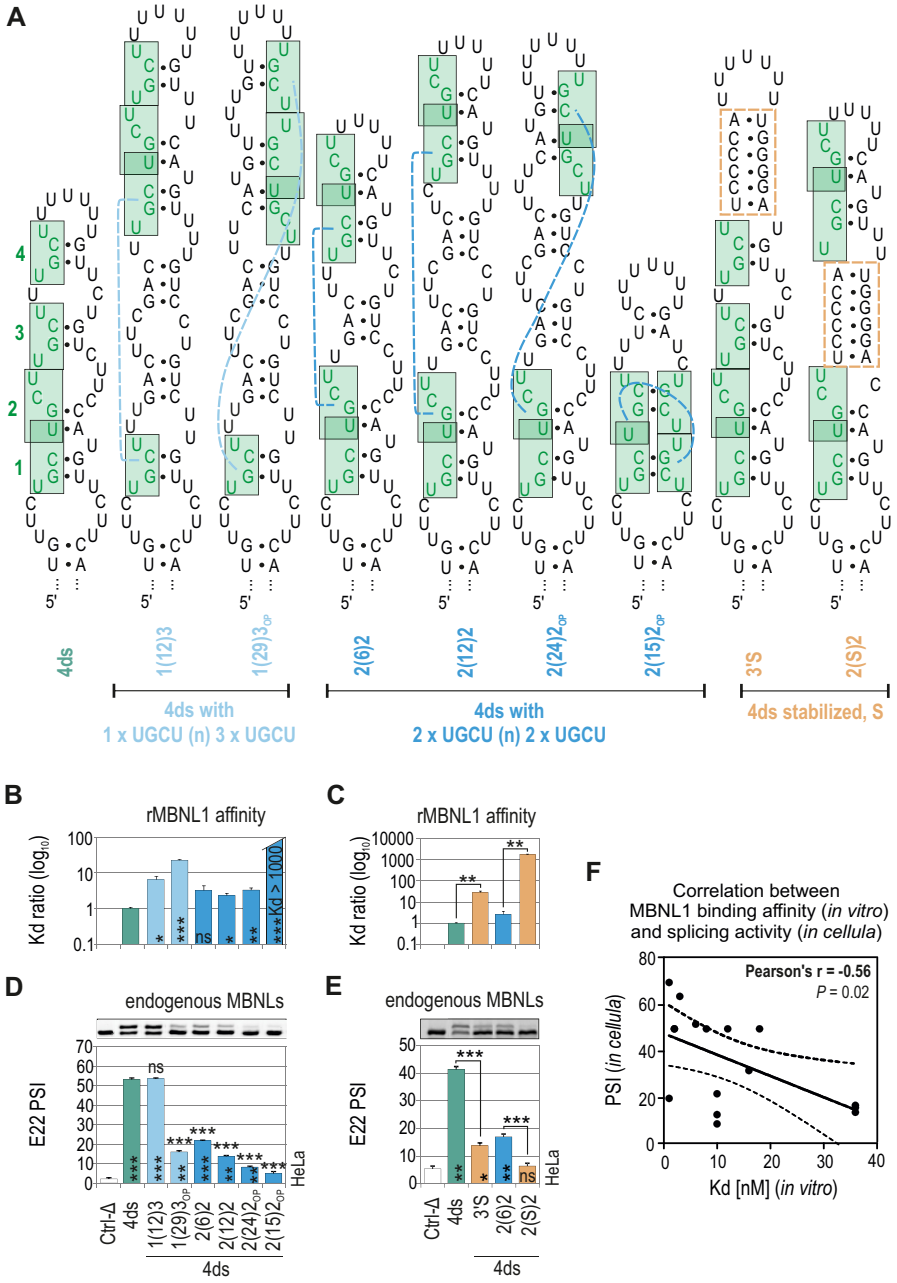


Figure 4. Effect of RNA structural context on MBNL splicing activity. (**A**) The RNA secondary structures of ds RNAs containing four UGCU motifs marked in green. Dashed blue lines depict a reciprocal position of UGCU motifs. Orange rectangles depict ds helical regions stabilizing the RNA secondary structure. The RNA names correspond to the distance between consecutive UGCU motifs which is marked in brackets surrounded by the number of adjacent motifs. (**B** and **C**) As in Figure 2D, but for a group of ds RNAs; $n = 2$–4 for each protein concentration (in the range of 0–200 nM of rMBNL1). (**D** and **E**) As in Figure 2C but for a group of ds RNAs; $n = 2$. Note the decreasing E22 inclusion upon separating tandems of motifs - the PSI of E22 for 4ds2(6)2 > 4ds2(12)2 > 4ds2(24)2. Vertical and horizontal asterisks denote the statistical significance of results in comparison to Ctrl-Δ and 4ds constructs, respectively. (**F**) The Pearson correlation coefficient of rMBNL1 binding affinity (K_d; nM) and splicing activity (PSI values) upon endogenous level of MBNLs for 20 artificial RNA structures analyzed in this study (based on the results shown in Figures 2–4). The RNAs represent two groups of structures (ss or ds) with the same sequence background which showed no effect on E22 inclusion in *Mbnl1*; *Mbnl2* DKO cells (Figure 2E). Dashed lines represent the 95% confidence interval.

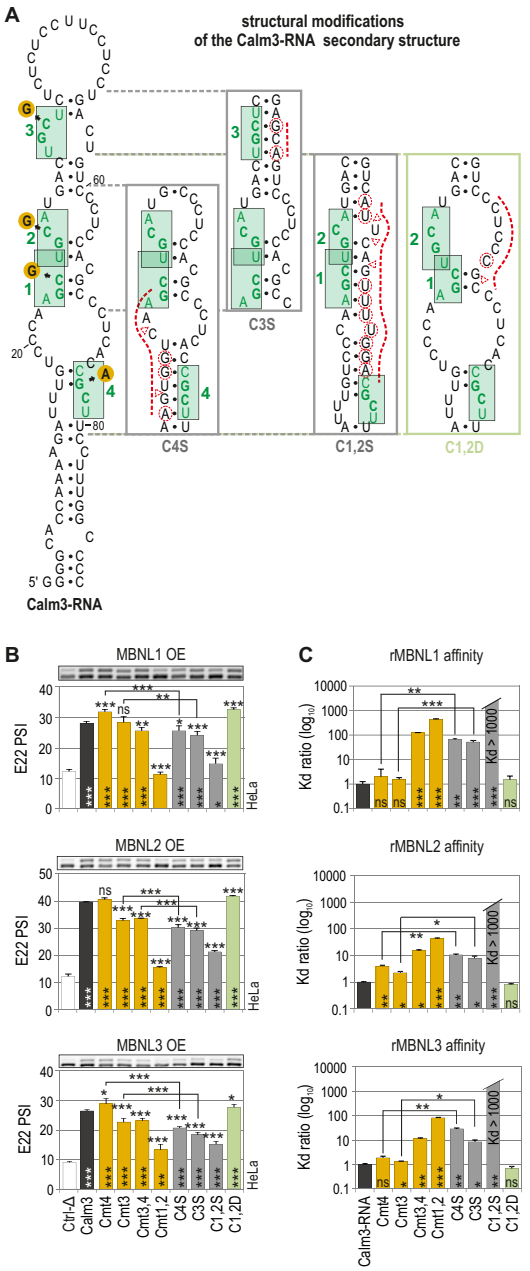


Figure 5. Natural RNA structural determinants that modulate MBNL activity. (**A**) A secondary structure of Calm3-RNA containing MBNL1-binding motifs (HGCH – H stands for U, C or A) marked in green and numbered. Point mutations within the motifs are marked with orange circles. The structural modifications of particular Calm3-RNA regions are presented in boxes. Substitutions and insertions or deletions are marked with red dashed circles or triangles, respectively. (**B**) RT-PCR results showing the splicing response of pre-mRNAs of hybrid minigenes with incorporated different Calm3 regulatory elements carrying HGCH mutations or structural modifications upon MBNL1, MBNL2 and MBNL3 OE in HeLa cells. The MBNL paralogs are identical with respect to splicing isoforms (lacking residues encoded by alternative exons 5 and 7); $n = 2$. Vertical and horizontal asterisks denote the statistical significance of results in comparison to Ctrl-Δ and Calm3 WT constructs, respectively. Ctrl-Δ, *Atp2a1-Δ* minigene. (**C**) Quantification of the biochemical assay showing binding affinity of recombinant rMBNL1, rMBNL2 and rMBNL3 to intact Calm3-RNA, HGCH motif and RNA secondary structure mutants normalized to Calm3-RNA WT; $n = 2$–4 for each protein concentration (in the range of 0–200 nM of rMBNL1). Vertical asterisks denote the statistical significance of results in comparison to intact Calm3-RNA.

Competition between MBNL proteins is driven by RNA regulatory element organization

Due to distinct properties of MBNL1 and MBNL3, including a different developmentally regulated pattern of expression and specific function (22,23), we selected these two paralogs to investigate their reciprocal impact on AS regulation. We previously showed that *Mbnl3* knockout in myoblasts (3KO) results in an increase in adult-like splicing profiles in hundreds of AS events compared to *Mbnl1*; *Mbnl2* DKO muscles, as observed by whole transcriptome analysis (58). In this study, we used RT-PCR to verify several new AS events changed upon MBNL3 loss in 3KO myoblasts in favor of MBNL1 and MBNL2-mediated direction of AS (Figure 6A and Supplementary Figure S10a). To confirm the MBNL3-dependence of these events, we compared the splicing profile of several MBNL-sensitive transcripts in human myoblasts treated with either siRNAs targeting both full-length and short MBNL3 isoforms (3KD cells) or siRNAs against MBNL1 and MBNL2 (DKD cells). The results showed the same direction of splicing changes upon MBNL3 or MBNL1 and MBNL2 silencing, but mostly revealed a larger percentage of adult-like AS isoforms in 3KD cells, in contrast to DKD cells, in which expectedly, the reversion to a developmentally immature splicing profile was observed (Figure 6B and Supplementary Figure S10b). These results were in agreement with observations in 3KO mouse myoblasts (Figure 6A) and were recapitulated in HeLa cells (Figure S10b). In the light of these observations we suppose that in the absence of MBNL3 in cells, the activity of MBNL1 and MBNL2 is enhanced either due to loss of MBNL3 splicing antagonism or to occupation of released binding sites by more active MBNL1 and MBNL2 paralogs.

To test the possibility that MBNL3 antagonizes MBNL1 activity, we analyzed several MBNL-mediated AS events upon MBNL3 and MBNL1 OE in *Mbnl1; Mbnl2* DKO fibroblasts and Cos7 cells. The direction of splicing changes was the same for all analyzed AS events (Figure 6C and Supplementary Figure S10c). These results rule out MBNL3 antagonism but support its competitive behavior in an RNA substrate dependent manner. In consistence with that, nuclear factor IX (NFIX) fetal exon 7 splicing decreased both when MBNL3 was silenced in myoblasts and upon MBNL3 OE in DKO fibroblasts (Figure 6B and C).

We next investigated whether MBNL3 occupancy impedes recruitment of MBNL1 to target RNAs, or *vice versa*, and whether the structural organization of RNA regulatory elements modulates competitive interactions between these paralogs. Because the primary structures of the examined endogenous pre-mRNAs vary in the number of MBNL binding motifs and nucleotide sequence composition, we utilized hybrid minigenes expressing RNAs with structurally distinct regulatory elements. We performed an *in cellula* competition assay, in which in the background of MBNL1 OE promoting E22 inclusion, we titrated different amounts of MBNL3 (Figure 6E). For pre-mRNAs with sequential and interspaced, but not for overlapping UGCU motifs, MBNL1-mediated splicing changes were strongly suppressed in an MBNL3 dose-dependent manner (Figure 6D and Supplementary Figure S11a). A control exper-

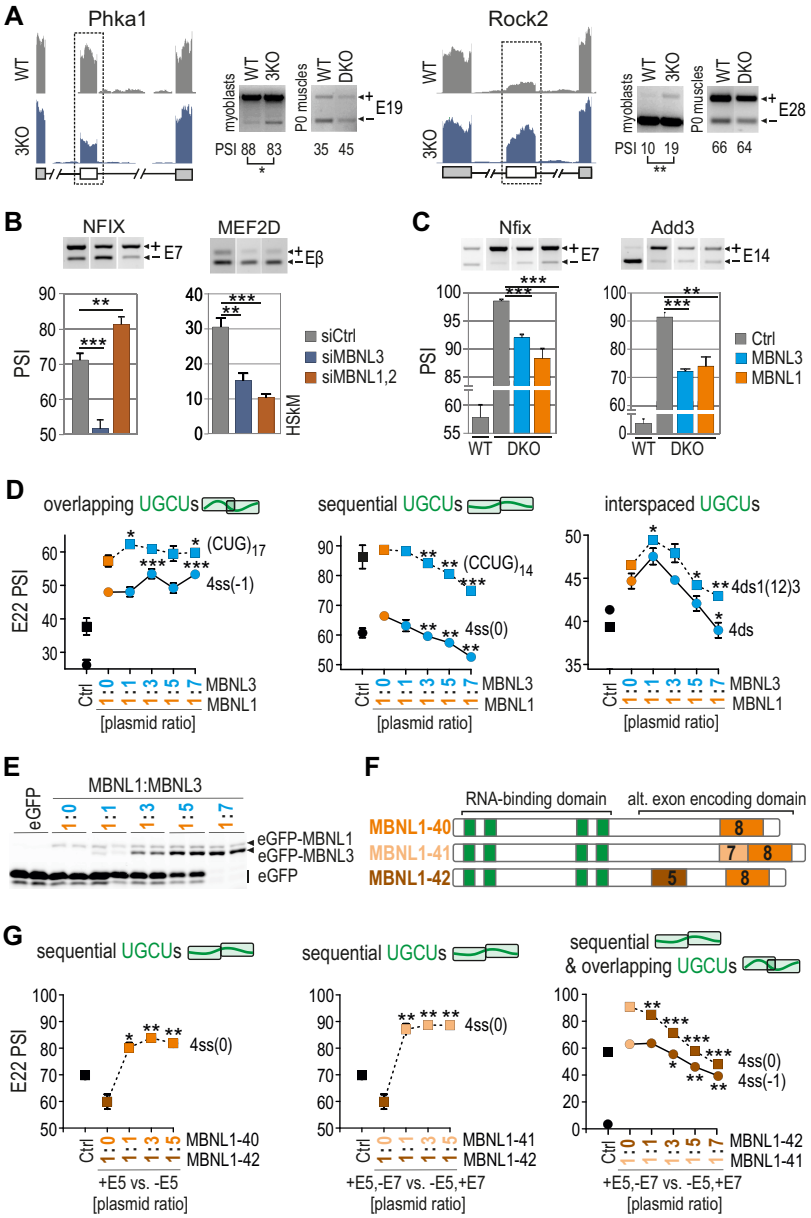


Figure 6. Competition between MBNL proteins is driven by RNA regulatory element organization. (**A**) RNA-seq read coverage across two MBNL-dependent alternative exons indicating their adult-like splicing profiles in *Mbnl3* knockout (3KO) myoblasts compared to wild-type cells and *Mbnl1*; *Mbnl2* DKO quadriceps muscle (postnatal, day 0); alternative and constitutive exons are marked with white or gray boxes, respectively; $n = 2$. Note that the effect of 3KO is opposite to the effect of DKO. (**B**) RT-PCR results of splicing changes in two MBNL1-dependent alternative exons which direction was opposite upon MBNL3 silencing (3KD) compared to MBNL1; MBNL2 double knockdown (DKD) with respect to cells treated with control siRNA in human skeletal myoblasts (HSkM); $n = 3$. (**C**) RT-PCR results of splicing changes in MBNL1-dependent alternative exons which direction was similar upon MBNL3 and MBNL1 OE in *Mbnl1*; *Mbnl2* DKO MEFs; Ctrl, the results from eGFP construct transfected cells (marked in grey); $n = 3$. (**D**) A competition assay between MBNL1 and MBNL3 paralogs. RT-PCR results showing the splicing response of pre-mRNAs with distinctly arranged UGCU motifs in intronic regulatory elements in constant background of MBNL1 OE (marked in orange) and with increasing concentration of MBNL3 (marked in blue) in HeLa cells. MBNL1 and MBNL3 paralogs are identical with respect to splicing isoforms (lacking residues encoded by alternative exons 5 and 7); Ctrl, the results from eGFP construct transfected cells (marked in black). The asterisks denote the statistical significance of results in comparison to cells with MBNL1 OE only (marked in orange); $n = 2$. Each graph represents the results obtained for two constructs belonging to the category of RNA regulatory elements with overlapping, sequential or interspaced UGCU motifs. (**E**) Image of polyacrylamide gels presenting the expression level of eGFP-MBNL1 and eGFP-MBNL3 fusion proteins and eGFP alone. After electrophoresis gels were scanned for eGFP fluorescence (protein samples were not heat-denatured before gel loading). (**F**) A schematic representation of MBNL1 splicing isoforms differing in the presence of residues encoded by alternative exons 5 and 7 (marked in brown and beige, respectively) but sharing the same RNA binding domain composed of four ZFs (marked in green). (**G**) As in Figure 6D but for MBNL1 isoform pairs. An equal amount of applied expression vectors resulted in 2:1 ratio of expressed proteins for MBNL1–42 versus MBNL1–40/41 (30). MBNL1–40, MBNL1–41 and MBNL1–42 are marked in orange, beige and brown, respectively. The asterisks denote the statistical significance of results in comparison to cells with overexpression of only one isoform; $n = 2$.

iment did not show a negative effect of increasing levels of MBNL1 on splicing in of E22 (Supplementary Figure S11b). For some minigenes, MBNL3 suppressed endogenous MBNL-induced inclusion of E22 [e.g. 4ss(0); Figure 6D], which was recapitulated in other cell lines (Supplementary Figure S11c). To further demonstrate the competitive behavior of MBNL proteins, we overexpressed MBNL3 together with hybrid minigenes in *Mbnl1*; *Mbnl2* DKD cells, and noted only the MBNL1-mediated direction of splicing changes (Supplementary Figure S11d).

We next asked whether the RNA structure-dependent competition between paralogs may be a resultant of their different binding affinity. We conducted biochemical studies with truncated (RNA binding domain only) and full-length recombinant MBNL3, although the results showed the reminiscent affinity of both rMBNL3 proteins to selected RNAs (Supplementary Figure S11e) consistently with the results for rMBNL1 (Figure 3).

Intrigued by the competitive behavior observed between MBNL paralogs, we tested if this phenomenon is an attribute of MBNL splicing isoforms as well. Hence, we examined the MBNL1 isoforms differing in the presence of MBNL1 exon 5, which encodes part of a multipartite nuclear localization signal (NLS) (39) and exon 7, encoding a sequence that is responsible for homotypic dimerization (59) (Figure 6F). To monitor differences in the splicing activity of MBNL1 isoforms, we applied a hybrid minigene-based cellular assay with overexpression of pairs of MBNL1 isoforms (Figure 6G and Supplementary Figure S11f). For pre-mRNAs with sequential UGCU motifs, isoform MBNL1–42 (containing exon 5) substantially suppressed E22 inclusion promoted by the endogenous MBNL pool, whereas titrations of MBNL1-40 or MBNL1-41 (which lack exon 5) were equally effective in eliminating this effect, leading to a large increase in E22 inclusion at a relatively low dose (Figure 6G, *left and middle chart*). Additionally, increasing levels of MBNL1–42 considerably mitigated the activity of MBNL1–41 for sequential and overlapping UGCU motifs (Figure 6G, *right chart*). A difference in the splicing activity of MBNL1–40 and MBNL1–41 varying in the presence of exon 7 was significant but exclusive to sequential UGCU motifs (Supplementary Figure S11f).

MBNL paralogs and their isoforms compete for the same RNA binding sites but the resultant splicing profile is a function of distinct paralog and isoform splicing activities and the variability of their C-terminal domains.

DISCUSSION

MBNL activity controlled by the number and distribution of YGCY RNA sequence motifs

A central question in RNA splicing regulation is the relative importance of primary sequence composition and structural organization of RNA regulatory elements within pre-mRNAs on splicing factor affinity and AS regulation. Here, we addressed that question for the MBNL family of developmental splicing regulators using hybrid minigenes with the same pre-mRNA backbone but differing in MBNL-binding UGCU regulatory cassettes to minimize interference by other *cis*- and *trans*-acting factors (44).

The results reported here show a positive correlation between the number of UGCU motifs (2–4) and the binding affinity and splicing activity of MBNL proteins in a sequence/structural context-dependent manner. At least 2xUGCU motifs are required for MBNL binding but the affinity is at least 10-times lower than for 4xUGCU motif-containing RNAs. Expectedly, the splicing of pre-mRNAs expressing 2xUGCU responds solely to high MBNL concentrations, in contrast to 4xUGCU motifs, which are sensitive to a low dose of MBNL. Our findings partially agree with previous results showing the sufficiency of just one such motif but the frequent requirement for multiple motifs (33,38) and the primary sequence composition dependency of MBNL1 binding (33,34). Generally, multiple copies of RNA-binding modules present in RBPs cooperate functionally and structurally (2); hence, the presence of several RNA motifs may allow binding of consecutive ZFs or ZF tandems and enhance interaction affinity as well as the specificity and stability of subsequent RBP-RNA complexes (34,60). However, in whole-transcriptome analyses, many MBNL-binding sites have been identified, but only a small percentage of them are functional and important for RNA processing pathways. For experimentally confirmed natural RNA targets, the number of YGCY motifs is often high and ranges from 9 for INSR I11 (61), 12 for MBNL1 E1 (45) and 15 for LDB3 E10 (44). Our studies show that YGCY arrangement and RNA structure have sometimes a greater impact on AS regulation than the number of YGCY motifs and that these three RNA hallmarks determine the functionality of potential RNA motifs. We observed, that MBNL1 affinity remains within the low-nanomolar K_d range and its splicing activity is high for contiguous 3x or 4xUGCU motifs as well as upon the separation of UGCU tandems within more and less stable structures (Figures 3 and 4) most likely due to a flexible linker which combines ZF tandems and enables their binding to separated RNA motifs (2,38,39). In contrast, overlapping UGCU motifs decrease MBNL1 splicing activity but not binding affinity, whereas separation of individual motifs affects both properties of MBNL1. Hence, we expect that further scattering of individual UGCU motifs will worsen MBNL1 function. ZF tandems of MBNL are not functionally equivalent but regulate splicing in an exclusive or mutually compensating manner (35). Most recent crystallographic studies report that the second ZF in each tandem specifically binds the YGCY motif, whereas the other ZF may interact with RNA less specifically (34). In contrast, another in-depth analysis describes ZF1-2 as a domain specifically recognizing the YGCY motif while ZF3-4 acts as a more general RNA binding domain (37). It could suggest that different YGCY motif arrangement may lead to the formation of MBNL-RNA complexes of distinct splicing capacities.

MBNL activity controlled by RNA secondary structures

Consistently with previous findings (28,34,38), we show that MBNL1 preferentially binds single-stranded RNAs, rather than more stable RNA structures with the same arrangement of YGCY motifs. Base-pairing of YGCY motifs (53) or their structural stabilization *via* substitution of

uridines with pseudouridines (42,54) was shown to impair splicing. Complementing these previous works, we showed that greater stability of RNA structure in regions adjacent to UGCU motifs introduced *via* mutations or occurring through natural interaction of RNA strands, plays a predominant role in modulation of MBNL-mediated AS regulation (Figures 1 and 4). The analyses of RNA structural determinants revealed high preferences of MBNL to contiguous UGCU motifs located on one side of the RNA hairpin. In contrast, neither binding nor splicing activity of MBNL1 was observed when 3x or 4xUGCU motifs are distributed on opposite sides of a stem of the RNA hairpin. Hence, we could expect that the splicing profile of MBNL-regulated alternative exons may result from the composition and arrangement of different local RNA structures as well as binding of other RBPs (44,62–65), the splicing of precursor RNAs and epigenetic parameters (66), which modulate the affinity and activity of MBNL proteins.

MBNL's splicing activity is not solely driven by the binding affinity

Our analysis of UGCU motif distribution in various RNA structural contexts revealed discrepancies between MBNL1 and MBNL3 binding affinity and splicing activity. The most prominent differences were observed for sequential arrangement of 4xUGCU motifs (5′…UGC*UU*GCU…), which in contrast to overlapping arrangement (5′…UGC*U*GCU…), highly promoted MBNL-dependent inclusion of alternative exons in different cellular models, although both served as preferable binding sites in biochemical assays in the present and previous studies (Figures 3–6 and Supplementary Figure S11) (28,30). Similar discrepancies have been described (35), but these reports did not correlate their observations with regulatory element organization and these prior analyses examined six different minigenes potentially introducing interfering variables. Moreover, we noticed that the same localization of distinct RNA regulatory elements in the same pre-mRNA backbone of the minigene induces various MBNL-determined maximal alternative exon inclusion rates, which differ between MBNL paralogs and their splicing isoforms. Our observations lead us to speculate that the process of MBNL binding to different UGCU motif arrangements and RNA structures alters the availability of protein regions previously reported as important in AS regulation. These MBNL regions include a linker sequence between ZF tandems (39,67), proline-rich elements (68), glutamine-rich regions downstream of ZF2 (35), alternative exon encoding sequences (39,67) and ZFs themselves (35) (Figure 7A). However, the observed lack of correlation between RNA binding *in vitro* and MBNL splicing activity *in cellula* could also result from different local RNA structural variations, including the formation of higher order RNA structures due to the impact of various cellular factors or biochemical conditions. The *in vitro* conditions described here are simplified and often limited to the presence of one major RNA structure. Moreover, certain structural/primary sequence arrangements and mutations may alter direct or indirect competition between MBNL1 and other RBPs. For example, polypyrimidine

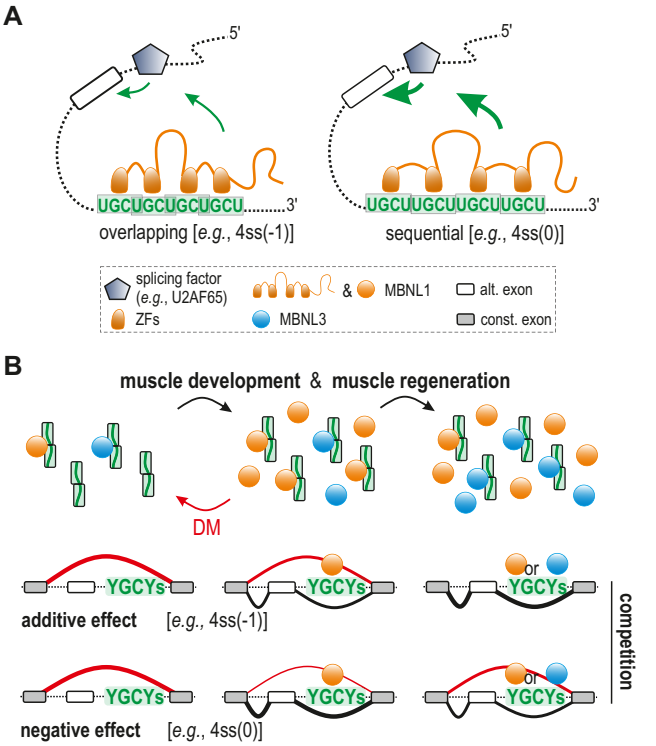


Figure 7. RNA regulatory element organization influences the outcome of MBNL-determined AS profile. (**A**) Upon MBNL binding to distinctly arranged or structured RNA motifs, changes in availability or conformation of splice-activation elements of the protein may occur and lead to alterations in the strength of the MBNL interaction with other splicing factors including U2 auxiliary factor 65 kDa subunit (U2AF65), resulting in weakened or enhanced alternative CE inclusion (depicted by the size of green arrows). (**B**) A model in which MBNL1 and MBNL3 compete at a dose- and RNA substrate-dependent manner during muscle development and regeneration as well as in pathological conditions (DM) modulating the AS profile; more explanation in the text. To simplify, MBNL splicing isoforms and MBNL2 paralog were omitted in the scheme.

tract-binding protein 1 (PTB) and U2AF65 may potentially recognize U-rich regions present in our single-stranded RNA models (62,63). RNA-binding fox-1 homolog (RBFOX) could interact with UGCAUG motif within our Calm3-derived models (69), whereas binding of CUGBP elav-like family member 1 (CELF1) or serine/arginine-rich splicing factor 1 (SRSF1) may modulate MBNL1 binding to RNA (44,64). However, our control experiments using *Mbnl1*; *Mbnl2* DKO cells do not indicate an essential role of other splicing factors in MBNL-mediated splicing (Figure 2 and Supplementary Figure S4).

Competition between MBNLs

Since the presence of three MBNL paralogs was discovered (13,70), much attention has focused on determining the cellular functions and distinguishable features of these paralogs. In contrast to a previously described antagonistic effect of MBNL3 on MBNL1 activity (13,17,71), in the current study, we discovered a competitive interaction between MBNL paralogs and their isoforms that is determined by the organization of RNA regulatory elements and modulates AS profile (Figure 6). Initially, we

observed that the loss of MBNL3 in myoblasts promotes the MBNL-mediated splicing profile (Figure 6). This splicing phenomenon was not evident either in adult heart or muscles of *Mbnl3* KO mice, probably due to the relatively low contribution of MBNL3 to the total pool of MBNL in adult tissues (17). However, while MBNL3 predominates in cells lacking MBNL1 and MBNL2, the paralog shifts the AS profile towards MBNL-induced (Figure 6 and Supplementary Figure S11) what confirms a convergent splicing activity of MBNL paralogs and suggests their competitive interaction while present concomitantly in cells. The competition seems to be RNA structure-influenced, since MBNL paralogs and isoforms competed with endogenous and overexpressed pool of MBNL for binding to both sequential and overlapping UGCU motifs within regulatory elements triggering a negative effect on splicing for mostly sequential UGCU arrangement (Figure 6d). One possible explanation of this phenomenon is that the splice-activation domains of MBNL paralogs vary, including a linker between ZF tandems that is merely 50% identical between the paralogs (67) and the distinct arrangement of proline-rich regions considered as a unique platform for association with other splicing factors (71). Moreover, MBNL3 lacks an alanine-rich region potentially important for RNA binding (72). MBNL isoforms share the same RNA-binding domain but have distinct C-terminal domains encoded by alternative exons which contribute to the protein activity either directly or by modulating protein regions involved in AS regulation (39,67,68,72).

The RNA organization-based competition of RBPs has not been described yet. However, a reminiscent observation has been reported for proteins belonging to the nuclear factor I (NFI) protein family, which binds to a specific sequence of DNA. NFI comprises four variants regulating gene expression in an antagonistic and synergistic way depending on their activities and distribution in cells (73). Future studies focused on MBNL1- and MBNL3-associated splicing factors should shed more light on the competitive mode of MBNL function.

In the light of the collected evidence, we propose a model for the competitive mode of action of MBNL paralogs dependent on RNA organization (Figure 7b). Due to the low active pool of MBNL paralogs occurring in initial stages of development (23) or under pathological conditions (25), MBNL-sensitive splicing isoforms are not promoted. After MBNL1 and MBNL3 levels increase during muscle development, MBNL-dependent splicing favors transcripts distinguished by a sequential-like YGCY arrangement. Under increased MBNL3 production during muscle regeneration and the saturation of many binding regions, MBNL3 displaces other MBNLs in protein-RNA complexes and causes either an additive effect on splicing (*e.g.*, for RNAs with the YGCY-overlapping arrangement) or a reverse direction of splicing changes (*e.g.*, for RNAs with the YGCY-sequential arrangement).

The discovery of complementary MBNL paralogs that become competitors for certain types of RNAs and in particular environments provides deeper insight into the molecular changes underlain by the sequestration and functional insufficiency of MBNLs in DM. Revealing the RNA structural features enabling MBNLs to efficiently bind the substrates with a low impact on MBNL splicing activity have potential therapeutic implications. It may lead to the development of an MBNL-derived molecule with high affinity to toxic CUG and CCUG repeats in DM, but negligible cellular activity due to limitations of interactions with other *trans*-acting factors.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Authors Contributions: Conception and design of experiments: K.S., K.T.. Acquisition, analysis and interpretation of data: P.C. (western blots), Ł.J.S. (RT-PCR verification of splicing profile in WT and 3KO myoblasts), J.D.T. (RNA-seq data processing), K.T. (the rest of experiments). Drafting or revising the manuscript: K.S., K.T., M.S.S., Ł.J.S., J.D.T.

FUNDING

Foundation for Polish Science TEAM program co-financed by the European Union within the European Regional Development Fund [TEAM/2011-7/10 to K.S.]; Polish National Science Centre [2011/01/B/NZ1/01603, 2014/15/B/NZ2/02453 to K.S., 2017/24/C/NZ1/00112 to K.T.]; post-doctoral fellowship award from the Myotonic Dystrophy and Wyck Foundations (to Ł.J.S.). 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) program for the years 2014–2019.
Conflict of interest statement. M.S.S. is a member of the scientific advisory board of Locana Bio. The other authors declare no competing interests.

REFERENCES

1. Kornblihtt,A.R., Schor,I.E., Allo,M., Dujardin,G., Petrillo,E. and Munoz,M.J. (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.*, **14**, 153–165.
2. Lunde,B.M., Moore,C. and Varani,G. (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.*, **8**, 479–490.
3. Wang,X.L., Vukovic,L., Koh,H.R., Schulten,K. and Myong,S. (2015) Dynamic profiling of double-stranded RNA binding proteins. *Nucleic Acids Res.*, **43**, 7566–7576.
4. Lewis,C.J.T., Pan,T. and Kalsotra,A. (2017) RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat. Rev. Mol. Cell Biol.*, **18**, 202–210.
5. 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.
6. Fleming,V.A., Geng,C., Ladd,A.N. and Lou,H. (2012) Alternative splicing of the neurofibromatosis type 1 pre-mRNA is regulated by the muscleblind-like proteins and the CUG-BP and ELAV-like factors. *BMC Mol. Biol.*, **13**, 35.
7. Wang,E.T., Cody,N.A.L., Jog,S., Biancolella,M., Wang,T.T., Treacy,D.J., Luo,S.J., Schroth,G.P., Housman,D.E., Reddy,S. *et al.* (2012) Transcriptome-wide regulation of Pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell*, **150**, 710–724.
8. Adereth,Y., Dammai,V., Kose,N., Li,R.Z. and Hsu,T. (2005) RNA-dependent integrin alpha(3) protein localization regulated by the Muscleblind-like protein MLP1. *Nat. Cell Biol.*, **7**, 1240–1247.

9. Du,H.Q., Cline,M.S., Osborne,R.J., Tuttle,D.L., Clark,T.A., Donohue,J.P., Hall,M.P., Shiue,L., Swanson,M.S., Thornton,C.A. *et al.* (2010) Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.*, **17**, 187–193.
10. Masuda,A., Andersen,H.S., Doktor,T.K., Okamoto,T., Ito,M., Andresen,B.S. and Ohno,K. (2012) CUGBP1 and MBNL1 preferentially bind to 3 ' UTRs and facilitate mRNA decay. *Scientific Rep.*, **2**, 1–10.
11. Batra,R., Manchanda,M. and Swanson,M.S. (2015) Global insights into alternative polyadenylation regulation. *RNA Biol.*, **12**, 597–602.
12. 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.
13. Squillace,R.M., Chenault,D.M. and Wang,E.H. (2002) Inhibition of muscle differentiation by the novel muscleblind-related protein CHCR. *Dev. Biol.*, **250**, 218–230.
14. 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.
15. Holt,I., Jacquemin,V., Fardaei,M., Sewry,C.A., Butler-Browne,G.S., Furling,D., Brook,J.D. and Morris,G.E. (2009) Muscleblind-Like proteins similarities and differences in normal and myotonic dystrophy muscle. *Am. J. Pathol.*, **174**, 216–227.
16. Charizanis,K., Lee,K.Y., Batra,R., Goodwin,M., Zhang,C.L., Yuan,Y., Shiue,L., Cline,M., Scotti,M.M., Xia,G.B. *et al.* (2012) Muscleblind-like 2-Mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron*, **75**, 437–450.
17. Choi,J., Dixon,D.M., Dansithong,W., Abdallah,W.F., Roos,K.P., Jordan,M.C., Trac,B., Lee,H.S., Comai,L. and Reddy,S. (2016) Muscleblind-like 3 deficit results in a spectrum of age-associated pathologies observed in myotonic dystrophy. *Scientific Rep.*, **6**, 1–10.
18. Zhang,B.W., Cai,H.F., Wei,X.F., Sun,J.J., Lan,X.Y., Lei,C.Z., Lin,F.P., Qi,X.L., Plath,M. and Chen,H. (2016) miR-30-5p regulates muscle differentiation and alternative splicing of Muscle-Related genes by targeting MBNL. *Int. J. Mol. Sci.*, **17**, 1–16.
19. Lee,K.S., Squillace,R.M. and Wang,E.H. (2007) Expression pattern of muscleblind-like proteins differs in differentiating myoblasts. *Biochem. Biophys. Res. Commun.*, **361**, 151–155.
20. Lee,K.Y., Li,M.Y., Manchanda,M., Batra,R., Charizanis,K., Mohan,A., Warren,S.A., Chamberlain,C.M., Finn,D., Hong,H. *et al.* (2013) Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol. Med.*, **5**, 1887–1900.
21. Lee,K.S., Smith,K., Amieux,P.S. and Wang,E.H. (2008) MBNL3/CHCR prevents myogenic differentiation by inhibiting MyoD-dependent gene transcription. *Differentiation*, **76**, 299–309.
22. Poulos,M.G., Batra,R., Li,M.Y., Yuan,Y., Zhang,C.L., Darnell,R.B. and Swanson,M.S. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. *Hum. Mol. Genet.*, **22**, 3547–3558.
23. Han,H., Irimia,M., Ross,P.J., Sung,H.K., Alipanahi,B., David,L., Golipour,A., Gabut,M., Michael,I.P., Nachman,E.N. *et al.* (2013) MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature*, **498**, 241–248.
24. 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.
25. 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.
26. Ranum,L.P.W., Rasmussen,P.F., Benzow,K.A., Koob,M.D. and Day,J.W. (1998) Genetic mapping of a second myotonic dystrophy locus. *Nat. Genet.*, **19**, 196–198.
27. Paul,S., Dansithong,W., Jog,S.P., Holt,I., Mittal,S., Brook,J.D., Morris,G.E., Comai,L. and Reddy,S. (2011) Expanded CUG repeats dysregulate RNA splicing by altering the stoichiometry of the muscleblind 1 complex. *J. Biol. Chem.*, **286**, 38427–38438.
28. Lambert,N., Robertson,A., Jangi,M., McGeary,S., Sharp,P.A. and Burge,C.B. (2014) RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell*, **54**, 887–900.
29. Wagner,S.D., Struck,A.J., Gupta,R., Farnsworth,D.R., Mahady,A.E., Eichinger,K., Thornton,C.A., Wang,E.T. and Berglund,J.A. (2016) Dose-Dependent regulation of alternative splicing by MBNL proteins reveals biomarkers for myotonic dystrophy. *PLoS Genet.*, **12**, 24.
30. Sznajder,L.J., Michalak,M., Taylor,K., Cywoniuk,P., Kabza,M., Wojtkowiak-Szlachcic,A., Matloka,M., Konieczny,P. and Sobczak,K. (2016) Mechanistic determinants of MBNL activity. *Nucleic Acids Res.*, **44**, 10326–10342.
31. Kalsotra,A., Xiao,X.S., Ward,A.J., Castle,J.C., Johnson,J.M., Burge,C.B. and Cooper,T.A. (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. Solana,J., Irimia,M., Ayoub,S., Orejuela,M.R., Zywitza,V., Jens,M., Tapial,J., Ray,D., Morris,Q.D., Hughes,T.R. *et al.* (2016) Conserved functional antagonism of CELF and MBNL proteins controls stem cell-specific alternative splicing in planarians. *Elife*, **5**, 29.
33. Goers,E.S., Purcell,J., Voelker,R.B., Gates,D.P. and Berglund,J.A. (2010) MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.*, **38**, 2467–2484.
34. Park,S., Phukan,P.D., Zeeb,M., Martinez-Yamout,M.A., Dyson,H.J. and Wright,P.E. (2017) Structural basis for interaction of the tandem zinc finger domains of human muscleblind with cognate RNA from human cardiac troponin T. *Biochemistry*, **56**, 4154–4168.
35. Purcell,J., Oddo,J.C., Wang,E.T. and Berglund,J.A. (2012) Combinatorial mutagenesis of MBNL1 zinc fingers elucidates distinct classes of regulatory events. *Mol. Cell. Biol.*, **32**, 4155–4167.
36. He,F., Dang,W., Abe,C., Tsuda,K., Inoue,M., Watanabe,S., Kobayashi,N., Kigawa,T., Matsuda,T., Yabuki,T. *et al.* (2009) Solution structure of the RNA binding domain in the human muscleblind-like protein 2. *Protein Sci.*, **18**, 80–91.
37. Hale,M., Richardson,J., Day,R., McConnell,O., Arboleda,J., Wang,E. and Berglund,J. (2018) An engineered RNA binding protein with improved splicing regulation. *Nucleic Acids Res.*, **46**, 3152–3168.
38. Cass,D., Hotchko,R., Barber,P., Jones,K., Gates,D.P. and Berglund,J.A. (2011) The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements. *BMC Mol. Biol.*, **12**, 1–7.
39. Tran,H., Gourrier,N., Lemercier-Neuillet,C., Dhaenens,C.M., Vautrin,A., Fernandez-Gomez,F.J., 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**, 12.
40. 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.
41. Goers,E.S., Voelker,R.B., Gates,D.P. and Berglund,J.A. (2008) RNA binding specificity of Drosophila muscleblind. *Biochemistry*, **47**, 7284–7294.
42. deLorimier,E., Coonrod,L.A., Copperman,J., Taber,A., Reister,E.E., Sharma,K., Todd,P.K., Guenza,M.G. and Berglund,J.A. (2014) Modifications to toxic CUG RNAs induce structural stability, rescue mis-splicing in a myotonic dystrophy cell model and reduce toxicity in a myotonic dystrophy zebrafish model. *Nucleic Acids Res.*, **42**, 12768–12778.
43. Fu,Y., Ramisetty,S.R., Hussain,N. and Baranger,A.M. (2012) MBNL1-RNA Recognition: Contributions of MBNL1 sequence and RNA conformation. *Chembiochem*, **13**, 112–119.
44. Cywoniuk,P., Taylor,K., Sznajder,L.J. and Sobczak,K. (2017) Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions. *Scientific Rep.*, **7**, 14.
45. Konieczny,P., Stepniak-Konieczna,E., Taylor,K., Sznajder,L.J. and Sobczak,K. (2017) Autoregulation of MBNL1 function by exon 1 exclusion from MBNL1 transcript. *Nucleic Acids Res.*, **45**, 1760–1775.
46. Gruber,A.R., Lorenz,R., Bernhart,S.H., Neuboock,R. and Hofacker,I.L. (2008) The vienna RNA websuite. *Nucleic Acids Res.*, **36**, 70–74.
47. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
48. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 9.

49. Chen,C.Z., Sobczak,K., Hoskins,J., Southall,N., Marugan,J.J., Zheng,W., Thornton,C.A. and Austin,C.P. (2012) Two high-throughput screening assays for aberrant RNA-protein interactions in myotonic dystrophy type 1. *Analyt. Bioanalyt. Chem.*, **402**, 1889–1898.
50. Wojciechowska,M., Taylor,K., Sobczak,K., Napierala,M. and Krzyzosiak,W.J. (2014) Small molecule kinase inhibitors alleviate different molecular features of myotonic dystrophy type 1. *RNA Biol.*, **11**, 742–754.
51. Childs-Disney,J.L., Stepniak-Konieczna,E., Tran,T., Yildirim,I., Park,H., Chen,C.Z., Hoskins,J., Southall,N., Marugan,J.J., Patnaik,S. *et al.* (2013) Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules. *Nat. Commun.*, **4**, 11.
52. 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.
53. Taliaferro,J.M., Lambert,N.J., Sudmant,P.H., Dominguez,D., Merkin,J.J., Alexis,M.S., Bazile,C.A. and Burge,C.B. (2016) RNA Sequence context effects measured in vitro predict in vivo protein binding and regulation. *Mol. Cell*, **64**, 294–306.
54. deLorimier,E., Hinman,M.N., Copperman,J., Datta,K., Guenza,M. and Berglund,J.A. (2017) Pseudouridine modification inhibits Muscleblind-like 1 (MBNL1) Binding to CCUG repeats and minimally structured RNA through reduced RNA flexibility. *J. Biol. Chem.*, **292**, 4350–4357.
55. 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.
56. Sobczak,K., Michlewski,G., de Mezer,M., Krol,J. and Krzyzosiak,W.J. (2010) Trinucleotide repeat system for sequence specificity analysis of RNA structure probing reagents. *Analyt. Biochem.*, **402**, 40–46.
57. Batra,R., Charizanis,K., Manchanda,M., Mohan,A., Li,M.Y., Finn,D.J., Goodwin,M., Zhang,C.L., Sobczak,K., Thornton,C.A. *et al.* (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-Mediated disease. *Mol. Cell*, **56**, 311–322.
58. Thomas,J.D., Sznajder,L.J., Bardhi,O., Aslam,F.N., Anastasiadis,Z.P., Scotti,M.M., Nishino,I., Nakamori,M., Wang,E.T. and Swanson,M.S. (2017) Disrupted prenatal RNA processing and myogenesis in congenital myotonic dystrophy. *Genes Dev.*, **31**, 1122–1133.
59. 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.
60. Edge,C., Gooding,C. and Smith,C.W.J. (2013) Dissecting domains necessary for activation and repression of splicing by muscleblind-like protein 1. *BMC Mol. Biol.*, **14**, 16.
61. Sen,S., Talukdar,I., Liu,Y., Tam,J., Reddy,S. and Webster,N.J.G. (2010) Muscleblind-like 1 (Mbnl1) Promotes insulin receptor exon 11 inclusion via binding to a downstream evolutionarily conserved intronic enhancer. *J. Biol. Chem.*, **285**, 25426–25437.
62. Warf,M.B., Diegel,J.V., von Hippel,P.H. and Berglund,J.A. (2009) The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9203–9208.
63. Echeverria,G.V. and Cooper,T.A. (2014) Muscleblind-like 1 activates insulin receptor exon 11 inclusion by enhancing U2AF65 binding and splicing of the upstream intron. *Nucleic Acids Res.*, **42**, 1893–1903.
64. Wang,E.T., Ward,A.J., Cherone,J.M., Giudice,J., Wang,T.T., Treacy,D.J., Lambert,N.J., Freese,P., Saxena,T., Cooper,T.A. *et al.* (2015) Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.*, **25**, 858–871.
65. Sellier,C., Cerro-Herreros,E., Blatter,M., Freyermuth,F., Gaucherot,A., Ruffenach,F., Sarkar,P., Puymirat,J., Bjarne,U., Day,J.W. *et al.* (2018) RbFOX1/MBNL1 competition for CCUG RNA repeats binding contributes to myotonic dystrophy type1/type 2 differences. *Nat. Commun.*, **9**, 1–16.
66. Luco,R.F., Allo,M., Schor,I.E., Kornblihtt,A.R. and Misteli,T. (2011) Epigenetics in Alternative Pre-mRNA Splicing. *Cell*, **144**, 16–26.
67. Grammatikakis,I., Goo,Y.H., Echeverria,G.V. and Cooper,T.A. (2011) Identification of MBNL1 and MBNL3 domains required for splicing activation and repression. *Nucleic Acids Res.*, **39**, 2769–2780.
68. Botta,A., Malena,A., Tibaldi,E., Rocchi,L., Loro,E., Pena,E., Cenci,L., Ambrosi,E., Bellocchi,M.C., Pagano,M.A. *et al.* (2013) MBNL1(42) and MBNL1(43) gene isoforms, overexpressed in DM1-patient muscle, encode for nuclear proteins interacting with Src family kinases. *Cell Death Dis.*, **4**, 12.
69. Conboy,J.G. (2017) Developmental regulation of RNA processing by Rbfox proteins. *Wiley Interdiscip. Rev.-RNA*, **8**, 17.
70. Fardaei,M., Larkin,K., Brook,J.D. and Hamshere,M.G. (2001) In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res.*, **29**, 2766–2771.
71. Lee,K.S., Cao,Y., Witwicka,H.E., Tom,S., Tapscott,S.J. and Wang,E.H. (2010) RNA-binding protein muscleblind-like 3 (MBNL3) disrupts myocyte enhancer factor 2 (Mef2) beta-exon splicing. *J. Biol. Chem.*, **285**, 33779–33787.
72. Pascual,M., Vicente,M., Monferrer,L. and Artero,R. (2006) The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation*, **74**, 65–80.
73. Perez-Casellas,L.A., Wang,X.Y., Howard,K.D., Rehage,M.W., Strong,D.D. and Linkhart,T.A. (2009) Nuclear Factor I transcription factors regulate IGF binding protein 5 gene transcription in human osteoblasts. *Biochim. Biophys. Acta-Gene Regul. Mech.*, **1789**, 78–87.

MBNL splicing activity depends on RNA binding site structural context

Katarzyna Taylor[1], Łukasz J. Sznajder[2], Piotr Cywoniuk[1], James D. Thomas[2,3], Maurice S. Swanson[2] and Krzysztof Sobczak[1,*]

[1] Laboratory of Gene Therapy, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland.

[2] Center for NeuroGenetics and the Genetics Institute, Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, 2033 Mowry Road, Gainesville, Florida 32610, USA.

[3] Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA.

* To whom correspondence should be addressed. Tel.: +4861 829 5958; Fax: +4861 829 5949; Email: ksobczak@amu.edu.pl.

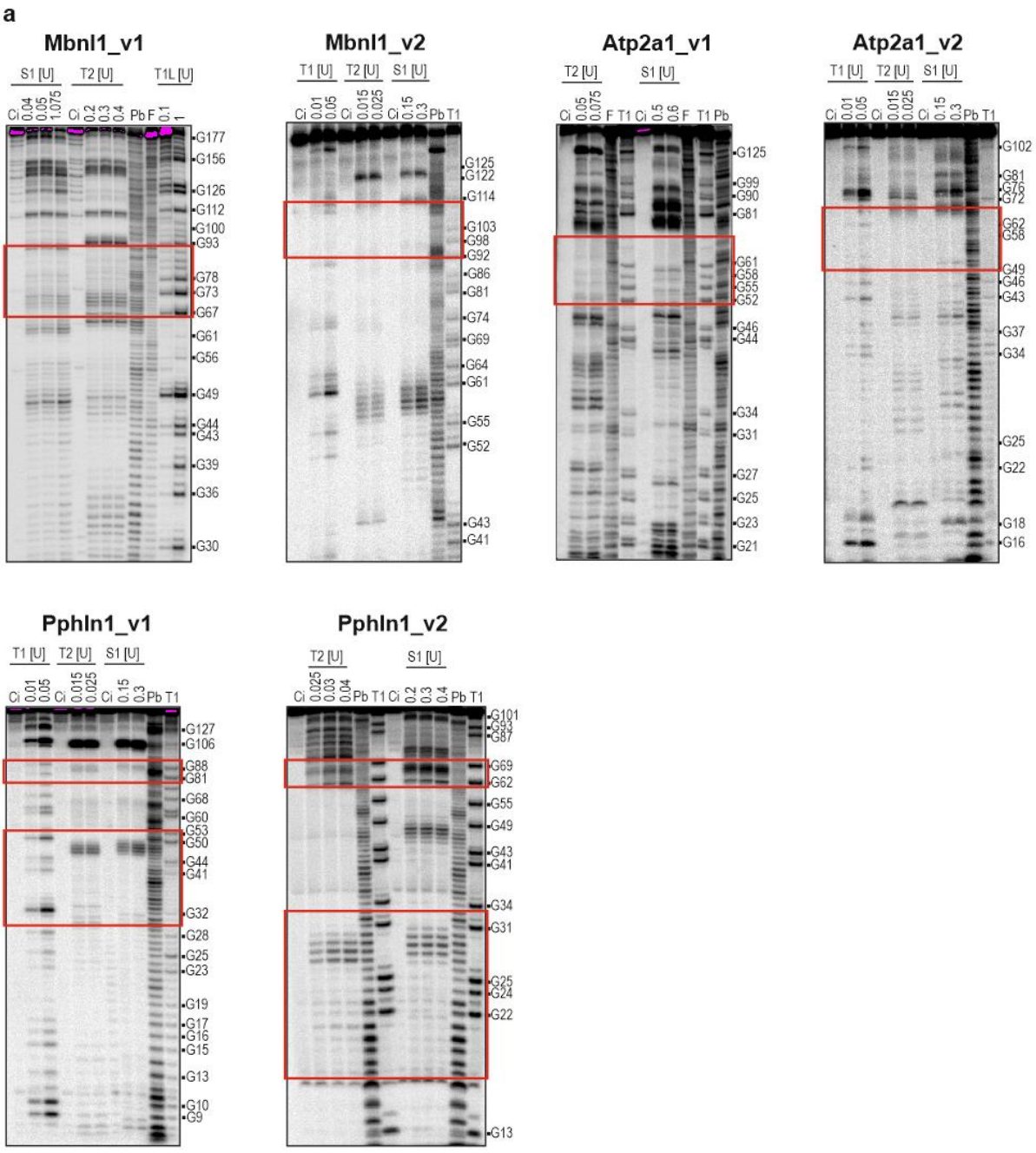
SUPPLEMENTARY FIGURES

Supplementary Figure S1. The secondary structure of RNA sequence length variants of the same transcripts.
Supplementary Figure S2. Different structures of the same RNA binding site affect MBNL1 binding and activity.
Supplementary Figure S3. The secondary structure determination of single-stranded RNAs.
Supplementary Figure S4. Distance between YGCY motifs modulates MBNL-mediated splicing regulation but not binding affinity.
Supplementary Figure S5. Stimulating or inhibitory effect of certain UGCU positions on MBNL binding and splicing activity.
Supplementary Figure S6. Lack of complex formation between rMBNL1 and ds RNAs containing two UGCU motifs.
Supplementary Figure S7. Stimulatory or inhibitory effect of position of three UGCU-containing RNAs on MBNL binding affinity.
Supplementary Figure S8. Significance of specific RNA structural determinants in natural MBNL targets - Calm3 example.
Supplementary Figure S9. Significance of specific RNA structural determinants in natural MBNL targets – Mbnl2 example.
Supplementary Figure S10. The effect of competition between MBNL paralogs on endogenous transcripts.
Supplementary Figure S11. The effect of competition between MBNL paralogs depends on RNA regulatory element organization.

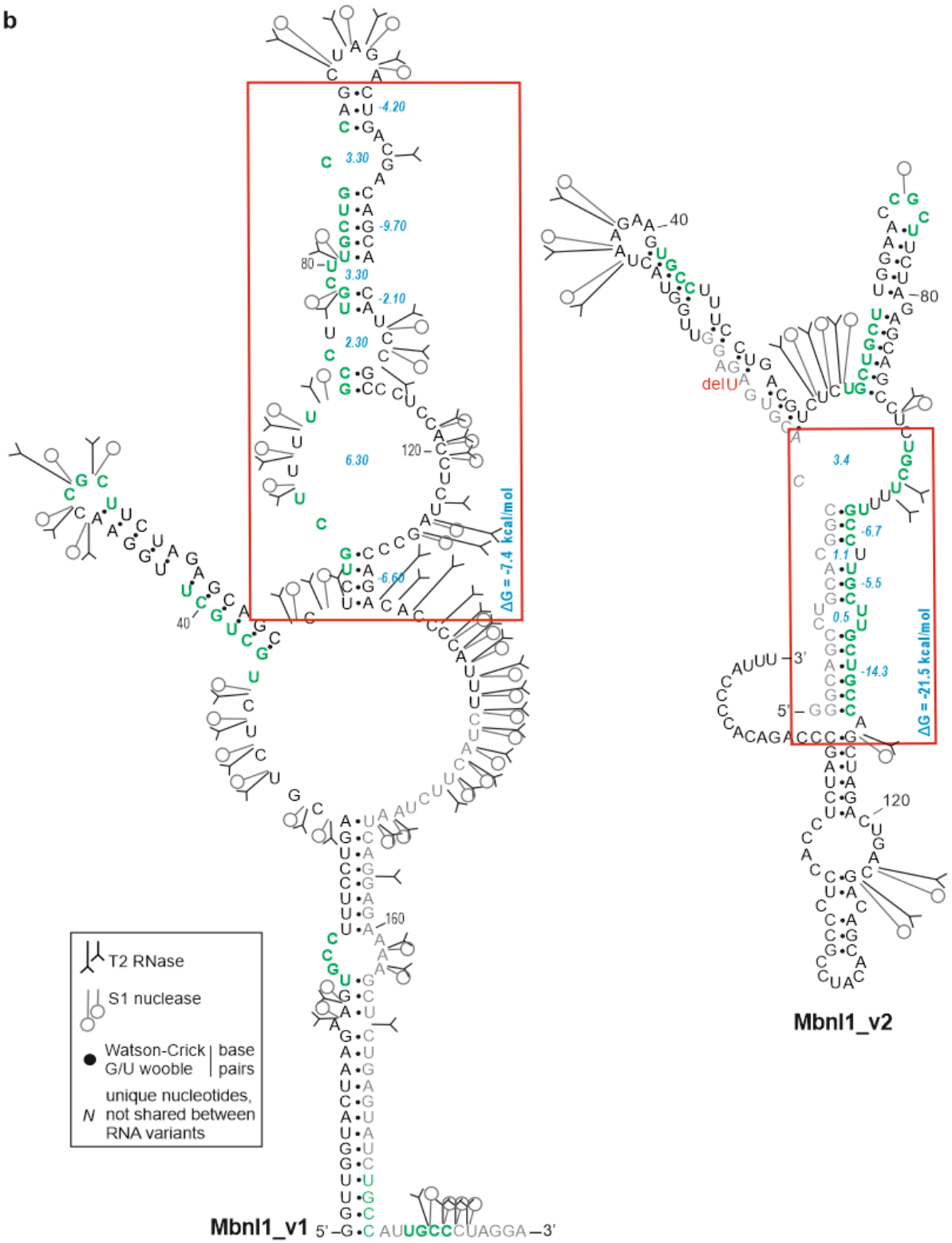
SUPPLEMENTARY MATERIALS

Supplementary Table 1. Sequences of single stranded DNA templates and primers for biochemical and *in cellulo* analysis.
Supplementary Table 2. Sequences of primers for splicing analysis of endogenous transcripts and RNA expression level.

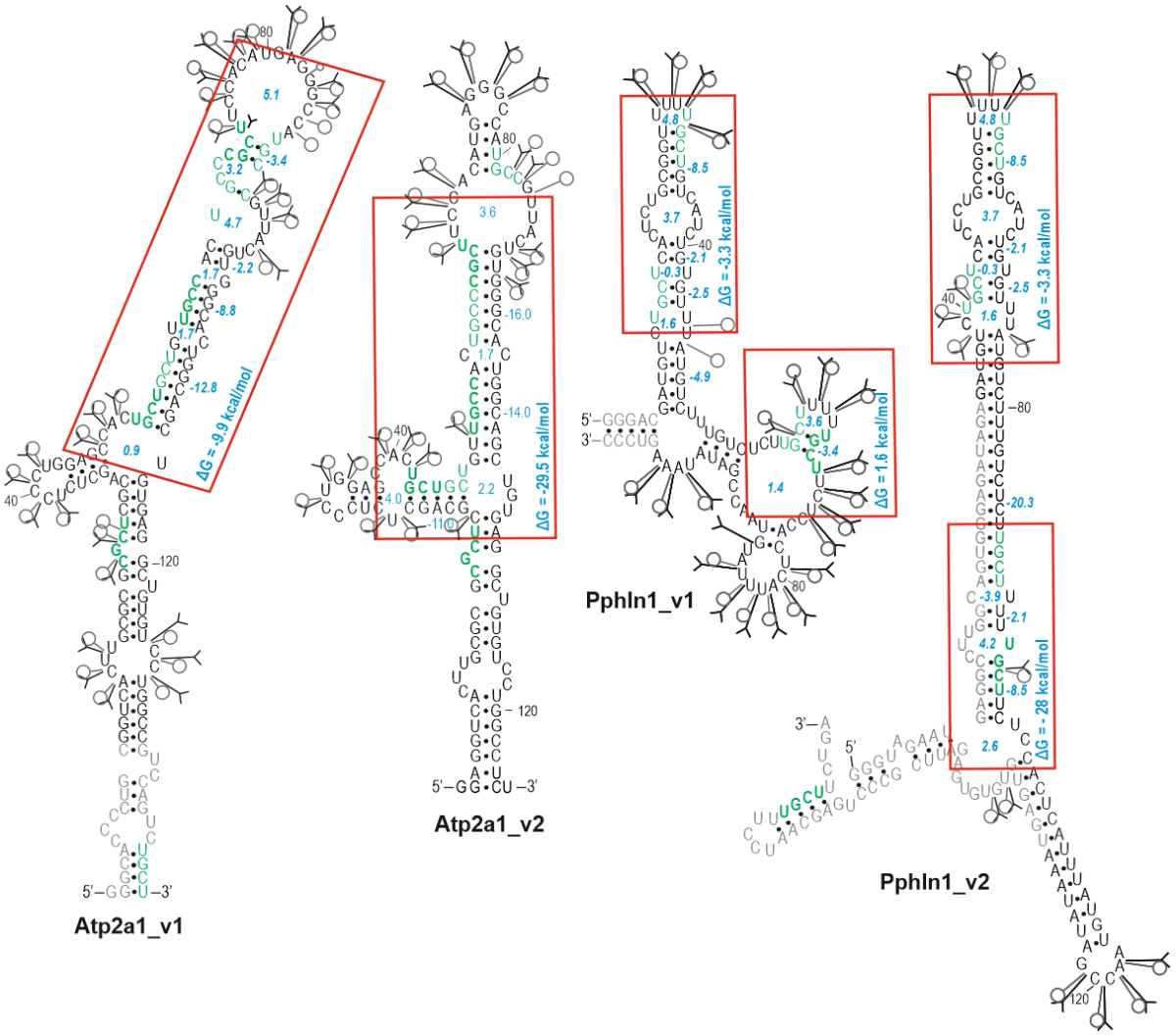
SUPPLEMENTARY FIGURES



Supplementary Figure S1



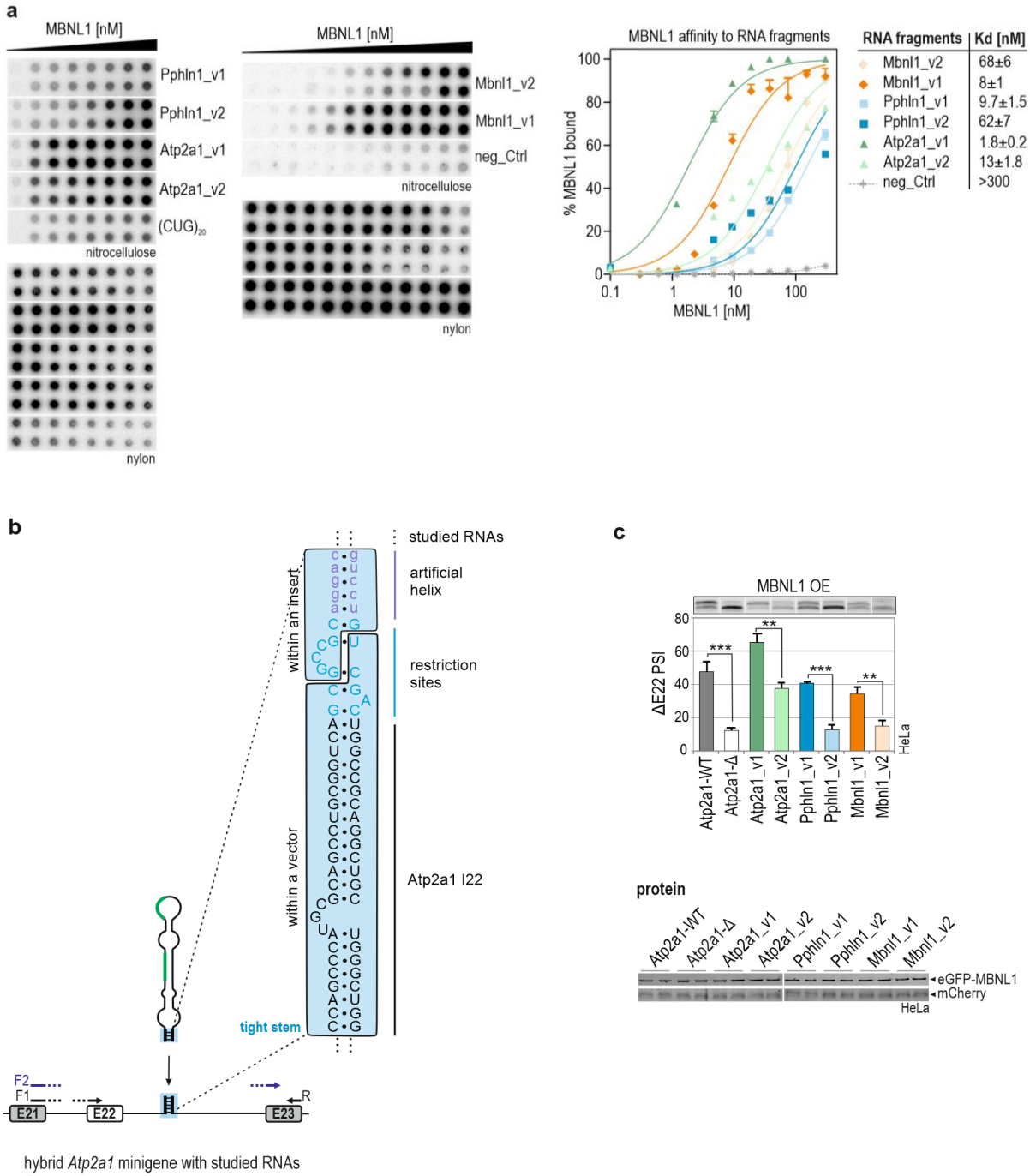
Supplementary Figure S1



Supplementary Figure S1 (refers to Figure 1)

The secondary structure of RNA sequence length variants of the same transcripts.

(**a**) The secondary structure probing of analyzed 5'-^{32}P-labeled RNAs using T1 and T2 RNases as well as S1 endonucleases, which recognize and cleave single-stranded RNA regions (RNase T1 recognizes exclusively G-residues while T2 and S1 recognize all unpaired nucleotides; U, enzyme unit). Ci, untreated RNA samples; Pb, lead ladder; T1L, RNase T1 in denaturing conditions as a G-residue ladder; particular guanosines are marked as G on the right of each electrophoretogram. (**b**) Proposed secondary structures of RNAs with MBNL1-binding motifs (YGCY) marked in green and in bold if conserved between mouse and human. Previously defined (1-3) functional MBNL1-binding motifs are marked with red boxes. The optimal thermodynamic stability of selected RNA modules is expressed in Gibbs free energy (ΔG) in kcal/mol for the reaction at 37°C using RNAfold software (4) and marked in blue. Unique nucleotides, not shared between RNA variants, are marked in grey.

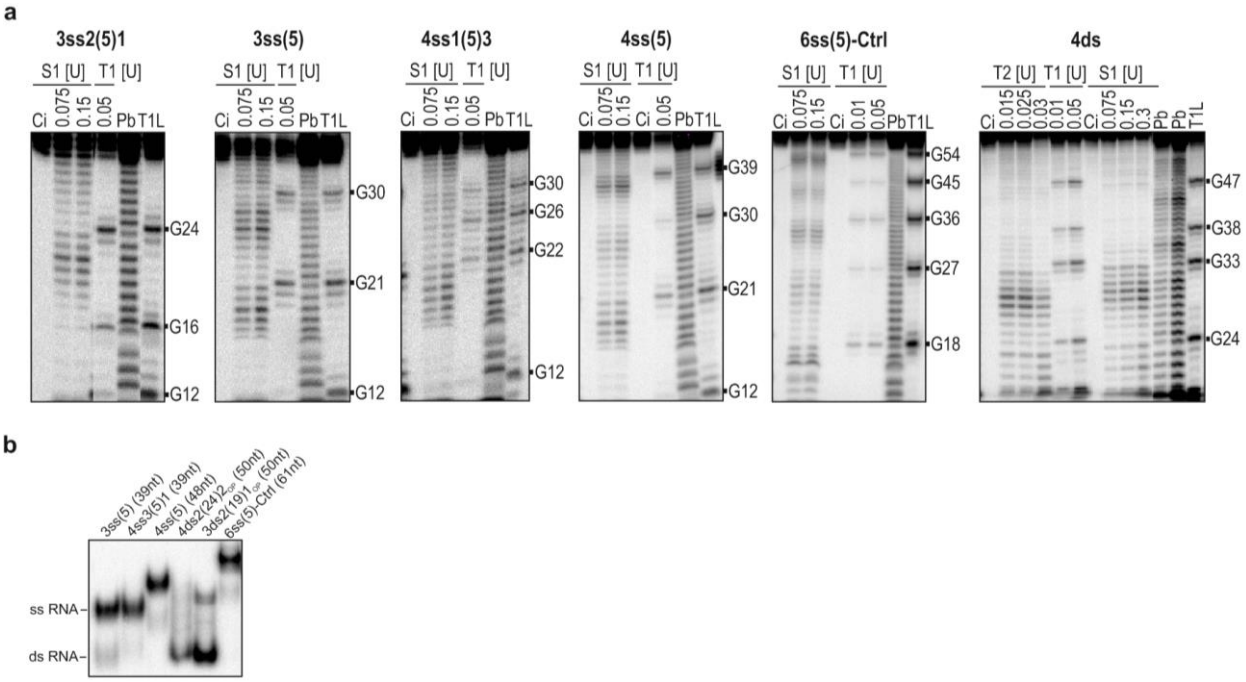


Supplementary Figure S2

Supplementary Figure S2 (refers to Figure 1)

Different structures of the same RNA binding site affect MBNL1 binding and activity.

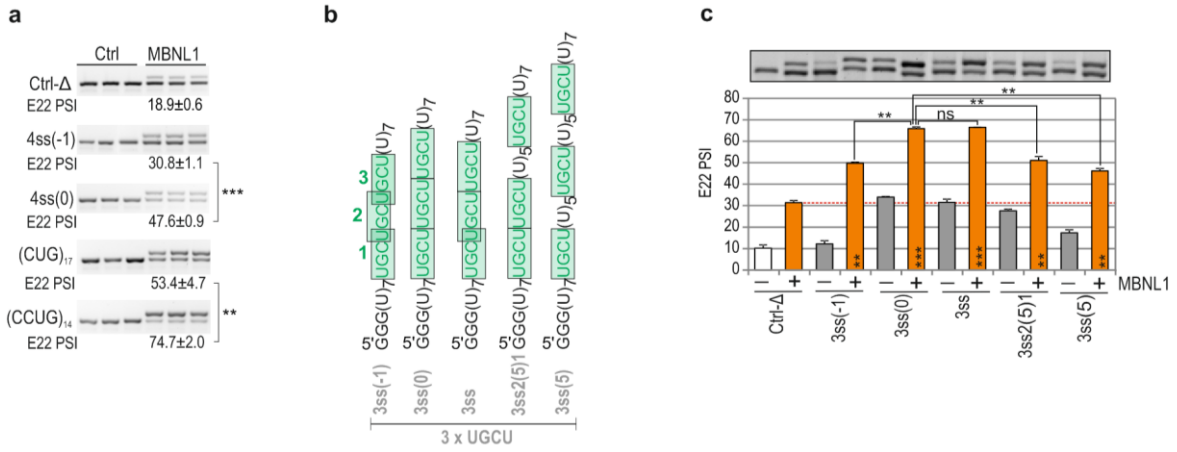
(**a**) *On the left*, raw data from a biochemical assay based on a double-membrane filtration method performed in the presence of 0.1 nM 5'-^{32}P-labeled RNAs and different concentrations of recombinant MBNL1 (rMBNL1) (for Pphln1_1,2 and a negative control, neg_Ctrl: 300, 75, 37.5, 18.7, 9.4, 4.7, 2.3, 1,2, 0 nM; for Atp2a1_1: 300, 75, 37.5, 18.7, 9.4, 4.7, 1.2, 0; for Atp2a1_2: 300, 150, 75, 37.5, 18.7, 9, 4.7, 0; for Mbnl1_1,2: 75, 37.5, 18.7, 9.4, 4.7, 2.3, 1,2, 0.6, 0.3, 0; for a positive control, $(CUG)_{20}$: 75, 37.5, 18.7, 9.4, 4.7, 2.3, 1,2, 0 nM of rMBNL1). A nitrocellulose membrane holds signals of MBNL-RNA complexes, whereas a nylon membrane shows amount of free RNA (see also Figure 1b with schematic presentation of the method). Neg_Ctrl constitutes an RNA sequence with no YGCY motifs. *On the right*, quantification of the biochemical assay showing diverse affinity of rMBNL1 to studied RNA fragments; dissociation constant (Kd) value of MBNL1-RNA complexes is listed on the right side of the chart; n = 2 for each protein concentration. (**b**) A schematic illustration of the secondary structure organization of a stable helical region formed by Atp2a1 I22, which preserves in cells the RNA secondary structure of inserts. The region constitutes a 14-bp-long helix derived from Atp2a1 I22, restriction sites used for cloning and a 5-bp-long artificial helix. It is distant by 12-90-nt from the MBNL-binding sites. The arrows indicate two forward and one reverse primers which were used in a 0.3:1:1 ratio for splicing analysis of hybrid minigenes. (**c**) *Upper panel*, RT-PCR showing the splicing response of ~~selected~~ hybrid *Atp2a1* minigenes upon overexpression (OE) of MBNL1-eGFP fusion protein in HeLa cells. Control cells were treated with eGFP expression vector and the average ΔPSI values were calculated from three independent experiments with SD. *Bottom panel*, image of SDS-polyacrylamide gel presenting the level of eGFP-MBNL1 fusion protein. MCherry as the loading control. The signal from the gels was scanned following excitement of eGFP and mCherry fluorescence, respectively (protein samples were not heat-denatured before gel loading).



Supplementary Figure S3 (refers to Figure 3)

The secondary structure determination of single-stranded RNAs.

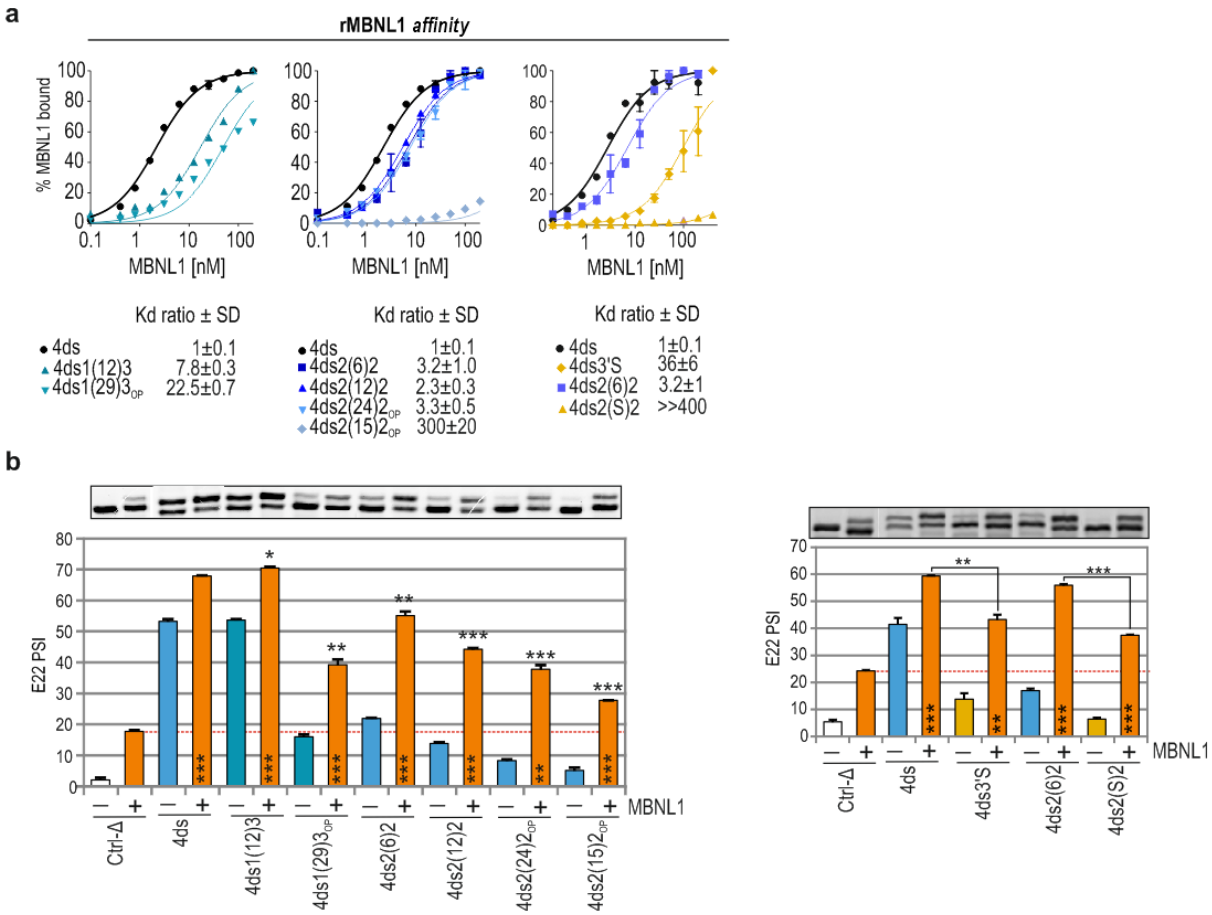
(**a**) The structure probing of representative RNA samples designed as single-stranded (ss) and subjected to limited cleavage with the S1 endonuclease recognizing ss RNA regions. Ci, untreated RNA samples; Pb, lead ladder; T1L, RNase T1 in denaturing conditions as a G ladder; particular guanosines are marked as G on the right of each electrophoretogram. 4ds RNA sample constitutes a double-stranded (ds) RNA control. (**b**) The electrophoretogram showing slower migration rate of representative ss RNAs in a native PA gel compared to ds RNA controls.



Supplementary Figure S4 (refers to Figure 3)

Distance between YGCY motifs modulates MBNL-mediated splicing regulation but not binding affinity.

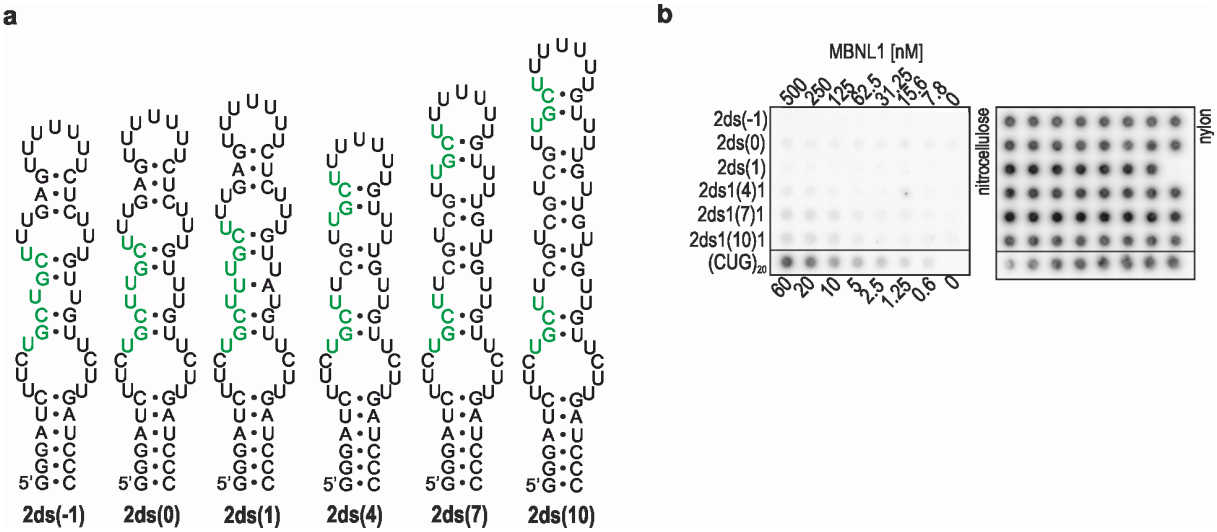
(**a**) RT-PCR showing the splicing response of hybrid minigenes expressing RNA regulatory elements distinguished by either overlapping or sequential arrangement of distinct number of UGCU motifs. The analysis was conducted in DKO MEFs co-transfected with hybrid minigenes and an either MBNL1 or eGFP expression vector (Ctrl). Ctrl-Δ, transfection with *Atp1a1*-Δ minigene; n = 3. (**b**) The secondary structures of studied ss RNA regulatory elements with differently interspaced three UGCU motifs marked in green. The RNA names correspond to the number of RNA motifs (3), structure type (ss) number of consecutive UGCU motifs and distance between them which is marked in brackets. (**c**) RT-PCR results showing the splicing response of hybrid minigenes expressing ss RNA regulatory elements upon MBNL1 (marked in orange) or eGFP OE (marked in white and grey) in HeLa cells. Vertical asterisks denote the statistical significance of results in comparison to a control experiment (Ctrl-Δ; transfection with *Atp1a1*-Δ minigene).



Supplementary Figure S5 (refers to Figure 4)

Stimulating or inhibitory effect of certain UGCU positions on MBNL binding and splicing activity.

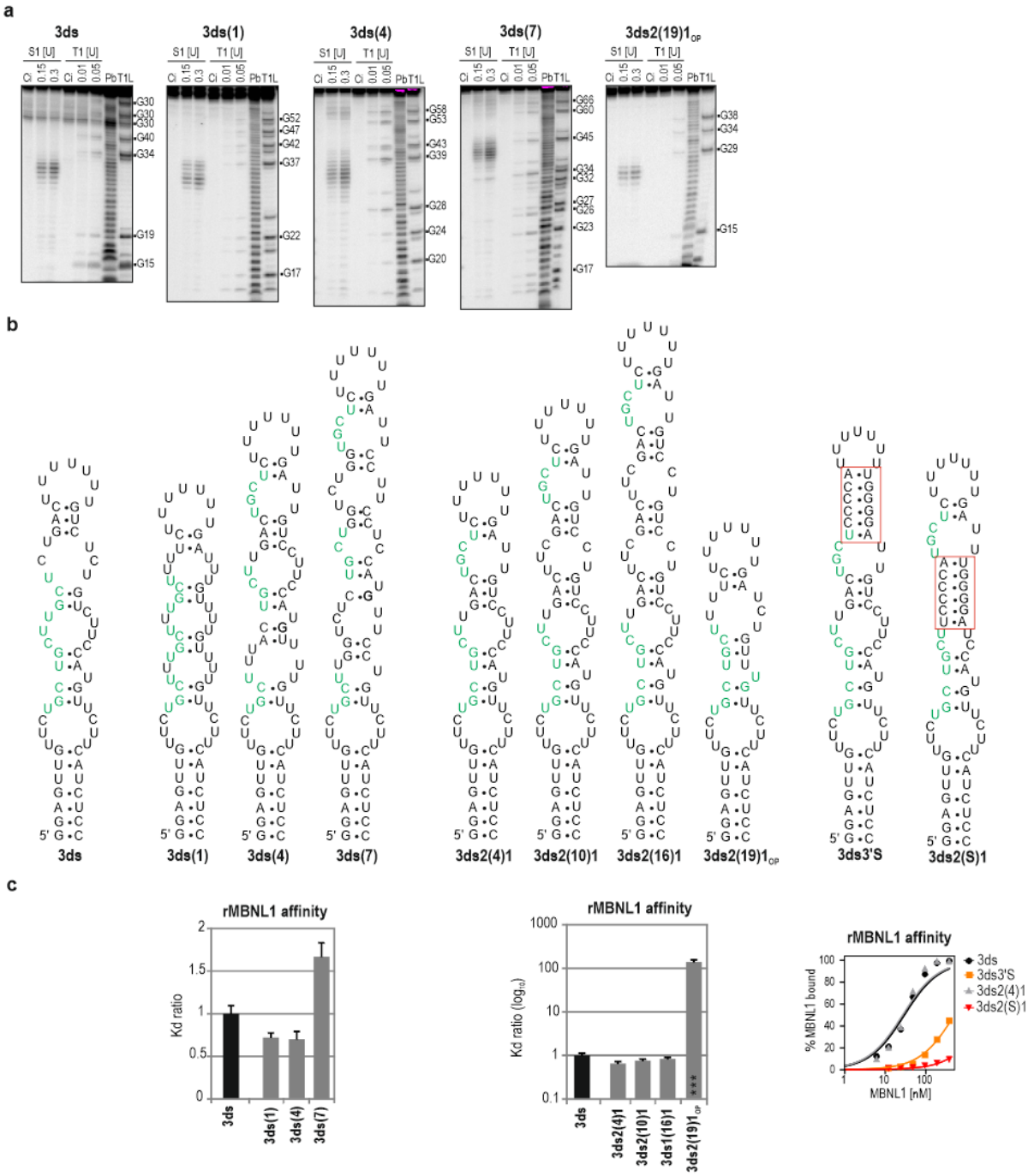
(**a**) Quantification of the biochemical assay showing rMBNL1 affinity to studied ds RNAs. Below each chart, relative Kd of rMBNL1-RNA complexes normalized to the Kd value for the 4ds RNA molecule, n = 2 for each protein concentration (in the range of 0-200 nM of rMBNL1). (**b**) RT-PCR results showing the splicing response of hybrid minigenes expressing ds RNA regulatory elements upon either endogenous level of MBNLs or MBNL1 OE. Vertical or horizontal asterisks denote the statistical significance of results in comparison to a control experiment (Ctrl-Δ; transfection with *Atp1a1*-Δ minigene) or 4ds construct, respectively.



Supplementary Figure S6 (refers to Figure 4)

Lack of complex formation between rMBNL1 and ds RNAs containing two UGCU motifs.

(**a**) The secondary structures of ds RNAs predicted *in silico* and containing two UGCU motifs marked in green. The name of RNAs corresponds to the number of RNA motifs (2), structure type (ds) and distance between consecutive UGCU motifs which is marked in brackets. (**b**) Raw data from a biochemical assay performed in the presence of 0.1 nM 5'-^{32}P-labeled RNAs and increasing concentrations of rMBNL1. $(CUG)_{20}$ constitutes a positive control.

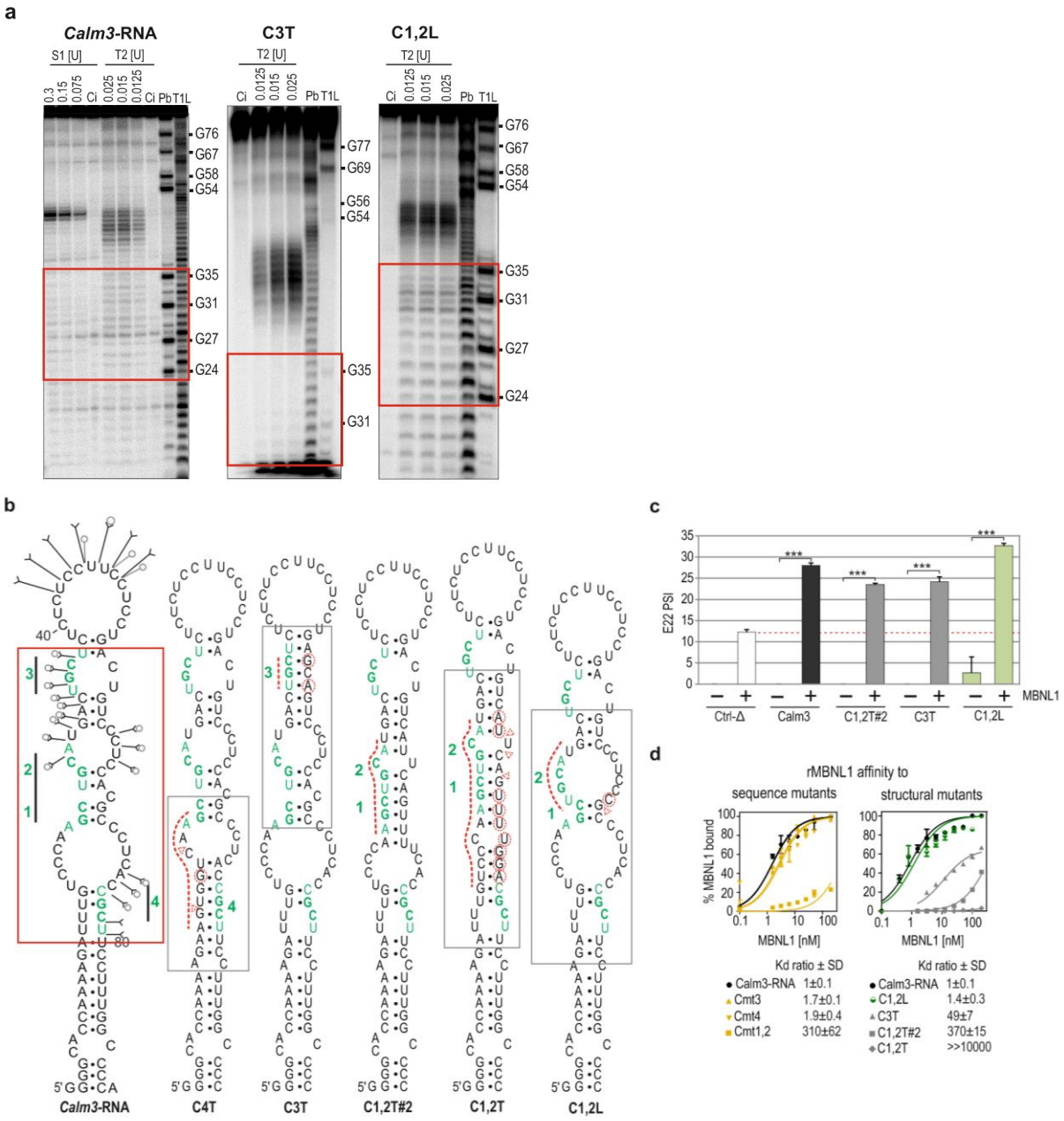


Supplementary Figure S7 (refers to Figure 4)

Stimulatory or inhibitory effect of position of three UGCU-containing RNAs on MBNL binding affinity.

(**a**) The secondary structure probing of analyzed 5'-^{32}P-labeled ds RNAs containing three UGCU motifs using T1 RNase (recognizing unpaired G-residues) and S1 endonucleases which cleave single-stranded RNA sequences. U, enzyme unit; Ci, untreated RNA samples; Pb, lead ladder; T1L, RNase T1 in denaturing conditions as a G-residue-specific ladder; particular guanosines are marked as G on the right of each electrophoretogram. The RNA names correspond to the number of RNA motifs (3), structure type (ds) and distance between consecutive UGCU motifs, which is marked in brackets and surrounded by the number of adjacent motifs. (**b**) Proposed secondary structures of RNAs with UGCU motifs marked in green. (**c**) Quantification of the biochemical assay showing relative rMBNL1 affinity

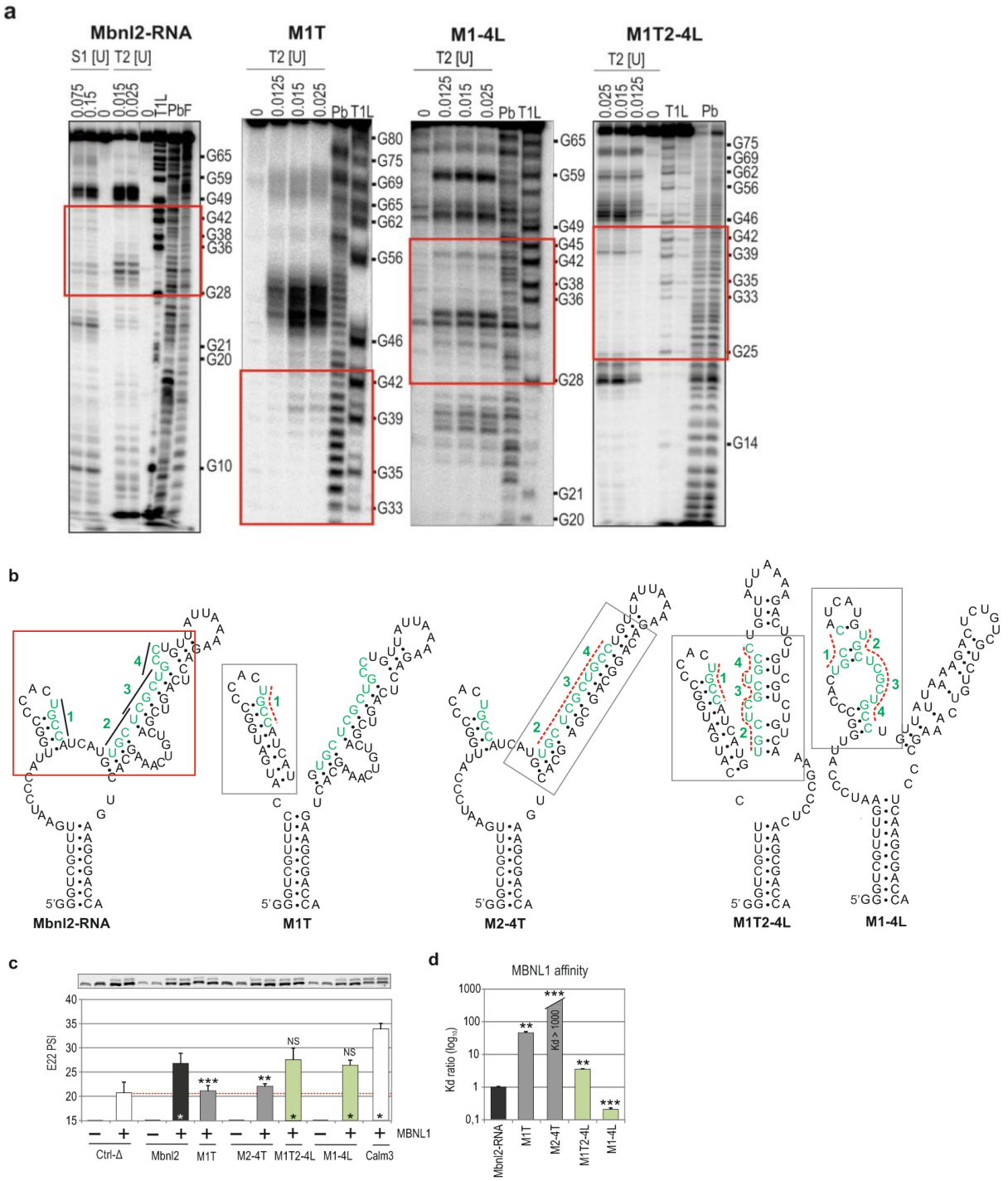
to RNA fragments and normalized to the Kd for the 4ds RNA molecule; n = 2 for each protein concentration (in the range of 0-200 nM of rMBNL1).



Supplementary Figure S8 (refers to Figure 5)

Significance of specific RNA structural determinants in natural MBNL targets - Calm3 example.

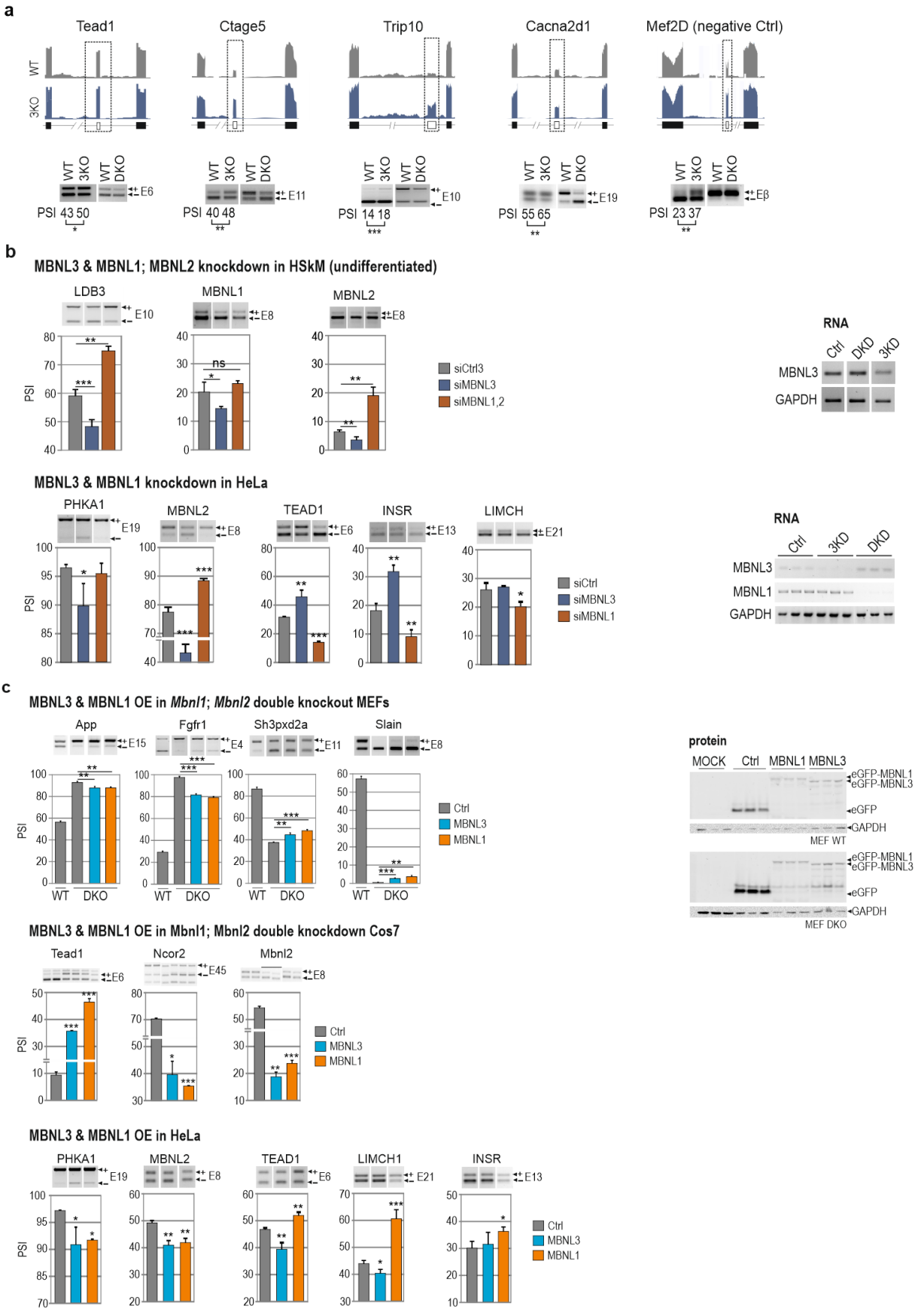
(**a** and **b**) As in Supplementary Figures S1a and S1b but for the Calm3-RNA fragment and its structural mutants. (**c**) RT-PCR results showing the splicing response of Calm3 minigene and selected structural mutants upon MBNL1 or eGFP OE. A red dashed line constitutes a splicing response threshold of hybrid minigene lacking a natural MBNL-binding region within I22 pre-mRNA (*Atp2a1*-Δ). (**d**) Quantification of the biochemical assay showing rMBNL1 binding affinity to intact Calm3-RNA as well as HGCH motif and RNA secondary structure mutants. Below each chart, the relative Kd of rMBNL1-RNA complexes normalized to the Kd value for Calm3-RNA molecule; n = 2 for each protein concentration (in the range of 0-200 nM of rMBNL1).



Supplementary Figure S9 (refers to Figure 5)

Significance of specific RNA structural determinants in natural MBNL targets - Mbnl2 example.

(**a** - **d**) As in Supplementary Figure S8 but for the RNA regulatory element derived from intron 4 of Mbnl2 transcript and its structural mutants.

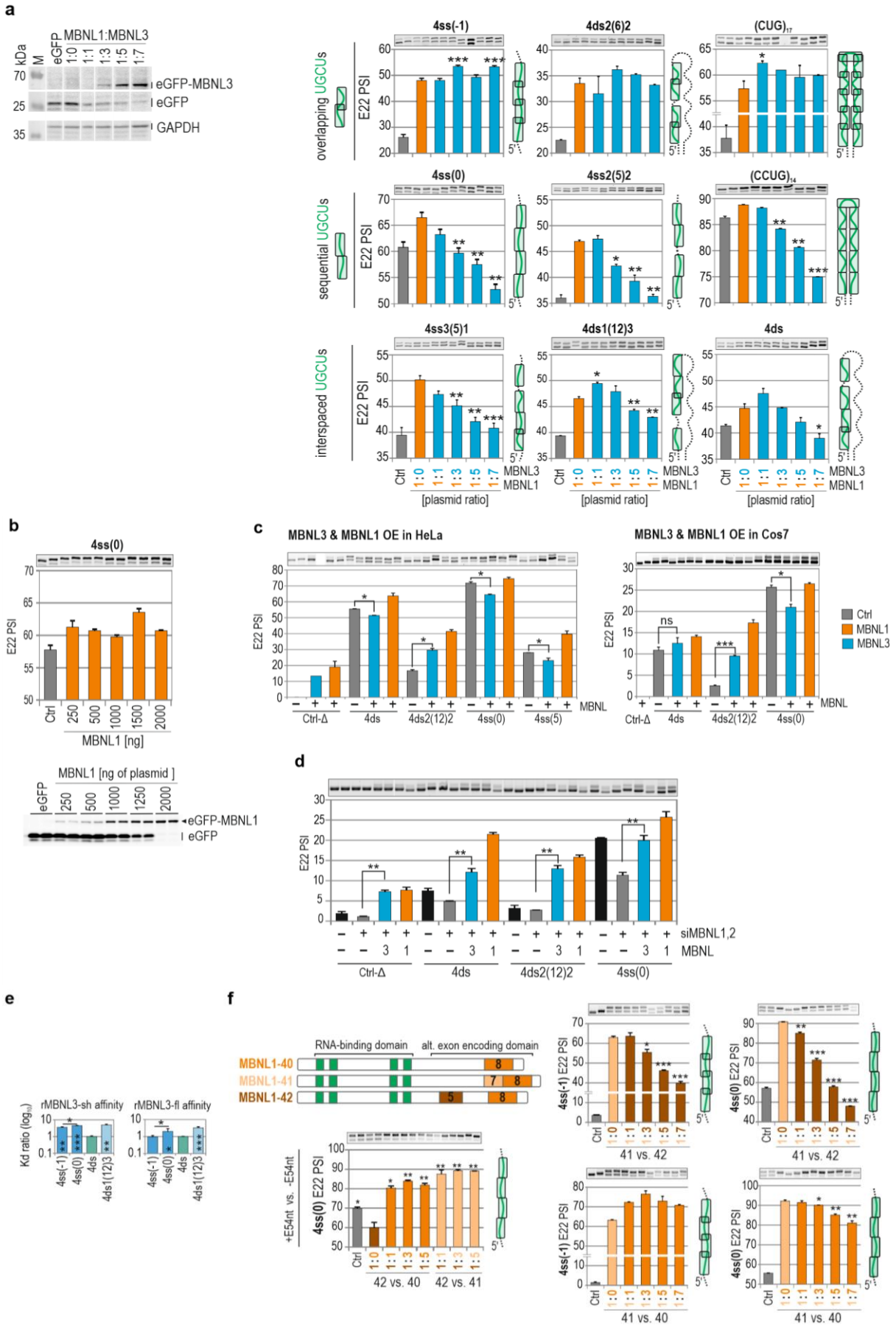


Supplementary Figure S10

Supplementary Figure S10 (refers to Figure 6)

The effect of competition between MBNL paralogs on endogenous transcripts.

(**a**) RNA-seq read coverage across MBNL-dependent alternative exons, indicating their adult-like splicing profiles in Mbnl3 knockout (3KO) myoblasts compared to wild-type cells and *Mbnl1*; *Mbnl2* DKO muscles; alternative and constitutive exons are marked with a white or grey box, respectively; n = 2. Note that the spicing effect observed in 3KO is opposite to the effect in DKO. (**b**) *Left panel*, RT-PCR results of splicing changes in MBNL1-dependent alternative exons whose direction was opposite upon MBNL3 silencing compared to MBNL1; MBNL2 double knockdown with respect to cells treated with control siRNA in human skeletal muscles (HSkMs) and in HeLa cells; n = 3 or 2, respectively. *Right panel*, RT-PCR results showing a reduced expression level of MBNL3 by nearly 40% and 80% upon siMBNL3 application in HSkM and HeLa cells, respectively*;* GAPDH levels are shown as a loading control; n = 3. (**c**) *Left panel*, RT-PCR results of splicing changes of four MBNL1 dependent alternative exons which direction was alike upon MBNL3 and MBNL1 OE in *Mbnl1*; *Mbnl2* DKO MEFs and Mbnl1; Mbnl2 double knockdown Cos7 cells; n = 3 or 2, respectively. RT-PCR results of splicing changes in MBNL1-dependent alternative exons whose direction was the same (PHKA1, MBNL2) or opposite (TEAD1, LIMCH1) upon MBNL3 and MBNL1 OE in the background of endogenous MBNLs in HeLa cells; Ctrl, the results from eGFP transfected cells (marked in grey); *Right panel*, eGFP-MBNL3 and eGFP-MBNL1 protein level was determined by Western-blotting; GAPDH protein levels are shown as the loading controls; n = 3.



Supplementary Figure S11

Supplementary Figure S11 (refers to Figure 6)

The effect of competition between MBNL paralogs depends on RNA regulatory element organization.

Cellular competition assay between MBNL1 and MBNL3 paralogs. (**a**) *Left panel*, Western-blot analysis depicting a gradual increase in eGFP-MBNL3 protein level in HeLa cells. GAPDH is the loading control. *Right panel*, RT-PCR results showing the splicing response of pre-mRNAs with distinctly arranged UGCU motifs in intronic regulatory elements in constant background of MBNL1 OE (marked in orange) and upon MBNL3 increasing concentration (marked in blue) in HeLa cells. MBNL1 and MBNL3 paralogs are identical with respect to splicing isoforms (lacking alternative exon 5 and 7 encoding residues); Ctrl, the results from eGFP transfected cells (marked in grey); the asterisks denote the statistical significance of results compared to cells with MBNL1 OE only (marked in orange); n = 2. On the right of each graph a scheme of a structural organization of UGCU motifs is presented. (**b**) *Top panel*, RT-PCR showing unchanged splicing response of 4ss(0) hybrid minigene upon increased concentration of MBNL1 expression vector. *Bottom panel*, the detection of MBNL1 protein level conducted by eGFP fluorescence read. (**c**) RT-PCR showing the splicing response of selected hybrid minigenes expressing different RNA regulatory elements upon MBNL1 or MBNL3 OE in HeLa (*left panel*) and Cos7 (*right panel*) cells in the background of endogenous MBNLs. Ctrl, the results from eGFP transfected cells (marked in grey); Ctrl-Δ; transfection with *Atp1a1*-Δ minigene. (**d**) RT-PCR results of splicing changes of selected hybrid minigenes upon MBNL3 and MBNL1 OE in Mbnl1; Mbnl2 double knockdown Cos7 cells; n = 2. (**e**) The quantification of the biochemical assay showing the affinity of short (RNA-binding domain only) and full-length rMBNL3 (rMBNL3-sh and rMBNL3-fl, respectively) to overlapping and sequentially arranged UGCU motifs. The Kd was normalized to the 4ds RNA molecule; n = 2–4 of each protein concentration (in the range of 0-200 nM of rMBNL3). (**f**) *Top left panel*, a schematic representation of MBNL1 splicing isoforms differing in the presence of alternative exon 5 and 7 encoding residues (marked in brown and beige, respectively) but sharing the same RNA binding domain composed of four ZFs (marked in green). *Bottom left panel*, as in Figure S11a but for MBNL1 isoform pairs. MBNL1-40, MBNL1-41, and MBNL1-42 are marked in orange, beige and brown, respectively. The asterisks denote the statistical significance of results compared to cells with overexpression of only one isoform; n = 2.

SUPPLEMENTARY MATERIALS

Supplementary Table S1

Sequences of single stranded DNA templates (_Tem) and primers for biochemical (_F/R & _TF/TR) and in cellulo (_cF/R) analysis. The sequence of T7 promoter or complementary to it is underlined.

Name	sequence 5' -> 3'
Atp2a1_F	GGTGGGATTTTCTCCCAACCT
Atp2a1_R	GCTGTATCTCCTTGCGCATC
Atp2a1_v1TF	TAATACGACTCACTATAGGGCACCCCTGCGGTCACT
Atp2a1_v1TR	AGCAGACTGGACGGCCA
Atp2a1_v1cF	ACTGCGGCCGCAGGACGCACCCCTGCGGTCACTTGCGCGCGCT
Atp2a1_v1cR	ATCGTCGACAGGACAGCAGACTGGACGGCCAGGACACAGCC
Atp2a1_v2TF	TAATACGACTCACTATAGGAGGAGGTCACCTTGCGCGCGCTCGCA
Atp2a1_v2TR	AGAGGCCAGGACACAGCCTCACAG
Atp2a1_v2cF	ACTGCGGCCGCAGGACGGTCACTTGCGCGCGCTCGCAGCTCT
Atp2a1_v2cR	ATCGTCGACAGGACGGCCAGGACACAGCCTCACAGCTGCC
Pphln1_F	AGGTAGTTGGTTGTCTGGGC
Pphln1_R	CTCAGACTTCGCCTTCCAGA
Pphln1_v1TF	TAATACGACTCACTATAGGGTAGAATGAGGCCTTGGCAG
Pphln1_v1TR	CAGAAGCAAAGGATTGCTCA
Pphln1_v1cF	ACTGCGGCCGCAGGACAGGACGTAGAATGAGGCCTTGGCA
Pphln1_v1cR	ATCGTCGACAGGACCAGAAGCAAAGGATTGCTC
Pphln1_v2TF	TAATACGACTCACTATAGGGACGATGTCTGCTCACTCTGCGGT
Pphln1_v2TR	GGGACTTTATATCGGTTACATAAATG
Pphln1_v2cF	ACTGCGGCCGCAGGACGATGTCTGCTCACTCTGCGGT
Pphln1_2vcR	ATCGTCGACAGGACTTTATATCGGTTACATAAATG
Mbnl1_1F	GATGGCTGGCTGCAATATGCC
Mbnl1_1R	CCTAGGGCAATGGCAGATACTC
Mbnl1_v1TF	TAATACGACTCACTATAGGGTTGGTACTAAGAAGTGCCT
Mbnl1_v1TR	CCTAGGGCAATGGCAGATACTC
Mbnl1_v1cF	ACTGCGGCCGCAGGACGTTGGTACTAAGAAGTGCCT
Mbnl1_v1cR	ATCGTCGACAGGACCCTAGGGCAATGGCAGATACTC
Mbnl1_2F	GGAGTCGCGATCCCACAAT
Mbnl1_2R	AAATGGGGTGTCTGGGCTAGAG
Mbnl1_v2TF	TAATACGACTCACTATAGGGCAGCCTGCACGGCCACGTG
Mbnl1_v2TR	AAATGGGGTGTCTGGGCTAGAG
Mbnl1_v2cF	ACTGCGGCCGCAGGACGGGCAGCCTGCACGGCCACGTGAGAG
Mbnl1_v2cR	ATCGTCGACAGGACAAATGGGGTGTcTGGGcTAGAGGTGG
Mbnl1_v2AGmcF	CGGCCACGTGTGGGTTGGTACTAAG
Mbnl1_v2AGmcR	CTTAGTACCAACCCACACGTGGCCG
Neg_Ctrl_F	CCAGCTTGTGGAGGGCCACTT
Neg_Ctrl_R	GCTGCCTCTCCCAGTGGTCTG
Neg_Ctrl_TF	TAATACGACTCACTATAGGGAGGCCGACTGTAAGGGT
Neg_Ctrl_TR	CACCCTTTCTGCTGGAGAGGC
CTG17_cF	GGCCGCGGCCC(CTG)17GGGCCG
CTG17_cR	TCGACGGCCC(CAG)17GGGCCGC
CTGG14_cF	GGCCGCGGCCC(CTGG)14GGGCCG

CTGG14_cR	TCGACGGCCC(CCAG)14GGGCCGC
Calm3_F	GACTGACGGACTCAACCCTG
Calm3_R	CTGAGCAGGGACAGGAATGG
Calm3_TF	TAATACGACTCACTATAGGGCACCAAAAGATTTGTCCCA
Calm3_TR	GGGGCCAAAGGAAGCGGTGAG
Calm3_cF	ACTGCGGCCGCAGGACGGGCACCAAAAGATTTGTCCCA
Calm3_cR	ATCGTCGACAGGACGGGGCCAAAGGAAGCGGTGAG
C4T_TF	TAATACGACTCACTATAGGGCACCAAAAGAAGTGGTCAAGCT
C4T_TR	GGGGCCAAAGGAAGCGGTGAG
C3T_TF	TAATACGACTCACTATAGGGCACCAAAAGATTTGTCCCA
C3T_TR	GGGGCCAAAGGAAGCGGTGAGGGCGTGGAGGGACTGCTCAGGA
C1,2T_TF	TAATACGACTCACTATAGGGCACCAAAAGATTTGTCCCA
C1,2T_TR	GGGGCCAAAGGAAGCGGTGAAAACTGAATGACAGT
C1,2L_TF	TAATACGACTCACTATAGGGCACCAAAAGATTTGTCCCAA
C1,2L_TR	GGGGCCAAAGGAAGCGGTGAGGGCGGGAGGG
C1,2T#2_TF	TAATACGACTCACTATAGGGCACCAAAAGATTTGTCCCA
C1,2T#2_TR	GGGGCCAAAGGAAGCGTCCAAAACTGAATGACAGT
Mbnl2_F	TGTGTATGCAAGCTGCTGTT
Mbnl2_R	ACAAATCTCGGCTGGTTCTCT
Mbnl2_TF	TAATACGACTCACTATAGGGTCGTTTGAATCCCATTG
Mbnl2_TR	GGTCGCTTCAGTGCTTTGACAGT
M1T_TF	TAATACGACTCACTATAGGGTCGTTTCCATGATGGCCCAC
M1T_TR	GGTCGCTTCAGTGCTTTGACAGT
M2-4T_TF	TAATACGACTCACTATAGGGTCGTTTGAATCCCATTGG
M2-4T_TR	GGTCGCTTCAGTGCTCGCTGCCTGTCTTTTAA
M1T2-4L_TF	TAATACGACTCACTATAGGGTCGTTTCCATGATGGCCCAC
M1T2-4L _TR	GGTCGCTTGAGGGCTTTGACAGACGACAGAGTCTTT
M1-4L_TF	TAATACGACTCACTATAGGGTCGTTTGAATCCCATTGG
M1-4L_TR	GGTCGCTTGAGGGCTTTGACAGACGACAGAGTCTTTTAA
T7_3G_F	TAATACGACTCACTATAGGGG
2ss	RNA was purchased from Future Synthesis
3ss	RNA was purchased from Future Synthesis
3ss2(5)1	AAAAAAAAGCAAAAAAAGCAAGCAAAAAAAAACCCTATAGTGAGTCGTATTA
3ss(5)	AAAAAAAACGAAAAAAAGCAAAAAAAGCAAAAAAAAACCCTATAGTGAGTCGTATTA
4ss(-1)	RNA was purchased from Future Synthesis
4ss(0)	RNA was purchased from Future Synthesis
4ss	RNA was purchased from Future Synthesis
4ss1(5)3	AAAAAAAAGCAAGCAAGCAAAAAAAAGCAAAAAAAAACCCTATAGTGAGTCGTATTA
4ss(5)	AAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAACCCTATAGTGAGTCGTATTA
6ss(5)-Ctrl	AAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAACCCTA TAGTGAGTCGTATTA
2ds(-1)_Tem	GGGATTCTTCTGCTGCTTTGAGTTTTTTTCTCTTTGTTGTTCTTGAATCCC
2ds(-1)_TF	TAATACGACTCACTATAGGGATTCTTCTGCT
2ds(-1)_TR	GGGATTCAAGAACAA
2ds(0)_Tem	GGGATTCTTCTGCTTGCTTTGAGTTTTTTTCTCTTTGTTTGTTCTTGAATCCC
2ds(0)_TF	TAATACGACTCACTATAGGGATTCTTCTGCT
2ds(0)_TR	GGGATTCAAGAACAA

2ds(1)_Tem	GGGATTCTTCTGCTTTGCTTTGAGTTTTTTTCTCTTTGTTTTGTTCTTGAATCCC
2ds(1)_TF	TAATACGACTCACTATAGGGATTCTTCTGCT
2ds(1)_TR	GGGATTCAAGAACAA
2ds(4)_TF	TAATACGACTCACTATAGGGATTCTTCTGCTTCGTTGCTTTTTT
2ds(4)_TR	GGGATTCAAGAACAACAAAACAAAAAAAGCAACGAAGCAGAA
2ds(7)_Tem	GGGATTCTTCTGCTTCGTCGTTGCTTTTTTTGTTTTGTTGTTGTTCTTGAATCCC
2ds(7)_TF	TAATACGACTCACTATAGGGATTCTTCTGCT
2ds(7)_TR	GGGATTCAAGAACAA
2ds(10)_TF	TAATACGACTCACTATAGGGATTCTTCTGCTTCGTCGTCGTTGCTTTTT
2ds(10)_TR	GGGATTCAAGAACAACAACAACAAAACAAAAAAAGCAACGACGACGAAG
3ds_TF	TAATACGACTCACTATAGGGATTCTTCTGCTGCTTGCTCCTCTTTTTTT
3ds_TR	GGGAATCAAGAACATGGAAGACAGATCTAAAAAAAAGAGGAGCAAGCAG
3ds(1)_TF	TAATACGACTCACTATAGGAGGTTGTTCTGCTTTGCTTTGCTTTTCTTTTTT
3ds(1)_TR	GGAGGATGAAGAACAAAACAAAACAAATCAAAAAAAGAAAAGCAAAGCAA
3ds(4)_TF	TAATACGACTCACTATAGGAGGTTGTTCTGCTTTACTGCTTGACTGCTCTTTTTT
3ds(4)_TR	GGAGGATGAAGAACAAAACATGGAAGGACAATCAAAAAAAGAGCAGTCAAGCAG
3ds(7)_TF	TAATACGACTCACTATAGGAGGTTGTTCTGCTTGGTCTCTGCTGGTCTGGTGCTCTT
3ds(7)_TR	GGAGGATGAGAACAGGAACATGAGGAAGGAATCAAAAAAAGAGCACCAGACCAGCAGA
3ds2(4)1_TF	TAATACGACTCACTATAGGGATTCTTCTGCTGCTTGACTGCTCTTTTTT
3ds2(4)1_TR	GGGAATCAAGAACATGGAAGGACAATcAAAAAAAGAGcAGTcAAGcAGc
3ds2(10)1_TF	TAATACGACTCACTATAGGGATTCTTCTGCTGCTTGACTTCGACTGCTCTTTTT
3ds2(10)1_TR	GGGAATCAAGAAcATGGAAGGACAGGACAATCAAAAAAAGAGCAGTCGAAGTCA
3ds2(16)1_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTGACTTCGACTTCGACTGCTCTTTTT
3ds2(16)1_TR	GGAGATGAAGAACATGGAAGGACAGGACAGGACAATCAAAAAAAGAGCAGTCGAAGTCG
3ds2(19)1_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTTTCTTTTT
3ds2(19)1_TR	GGAGATGAAGAGCAAACAGATCAAAAAAAGAAAAGCAGCAGAAC
3ds3'S_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTGACTGCTCCCCATTTTTT
3ds3'S_TR	GGAGATGAAGAACATGGAAGGACAATCCCCAAAAAAAATGGGGAGCAGTCAA
3ds2(S)1_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTCCCCATGCTCTTTTT
3ds2(S)1_TR	GGAGATGAAGAACATGGATCCCCAAATCAAAAAAAGAGCATGGGGAAGCA
4ds_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTGCTTTGCTTTTTT
4ds_TR	GGAGATGAAGAACATGGAAGACAGAACAAAAAAAGCAAAGCAAGCAGC
4ds1(12)3_Tem	GGAGTTGTTCTGCTTTGACTTCGACTTGCTGCTTGCTTTTTTTGTTTCATGTTTTGTCCTGTCTTGTTC TCATCTCC
4ds1(12)3_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTTTGACTTCG
4ds1(12)3_TR	GGAGATGAGAACAAGACAGGACAA
4ds1(29)3_Tem	GGAGTTGTTCTGCTTTGACTTCGACTTCATGTTTTGTTTTTTTTGCTTGCTGCTTGTCCTGTCTTGTTC TCATCTCC
4ds1(29)3_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTTTGACTTCG
4ds1(29)3_TR	GGAGATGAGAACAAGACAGGACAA
4ds2(6)2_TF	TAATACGACTCACTATA GGAGTTGTTCTGCTGCTTGACTTTGCTGCTTTTTTC
4ds2(6)2_TR	GGAGATGAAGAACATGGAAGGACAGAACTTGAAAAAAGCAGCAAAGTCAAGCA
4ds2(12)2_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTGACTTCGACTCTGCTGCTTTTT
4ds2(12)2_TR	GGAGATGAAGAACATGGAAGGACAGGACAAAACATGAAAAAAGCAGCAGAGTCGAAG
4ds2(24)2_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTGACTTCGACTTTCATGTTTTTT
4ds2(24)2_TR	GGAGATGAAGAACATGGAAGGACAGGACAAGCAGCAAAAAAAAACATGAAAGTCGAAG
4ds2(15)2_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTTTCTTTTT

4ds2(15)2_TR	GGAGATGAAGAGCAAGCAGATCAAAAAAAGAAAAGCAGCAGAAC
4ds3'S_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTGCTTTGCTTCCCCATTTTTT
4ds3'S_TR	GGAGATGAAGAACATGGAAGACAGAACATCCCCAAAAAAAATGGGGAAGCAAAGC
4ds2(S)2_TF	TAATACGACTCACTATAGGAGTTGTTCTGCTGCTTCCCCATGCTGCTTTTT
4ds2(S)2_TR	GGAGATGAAGAACATGGTCCCCAAAACATGAAAAAAGCAGCATGGGGAAGCA
2ds_cF	GGCCGCAGGACGGAGTTGTTCTGCTTGCTTTGAGTTTTTTTCTCTTTGTTTGTTCTTCATCTCCGTCCTG
2ds_cR	TCGACAGGACGGAGATGAAGAACAAACAAAGAGAAAAAAACTCAAAGCAAGCAGAACAACTCCGTCCTGC
2ss(-1)_cF	GGCCGCGTTTTTTTTTTTGCTGCTTTTTTTTTTTG
2ss(-1)_cR	TCGACAAAAAAAAAAAAGCAGCAAAAAAAAAAAcGC
2ss(0)_cF	GGCCGCGTTTTTTTTTTGCTTGCTTTTTTTTTG
2ss(0)_cR	TCGACAAAAAAAAAGCAAGCAAAAAAAAAAcGC
2ss(5)_cF	GGCCGCGTTTTTTTTTTGCTTTTTTTGCTTTTTTTTTG
2ss(5)_cR	TCGACAAAAAAAAAGCAAAAAAAGCAAAAAAAAAAcGC
3ds_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTGCTCCTCTTTTTTTTAGATCTGTCTTCCATGTTCTTCATCTCCGTCCTG
3ds_cR	TCGACAGGACGGAGATGAAGAACATGGAAGACAGATCTAAAAAAAAGAGGAGCAAGCAGCAGAACAACTCCGTCCTGC
3ss_cF	GGCCGCGTTTTTTTTTTTGCTGCTTGCTTTTTTTTTTTG
3ss_cR	TCGACAAAAAAAAAAAGCAAGCAGCAAAAAAAAAAAcGC
3ss(-1)_cF	GGCCGCGTTTTTTTTTTTGCTGCTGCTTTTTTTTTTTG
3ss(-1)_cR	TCGACAAAAAAAAAAAGCAGCAGCAAAAAAAAAAAcGC
3ss(0)_cF	GGCCGCGTTTTTTTTGCTTGCTTGCTTTTTTTTG
3ss(0)_cR	TCGACAAAAAAAAGCAAGCAAGCAAAAAAAAcGC
3ss(5)_cF	GGCCGCGTTTTTTTTGCTTTTTTTGCTTTTTTTGCTTTTTTTTG
3ss(5)_cR	TCGACAAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAAcGC
3ss2(5)1_cF	GGCCGCGTTTTTTTTGCTTGCTTTTTTGCTTTTTTTTG
3ss2(5)1_cR	TCGACAAAAAAAAGCAAAAAAGCAAGCAAAAAAAAcGC
4ds_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTGCTTTGCTTTTTTTGTTCTGTCTTCCATGTTCTTCATCTCCGTCCTG
4ds_cR	TCGACAGGACGGAGATGAAGAACATGGAAGACAGAACAAAAAAAGCAAAGCAAGCAGCAGAACAACTCCGTCCTGC
4ds2(15)2_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTTTCTTTTTTTGATCTGCTTGCTCTTCATCTCCGTCCTG
4ds2(15)2_cR	TCGACAGGACGGAGATGAAGAGCAAGCAGATCAAAAAAAGAAAAGCAGCAGAACAACTCCGTCCTGC
4ds2(6)2_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTGACTTTGCTGCTTTTTTCAAGTTCTGTCCTTCCATGTTCTTCATCTCCGTCCTG
4ds2(6)2_cR	TCGACAGGACGGAGATGAAGAACATGGAAGGACAGAACTTGAAAAAAGCAGCAAAGTCAAGCAGCAGAACAACTCCGTCCTGC
4ds2(12)2_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTGACTTCGACTCTGCTGCTTTTTTCATGTTTTGTCCTGTCCTTCCATGTTCTTCATCTCCGTCCTG
4ds2(12)2_cR	TCGACAGGACGGAGATGAAGAACATGGAAGGACAGGACAAAACATGAAAAAAGCAGCAGAGTCGAAGTCAAGCAGCAGAACAACTCCGTCCTGC
4ds2(24)2_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTGACTTCGACTTTCATGTTTTTTTTGCTGCTTGTCCTGTCCTTCCATGTTCTTCATCTCCGTCCTG
4ds2(24)2_cR	TCGACAGGACGGAGATGAAGAACATGGAAGGACAGGACAAGCAGCAAAAAAAACATGAAAGTCGAAGTCAAGCAGCAGAACAACTCCGTCCTGC
4ds1(12)3_cF	GGCCGCAGGACGGAGTTGTTCTGCTTTGACTTCGACTTGCTGCTTGCTTTTTTTGTTTCATGTTTTGTCCTGTCTTGTTCTCATCTCCGTCCTG

4ds1(12)3_cR	TCGACAGGACGGAGATGAGAACAAGACAGGACAAAACATGAAACAAAAAAAGCAAGCAGCAAGT CGAAGTCAAAGCAGAACAACTCCGTCCTGC
4ds1(29)3_cF	GGCCGCAGGACGGAGTTGTTCTGCTTTGACTTCGACTTCATGTTTTGTTTTTTTTGCTTGCTGCTTGT CCTGTCTTGTTCTCATCTCCGTCCTG
4ds1(29)3_cR	TCGACAGGACGGAGATGAGAACAAGACAGGACAAGCAGCAAGCAAAAAAAACAAAACATGAAGT CGAAGTCAAAGCAGAACAACTCCGTCCTGC
4ds2(S)2_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTCCCCATGCTGCTTTTTTCATGTTTTGGGGACCATGTTC TTCATCTCCGTCCTG
4ds2(S)2_cR	TCGACAGGACGGAGATGAAGAACATGGTCCCCAAAACATGAAAAAAGCAGCATGGGGAAGCAGC AGAACAACTCCGTCCTGC
4ds3'S_cF	GGCCGCAGGACGGAGTTGTTCTGCTGCTTGCTTTGCTTCCCCATTTTTTTTGGGGATGTTCTGTCTTC CATGTTCTTCATCTCCGTCCTG
4ds3'S_cR	TCGACAGGACGGAGATGAAGAACATGGAAGACAGAACATCCCCAAAAAAAATGGGGAAGCAAAG CAAGCAGCAGAACAACTCCGTCCTGC
4ss_cF	GGCCGCGTTTTTTTTTTTGCTGCTTGCTTTGCTTTTTTTTTTTG
4ss_cR	TCGACAAAAAAAAAAAAGCAAAGCAAGCAGCAAAAAAAAAAAcGC
4ss(-1)_cF	GGCCGCGTTTTTTTTTTTGCTGCTGCTGCTTTTTTTTTTTG
4ss(-1)_cR	TCGACAAAAAAAAAAAAGCAGCAGCAGCAAAAAAAAAAAcGC
4ss(0)_cF	GGCCGCGTTTTTTTTTTTGCTTGCTTGCTTGCTTTTTTTTTTTG
4ss(0)_cR	TCGACAAAAAAAAAAAAGCAAGCAAGCAAGCAAAAAAAAAAAcGC
4ss(5)_cF	GGCCGCGTTTTTTTTGCTTTTTTTGCTTTTTTTGCTTTTTTTGCTTTTTTTTG
4ss(5)_cR	TCGACAAAAAAAAAGCAAAAAAAAGCAAAAAAAAGCAAAAAAAAGCAAAAAAAAAcGC
4ss2(5)2_cF	GGCCGCGTTTTTTTTTTTTGCTTGCTTTTTTTGCTTGCTTTTTTTTTTTTG
4ss2(5)2_cR	TCGACAAAAAAAAAAAAGCAAGCAAAAAAAAGCAAGCAAAAAAAAAAAAcGC
4ss1(5)3_cF	GGCCGCGTTTTTTTTTGCTTTTTTTTTGCTTGCTTGCTTTTTTTTTG
4ss1(5)3_cR	TCGACAAAAAAAAAGCAAGCAAGCAAAAAAAAAGCAAAAAAAAAcGC

Supplementary Table S2

Sequences of primers for splicing analysis of *Atp2a1* minigene transcript and endogenous transcripts and their expression level

a) Splicing analyses

Primer name	**Sequence 5' → 3'**	**Organism**	**Ta**
HybridMiniG_F1	ACTATCTGGAGGGATAACCACC		60°C
HybridMiniGn_F2	ACTATCTGGAGGATCCAGAAGA		
HybridMiniG_R	AGCAATCAGCTAGTCAGTTGCC		
MBNL1_E5_F	GCTGCCCAATACCAGGTCAAC	*Hs/Mm*	55°C
MBNL1_E5_R	TGGTGGGAGAAATGCTGTATGC	*Hs/Mm*	
MBNL2_E5_F	ATTTTCACCCTCCTGCACAC	*Hs/Mm/Chs*	55°C
MBNL2_E5_R	CAAGACGCTGGGGTTAAAGA	*Hs/Mm/Chs*	
INSR_E13_F	CCAAAGACAGACTCTCAGAT	*Hs*	55°C
INSR_E13_R	AACATCGCCAAGGGACCTGC	*Hs*	
PHKA1_E19_F	TGCACACACTTGAGCTTCATGGA	*Hs/Mm*	55°C
PHKA1_E19_R	AAAGTCCACCTCCCCAGACTGGTC	*Hs/Mm*	
TEAD1_E6_F	GGAGGCCCTGGCTATCTATC	*Hs/Mm/Chs*	55°C
TEAD1_E6_R	GCTTGGAATGAAAATCACGAG	*Hs/Mm/Chs*	
LIMCH_E21_F	CGGAAGCTGCCAGATGTGAAGAAG	*Hs*	55°C

LIMCH_E21_R	CCTCCTCACACCGCATGTCAAA	*Hs*	
LDB3_10_F	AGCCAGAAGATGAGGCTGACGA	*Hs*	55°C
LDB3_10_R	CTGCACAAACTCCTTCAGCGAGTC	*Hs*	
MEF2D_EB_F	TACCCACAGCACCCAGCTT	*Hs*	55°C
MEF2D_EB_R	TAGACTGGGAGACCCAAGG	*Hs*	
NFIX_ E7_F	GAGCCCTGTTGATGACGTGTTCTA	*Hs/Mm*	55°C
NFIX_ E7_R	GAGTCCAGTAGATGATGTGTTCTA	*Hs/Mm*	
Ncor2d_45_F	ACACCCACAACCGGAATGAGCCTG	*Chs*	55°C
Ncor2d_45_R	GGACTTGGCTTTTCGGCTGCTG	*Chs*	
Add3_E14_F	TCCTAACCCATTCAGCCACC	*Mm*	55°C
Add3_E14_R	AGCAAGCTCGTCTTCGACAT	*Mm*	
App_E15_F	CTACGGAAACGACGCTCTCA	*Mm*	55°C
App_E15_R	CCGTCTTGATGTTTGTCAGCC	*Mm*	
Fgfr1_E4_F	CCCCAACCTTGCCTGAACAA	*Mm*	55°C
Fgfr1_ E4_R	GGTTTGGTGTTGTCCGTCTC	*Mm*	
Sh3pxd2a_F	CAGAATGGCACACGAGACGA	*Mm*	55°C
Sh3pxd2a_R	TCTCGTCTTTGCTCTGGCTG	*Mm*	
Slain_E8_F	ATACCTCGAATGCAGCCTCAG	*Mm*	55°C
Slain_ E8_R	GGTGAGGAGGTGGTTTGTGT	*Mm*	
Mef2d_EB_F	GATCTGCGGGTCATCACTTC	*Mm*	55°C
Mef2d_EB_R	GGTGAGCGAGTGGGTAGACT	*Mm*	
Trip10_E10_F	CGTTGGCCTTTCGGGAAAAA	*Mm*	55°C
Trip10_E10_R	GGTCCTCCTCCTTCTGCAAC	*Mm*	
Cacna2d1_E19_F	TGTCCCAATGGCTACTATTTTGC	*Mm*	55°C
Cacna2d1_E19_R	CTCAGCATCGAGAAAATCCAGTG	*Mm*	
Ctage_E11_F	GGAGTCAACAGAAGCACAAAGT	*Mm*	55°C
Ctage_E11_R	TAGCTCGCTCTTCTTCCAGCTT	*Mm*	
Rock2_34_F	CCCGATCATCCCCTAGAACT	*Mm*	55°C
Rock2_34_R	CGATCACCTTCAATAACTGCTTT	*Mm*	

b) RNA expression level:

Primer name	**Sequence 5' → 3'**	**Organism**	**Ta**
MBNL3 _F	GGCCCAAAACTGATGCGTTC	*Hs*	55°C
MBNL3_ R	TCAGTAGGGTGAGCATAGCG	*Hs*	
MBNL1_ F	CTGCCGAACATCTGACTAGC	*Hs*	55°C
MBNL1_ R	TTGCTTGGCACATTCATCTC	*Hs*	
GAPDH_F_240	GAGTCAACGGATTTGGTCGT	*Hs*	55°C
GAPDH_R_240	TTGATTTTGGAGGGATCTCG	*Hs*	

Hs, *Homo sapiens*; *Mm*, *Mus musculus; Chs, Chlorocebus sabaeus*

SUPPLEMENTARY REFERENCES

1. Sznajder, L.J., Michalak, M., Taylor, K., Cywoniuk, P., Kabza, M., Wojtkowiak-Szlachcic, A., Matloka, M., Konieczny, P. and Sobczak, K. (2016) Mechanistic determinants of MBNL activity. *Nucleic Acids Research*, **44**, 10326-10342.
2. Konieczny, P., Stepniak-Konieczna, E., Taylor, K., Sznajder, L.J. and Sobczak, K. (2017) Autoregulation of MBNL1 function by exon 1 exclusion from MBNL1 transcript. *Nucleic Acids Research*, **45**, 1760-1775.
3. Cywoniuk, P., Taylor, K., Sznajder, L.J. and Sobczak, K. (2017) Hybrid splicing minigene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions. *Scientific Reports*, **7**, 14.
4. Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuboock, R. and Hofacker, I.L. (2008) The Vienna RNA Websuite. *Nucleic Acids Research*, **36**, 70-74.

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper ,,**MBNL splicing activity depends on RNA binding site structural context**" Katarzyna Taylor, Łukasz J. Sznajder, Piotr Cywoniuk, James D. Thomas, Maurice S. Swanson & Krzysztof Sobczak in ***Nucleic Acids Research*** 2018, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who analyzed the material from experiment presented in Fig. 1E as well as performed all western blot and eGFP fluorescence detection experiments (Fig. S2c, 6E, S10c, S11a, b).

Ł.J.S. (Łukasz J. Sznajder) performed RT-PCR verification of splicing profile in WT and 3KO myoblasts (Fig. 6A and S10a).

J.D.T. (James D. Thomas) shared the material from 3KO myoblasts and DKO muscles.

K.T. (Katarzyna Taylor) performed all remaining experiments, all figures, wrote, and drafted the manuscript.

M.S.S. (Maurice S. Swanson) Drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 20.09. 2018, Poznań

Name: Katarzyna Taylor

Signature: Taylor

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**MBNL splicing activity depends on RNA binding site structural context**" Katarzyna Taylor, Łukasz J. Sznajder, Piotr Cywoniuk, James D. Thomas, Maurice S. Swanson & Krzysztof Sobczak in ***Nucleic Acids Research*** 2018, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who analyzed the material from experiment presented in Fig. 1E as well as performed all western blot and eGFP fluorescence detection experiments (Fig. S2c, 6E, S10c, S11a, b).

Ł.J.S. (Łukasz J. Sznajder) performed RT-PCR verification of splicing profile in WT and 3KO myoblasts (Fig. 6A and S10a).

J.D.T. (James D. Thomas) shared the material from 3KO myoblasts and DKO muscles.

K.T. (Katarzyna Taylor) performed all remaining experiments, all figures, wrote, and drafted the manuscript.

M.S.S. (Maurice S. Swanson) Drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 27.09.2018, Gainesville, FL

Name: Łukasz J. Sznajder

Signature:

Łukasz Sznajder

CO-AUTHOR STATEMENT

I declare that the work in the paper ,,**MBNL splicing activity depends on RNA binding site structural context**" Katarzyna Taylor, Łukasz J. Sznajder, Piotr Cywoniuk, James D. Thomas, Maurice S. Swanson & Krzysztof Sobczak in ***Nucleic Acids Research*** 2018, of which I am a co-author, is a part of my PhD dissertation (P.C.). I analyzed the material from experiment presented in Fig. 1E as well as performed all western blot and eGFP fluorescence detection experiments (Fig. S2c, 6E, S10c, S11a, b).

Ł.J.S. (Łukasz J. Sznajder) performed RT-PCR verification of splicing profile in WT and 3KO myoblasts (Fig. 6A and S10a).

J.D.T. (James D. Thomas) shared the material from 3KO myoblasts and DKO muscles.

K.T. (Katarzyna Taylor) performed all remaining experiments, all figures, wrote, and drafted the manuscript.

M.S.S. (Maurice S. Swanson) Drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 15.10.2018, Poznań

Name: Piotr Cywoniuk

Signature: Piotr Cywoniuk

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper ,,**MBNL splicing activity depends on RNA binding site structural context''** Katarzyna Taylor, Łukasz J. Sznajder, Piotr Cywoniuk, James D. Thomas, Maurice S. Swanson & Krzysztof Sobczak in ***Nucleic Acids Research*** 2018, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who analyzed the material from experiment presented in Fig. 1E as well as performed all western blot and eGFP fluorescence detection experiments (Fig. S2c, 6E, S10c, S11a, b).

Ł.J.S. (Łukasz J. Sznajder) performed RT-PCR verification of splicing profile in WT and 3KO myoblasts (Fig. 6A and S10a).

J.D.T. (James D. Thomas) shared the material from 3KO myoblasts and DKO muscles.

K.T. (Katarzyna Taylor) performed all remaining experiments, all figures, wrote, and drafted the manuscript.

M.S.S. (Maurice S. Swanson) Drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript**.**

Date: 27.09.2018, Seattle, WA

Name: James D. Thomas

Signature: 

CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**MBNL splicing activity depends on RNA binding site structural context**" Katarzyna Taylor, Łukasz J. Sznajder, Piotr Cywoniuk, James D. Thomas, Maurice S. Swanson & Krzysztof Sobczak in ***Nucleic Acids Research*** 2018, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who analyzed the material from experiment presented in Fig. 1E as well as performed all western blot and eGFP fluorescence detection experiments (Fig. S2c, 6E, S10c, S11a, b).

Ł.J.S. (Łukasz J. Sznajder) performed RT-PCR verification of splicing profile in WT and 3KO myoblasts (Fig. 6A and S10a).

J.D.T. (James D. Thomas) shared the material from 3KO myoblasts and DKO muscles.

K.T. (Katarzyna Taylor) performed all remaining experiments, all figures, wrote, and drafted the manuscript.

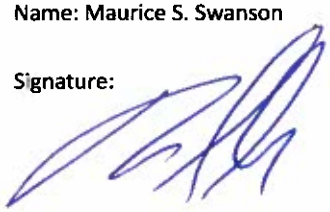
M.S.S. (Maurice S. Swanson) Drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 27.09.2018, Gainesville, FL

Name: Maurice S. Swanson

Signature:



CO-AUTHOR STATEMENT

I declare that I am aware that the work in the paper „**MBNL splicing activity depends on RNA binding site structural context**" Katarzyna Taylor, Łukasz J. Sznajder, Piotr Cywoniuk, James D. Thomas, Maurice S. Swanson & Krzysztof Sobczak in ***Nucleic Acids Research*** 2018, of which I am a co-author, is a part of the PhD dissertation by Piotr Cywoniuk (P.C.) who analyzed the material from experiment presented in Fig. 1E as well as performed all western blot and eGFP fluorescence detection experiments (Fig. S2c, 6E, S10c, S11a, b).

Ł.J.S. (Łukasz J. Sznajder) performed RT-PCR verification of splicing profile in WT and 3KO myoblasts (Fig. 6A and S10a).

J.D.T. (James D. Thomas) shared the material from 3KO myoblasts and DKO muscles.

K.T. (Katarzyna Taylor) performed all remaining experiments, all figures, wrote, and drafted the manuscript.

M.S.S. (Maurice S. Swanson) Drafted and revised the manuscript.

K.S. (Krzysztof Sobczak) designed the study, supervised the experiments, drafted and revised the manuscript.

Date: 20.09. 2018, Poznań

Name: Krzysztof Sobczak

Signature: