

Direct Evidence that EXP/muscleblind Interacts with CCUG Tetranucleotide Repeats

Yoshihiro Kino, Yoko Oma, Yuya Takeshita, Daisuke Mori, Noboru Sasagawa and Shoichi Ishiura

Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

Abstract

Myotonic Dystrophy (DM), the most common form of adult-onset muscular dystrophy, comprises at least 2 sub-types, DM1 and DM2. DM1 is caused by the expansion of a CTG repeat located in the 3' untranslated region (3'UTR) of the DM protein kinase (*DMPK*) gene. Recently, the expansion of a CCTG tetranucleotide repeat located in the intron of the *ZNF9* gene was identified as the mutation responsible for DM2. Since both DM1 and DM2 are caused by the expansion of repetitive sequences, some common factors that interact with these sequences might be involved in the pathogenesis of DM. EXP/muscleblind is a candidate for such factors and is thought to be sequestered by the expanded forms of DM transcripts. Here we performed yeast three-hybrid assays that can detect protein-RNA interactions *in vivo* to confirm the binding feature of EXP to several kinds of repetitive RNA sequences. We found that both EXP40 and EXP42, splice variants of EXP, showed significant interactions with both CCUG and CUG repeats, but not with UG repeats. These results strongly suggest that EXP might be involved in the pathogenesis of both DM1 and DM2. We found the difference in the intracellular localization of EXP40 and EXP42, though their amino acid sequences differ slightly.

Key words: myotonic dystrophy, RNA-binding protein, EXP, CCUG repeat, yeast three-hybrid system.

Basic Appl Myol 13(6) 293-298, 2003

Myotonic dystrophy (DM) is an autosomal dominant inherited disease. DM consists of at least two sub-types, DM1 and DM2, which share many symptoms such as myotonia (muscle hyperexcitability), muscle weakness and cataracts [14]. DM1 is caused by an expansion of a CTG repeat located in the 3' untranslated region (3'UTR) of the DM protein kinase (*DMPK*) gene [3,6,10]. Recently, the mutation responsible for DM2 was identified in the intron of the *ZNF9* gene. Interestingly, it is the expansion of a novel type of repeat, a CCTG tetranucleotide repeat [9].

One major question is how a mutation not located in the open reading frame of a gene leads to dominant inheritance, and several models have been proposed to explain it. An RNA gain-of-function model is one such model, in which expanded CUG/CCUG repeats in the mRNA or pre-mRNA transcribed from a mutated allele sequester some proteins or RNAs, which interact with these repeats and prevent their normal function. Indeed, Taneja and several others have reported the observation of CUG repeat RNA foci stained by CAG repeat probes

localized in the nucleus in FISH (fluorescent *in situ* hybridization) analysis of DM cells [24], suggesting that mRNAs with expanded repeats form aggregative complexes that cannot be exported to the cytoplasm. This model can account for the dominant inheritance of DM, and the retention of RNA in the nucleus might also cause a reduction in its translation, which might result in the haplo-insufficiency of the *DMPK* or *ZNF9* protein.

Several genes have been proposed to be involved in the pathogenesis of DM1. One of such gene, CUG-BP1 (CUG repeat-binding protein 1), was found in a gel-shift assay using a CUG₈ probe *in vitro* [25]. CUG-BP1 has been reported to be involved in the regulation of alternative splicing and deadenylation [17,20,27]. There is evidence to indicate the involvement of CUG-BP1 in the pathogenesis of DM1, including a disorder in muscle differentiation [17,27], insulin resistance [20] and myotonia [4,12] through irregular splicing. The expression level and functional activity of CUG-BP1 are altered by an expanded CUG repeat [4,20,26]. There are other reports, however, that CUG-BP1 does not interact

EXP Interacts with CCUG Repeats

with CUG repeats in a length-dependent manner [15], and it didn't co-localize with CUG repeat RNA foci in the FISH analysis of DM1 cells [5]. Previously we reported that CUG-BP1 binds to UG dinucleotide repeats more specifically than CUG repeats in the yeast three-hybrid system [23]. If this is the case, CUG-BP1 is not a suitable candidate for the sequestration model, even though it might be involved in the symptoms of DM in the downstream of the regulatory pathway induced by the expanded repeats.

EXP/MBNL, the human ortholog of *muscleblind* in *Drosophila*, is another potential CUG-repeat interacting protein that was identified as a protein interacting with double-stranded CUG repeats in a length-dependent manner [16]. Moreover, EXP has been reported to co-localize not only with the CUG repeat in DM1 cells [5], but also with the foci stained by the CAG probe in the cells of DM2 patients [13], suggesting that this protein interacts with the expanded CCUG repeat. Although the function of EXP in the human body is not understood, *muscleblind* is known to be involved in the terminal differentiation of photoreceptor cells in the eye [2] and muscles [1]. Because eyes and muscles are affected in DM patients, it is possible that EXP accounts for some of the features seen in DM.

We tried to examine the binding ability of EXP to repetitive RNAs using a yeast three-hybrid system. This system has an advantage in detecting protein-RNA interactions *in vivo* [21]. Although co-localization of EXP and repetitive RNAs has been observed by FISH analysis, this system would provide a more controlled condition in which length and species of RNA can be

manipulated. Here we describe that EXP shows significant interactions with both CCUG and CUG repeats, providing evidences which suggest that EXP might be involved in the pathogenesis of DM. And, we found the difference in the intracellular localization of EXP splice variants.

Material and Methods

Vector construction

EXP was amplified from a human muscle cDNA library by PCR (polymerase chain reaction) with the primers 5'-CTAAACATGGCTGTTAGTGTAC-3' as a sense primer and 5'-CTACATCTGGGTAACAT-CTTGTGG-3' as an antisense primer. The PCR fragment was ligated into the *HincII* site of pUC118 (TAKARA, Tokyo, Japan) and sequenced in the usual procedure. We confirmed several kinds of EXP splice variants, including EXP40 and EXP42. EXP40 and EXP42 were amplified by PCR with the forward primer 5'-CCCGGGATCCATGGCTGTTAGTGTACACC-3' and the reverse primer 5'-GCAGGCTGACCATCTGGGTAACATAC-3', and then digested with *SalI*. The fragments were ligated into the *SmaI* and *SalI* sites of pGAD424 (CLONTECH, Tokyo, Japan), and the sequences were confirmed again.

Repetitive sequences were synthesized as previously described [22]. In brief, pairs of TG₁₂/CA₁₂, CTC₇/CAG₇ and CCTG₇/CAGG₇ oligo-DNAs were synthesized (Prologo, Kyoto, Japan). Each pair was amplified by PCR in the mixture and under the conditions as described [22]. The PCR fragments were ligated into the *HincII* site of pUC118 and sequenced. Then the vectors were digested with *EcoRI* and *PstI* and the repeat-containing fragments were blunt-ended with T₄ DNA polymerase (TAKARA). The repeat fragments were ligated into the *SmaI* site of

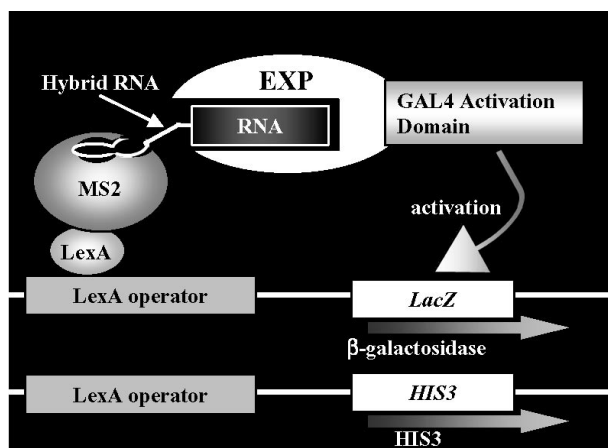


Fig. 1. Scheme of Yeast three-hybrid system

Repetitive RNAs are bound to LexA-fused MS2 coat protein and located near the LexA operator. EXP is expressed in yeast as a fused protein with GAL4 activation domain. When EXP interacts with the repetitive RNA, GAL4 activation domain activates the reporter genes downstream of LexA operator. In this study, we used β -galactosidase and HIS3 as reporter genes.

	MS2-2	IRE	UG 24	CUG 7	CUG 21	CUG 37	CCUG 7	CCUG 50
pGAD	-	-	-	n.d.	n.d.	n.d.	n.d.	-
IRP	-	+++	-	n.d.	n.d.	n.d.	n.d.	-
EXP40	-	n.d.	-	+	+++	+++	+	+++
EXP42	-	n.d.	-	+	+++	+++	+	+++

... survived
 + survived but grew very slowly
 - not survived
 n.d. not determined

Table. 1. Histidine selective assay

Colony viability is classified into 3 categories, (+++, survived), (+, survived but grew very slowly), and (-, not survived). Hybrid RNAs and proteins used for making combinational transformants are denoted to the left and top of the panels, respectively. MS2-2, empty vector for hybrid RNA expression; pGAD, empty vector for hybrid protein expression. The combination of IRE (iron responsive element) and IRP (iron regulatory protein) was used as a positive control.

EXP Interacts with CCUG Repeats

pIII/MS2-2 and sequenced. Positive control plasmids, pIII/IRE-MS2 and pAD-IRP [21], were gifts from Dr. Marvin Wickens (University of Wisconsin-Madison, US).

Transformation of yeast was performed using the general lithium-acetate method. Combinations of the constructs for RNA expression and those for protein expression described above were co-transformed and selected on leucine and uracil selective plates. Five colonies of each combination were picked up and spotted onto a histidine selective plate, which lacks leucine, uracil and histidine. The plates were incubated for about 10 days and the viability of each colony was observed and classified into one of three categories (see Results). The β -galactosidase assay was performed as described [21,23]. The results were statistically analyzed by a two-tailed Student's t-test.

Intracellular localization

EXP40 and EXP42 were cut off from pGAD vectors used above by *Bam*HI and *Sal*I and ligated into the *Bg*III and *Sal*I site of pEGFP-C1 (CLONTECH). Resulting plasmids were transfected into COS-7 cells with FuGENE6 (Roche Diagnostics, Tokyo, Japan) as described by the manufacturer. Fluorescence was visualized 24 hour posttransfection by conventional fluorescent microscopy (OLYMPUS, Tokyo, Japan). Hoechst 33342 was used for the staining of nucleus.

Results

We cloned several splice variants of EXP from a human skeletal muscle cDNA library. Among them, EXP40 and EXP42 were used in the following study. To detect interactions between protein and RNA, we employed yeast three-hybrid assays in which if an objective RNA (bait) binds to a protein (prey), the expressions of two reporter genes, *HIS3* and *LacZ*, are activated by the Gal4 activation domain fused to the prey (Fig. 1). The activity of *HIS3* can be detected as the viability of yeast transformants on histidine selective plates. The expression of *LacZ* can be quantified by the enzymatic activity of β -galactosidase in the lysate of

transformants.

We synthesized several repetitive sequences as RNA bait and assayed with EXP40 or EXP42 as prey. Table 1 shows the result of the histidine selective assay. In order to assess the binding ability, we classified the degree of growth into one of three categories: (+++) survived (positive), (+) survived but grew very slowly (survival could be seen after more than 1 week), and (-) not survived (negative). The colonies in category (+) could represent a very faint interaction between RNA and protein, but it could also be false positives, because even the combination of the negative control sometimes showed a false positive for unknown reasons. The combination of IRE (iron responsive element) and IRP (iron regulatory protein), which are classically known to bind, were used as a positive control to confirm that the system normally worked.

On the histidine selective plate, both EXP40 and EXP42 showed some viability in combination with CCUG and CUG repeats, but not with UG repeat (Table 1). The results of the β -galactosidase assay also suggested that EXP40 interacts significantly with CCUG and CUG repeats (Fig. 3) in consistent with the results of histidine selective assay. In both experiments, EXP showed slight or almost no interaction with shorter CUG/CCUG repeats, suggesting that the binding of EXP requires longer repeats. The binding feature of EXP42 was the same as EXP40 (Table 1).

We next examined the intracellular localization of EXP40 and EXP42. They were expressed as a fusion protein with EGFP (enhanced green fluorescent protein) in COS-7 cells. Interestingly, EXP42 localized exclusively in the nucleus, while EXP40 distributed both

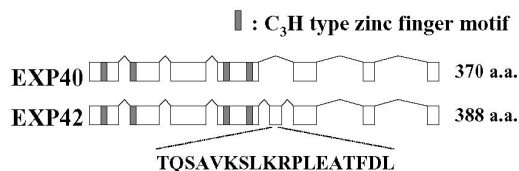


Fig. 2 Comparison of EXP40 and EXP42

Each open box corresponds to an exon. EXP42 differs from EXP40 by an insertion of an exon which encodes 12 amino acids as indicated. Both of them have four C₃H type zinc finger motifs as indicated by gray boxes.

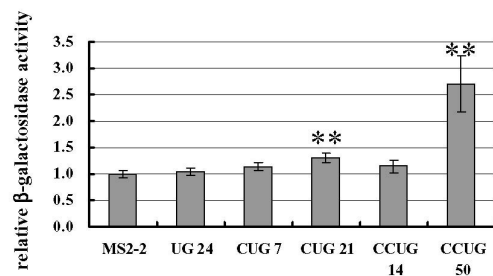


Fig. 3. Interaction of RNA repeats with EXP40 (β -galactosidase assay)

Relative β -galactosidase activities of EXP40 with repetitive RNA sequences are shown in this figure. The β -galactosidase activity of each transformant is shown as a relative value compared to that of the combination with MS2-2 (normalized to 1.0). The RNA sequences used are indicated under the respective bars. Error bars denote the standard deviations ($n=4$). **: $P<0.005$ relative to the combination with MS2-2 by Student's t-test.

EXP Interacts with CCUG Repeats

cytoplasm and nucleus as shown in Fig. 4. These results were confirmed by the immunostaining of myc-tagged EXP40/42 (data not shown).

Discussion

The identification of a CCTG repeat on chromosome 3q21 as a causative factor of DM2 had a considerable impact because several pathogenic models of DM relied on the gene context of 19q13.3 including *DMPK* and assumed the haploinsufficiency of *DMPK* and its neighboring genes. From results in *DMPK*-deficient mice, the reduction in the DMPK protein might be involved in some cardiac conduction defects of DM, but it could not explain all the symptoms variously seen in DM [7,18]. *Six5*, the gene downstream of *DMPK*, has been suggested to be involved in cataract formation, a well-known symptom of DM [8]. But cataracts have been reported in DM2 patients [14], in which the expression of *Six5* might not be altered. Although the involvement of these genes in the pathogenesis of DM cannot be excluded, the importance of the RNA gain-of-function model has become greater because the expression of aberrant RNA could be common in both DM1 and DM2 and the trans-action of RNA can reasonably explain the dominant heredity of DM. Indeed, transgenic mice harboring a 250 CTG repeat in the 3'UTR of the human skeletal actin gene develop myotonia and myopathy [19], in which the expression of *DMPK* and its neighboring gene might not be directly affected, suggesting the expression of an expanded CUG repeat is sufficient to cause some of the symptoms of DM. The simplest effect of the RNA-gain-of-function can be thought of as the sequestration of interactive factors with the expanded RNA repeat. These factors might lose their normal localization and function, leading to the pathogenesis. In this context, the identification and

characterization of RNA repeat-binding proteins seems to be important.

We tried to examine EXP whether it interacts with CCUG and CUG repeats in a yeast three-hybrid system. For this purpose, several repetitive sequences were synthesized as RNA baits and assayed with EXP as prey. The HIS3 assay is suitable for processing many samples at the same time, but it is not quantitative. On the other hand, the LacZ assay allows quantitative analysis.

We showed that EXP40 and EXP42 bind to both CCUG and CUG repeats. This result strongly suggests that several splice variants of EXP may be involved in the pathogenesis of DM. Miller et al. reported EXP as a binding factor of a double-stranded CUG repeat [16]. EXP has also been suggested to localize in the nuclear foci stained by the CAG probe in DM1 and DM2 cells [13]. On the other hand, EXP did not interact with UG (Fig. 3), CA or CAGG repeats (data not shown). Predictions of the secondary structures of RNA sequences (MFOLD v3.1 software, <http://bioinfo.math.rpi.edu/~mfold/>) reveals that not only CUG repeats but also CCUG repeats, form double-stranded structures, while UG and CA repeats do not seem to have such structures. The hairpin might become more stable as the repeat expands. The involvement of EXP in the symptoms of DM is still unclear, but this protein is most suitable for the sequestration model. The identification of its real target and function might explain the pathogenesis of DM.

EXP40 and EXP42 have four zinc finger motifs in common, and differ from each other by 18 amino acids in the downstream of the 4th zinc finger motif. Our results suggest that this insertion might not alter the binding ability of EXP, but alter the intracellular localization. There is no obvious nuclear localization signal (NLS) in the 18 amino acid insertion, but two lysines and an arginine residues, which are characteristic residues found in NLSs, exist in this region (Fig. 2). Although physiological significance of this difference remains to be explained, this finding is interesting because nuclear foci of expanded repeats are hallmark of DM. Some of EXP splice variants are possibly affected by the expanded repeat while others are not.

Acknowledgements

We would like to thank Dr. Marvin Wickens of the University of Wisconsin at Madison for kindly providing us with the yeast three-hybrid system. This work is supported by research grants from the Ministry of Health, Labor and Welfare, Japan, and "Ground-based Research Announcement for Space Utilization" promoted by Japan Space Forum.

Address correspondence to:

Shoichi Ishiura, Department of Life Sciences, Graduate School of Arts and Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan, fax +81

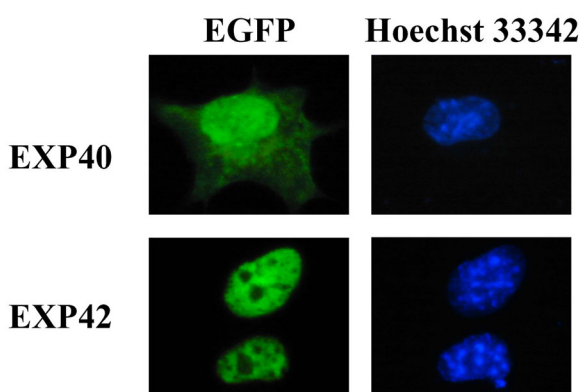


Fig. 4. Intracellular localization of EXP
EGFP-fused EXP40 and EXP42 were expressed in COS-7 cells. The fluorescence of EGFP was observed by fluorescent microscopy. Nucleus were stained by Hoechst 33342. EXP42 was localized exclusively in the nucleus, while EXP40 was distributed in the cytoplasm as well as in the nucleus. 400x.

EXP Interacts with CCUG Repeats

3-5454-6739, Tel +81 3-5454-6739, Email cishiura@mail.ecc.u-tokyo.ac.jp.

References

- [1] Artero R, Prokop A, Paricio N, Begemann G, Pueyo I, Mlodzik M, Perez-Alonso M, Baylies MK: The *muscleblind* gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by *Dmef2*. *Dev Biol* 1998; 195: 131-43.
- [2] Begemann G, Paricio N, Artero R, Kiss I, Perez-Alonso M, Mlodzik M: *muscleblind*, a gene required for photoreceptor differentiation in *Drosophila*, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development* 1997; 124: 4321-31.
- [3] Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Jones C, Juvonen V, Johnson K, Harper PS, Shaw DJ, and Housman DE: 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* 1992; 68: 799-808.
- [4] Charlet-B N, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA: Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol Cell* 2002; 10: 45-53.
- [5] Fardaei M, Larkin K, Brook JD, Hamshire MG: *In vivo* co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res* 2001; 29: 2766-71.
- [6] Fu YH, Friedman DL, Richards S, Pearlman JA, Gibbs RA, Pizzuti A, Ashizawa T, Perryman MB, Scarlato G, Fenwick RG Jr., Caskey CT: Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 1993; 260: 235-8.
- [7] Jansen G, Groenen PJ, Bachner D, Jap PH, Coerwinkel M, Oerlemans F, van den Broek W, Gohlsch B, Pette D, Plomp JJ, Molenaar PC, Nederhoff MG, van Echteld CJ, Dekker M, Berns A, Hameister H, Wieringa B: Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat Genet* 1996; 13: 316-24.
- [8] Klesert TR, Cho DH, Clark JI, Maylie J, Adelman J, Snider L, Yuen EC, Soriano P, Tapscott SJ: Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. *Nat Genet* 2000; 25: 105-9.
- [9] Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP: Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 2001; 293: 864-7.
- [10] Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barcelo J, O'Hoy K, Leblond S, Earle-Macdonald J, de Jong PJ, Wieringa B, Korneluk RG: Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 1992; 255: 1253-5.
- [11] Mankodi A, Logigian E, Callahan L, McClain C, White R, Henderson D, Krym M, Thornton CA: Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 2000; 289: 1769-73.
- [12] Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA: Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell* 2002; 10: 35-44.
- [13] Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, Henderson D, Schalling M, Swanson MS, Thornton CA: Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum Mol Genet*. 2001; 10: 2165-70.
- [14] Meola G: Clinical and genetic heterogeneity in myotonic dystrophies. *Muscle Nerve*. 2000; 23: 1789-99.
- [15] Michalowski S, Miller JW, Urbinati CR, Paliouras M, Swanson MS, Griffith J: Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res* 1999; 27: 3534-42.
- [16] Miller JW, Urbinati CR, Teng-Ummuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS: Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* 2000; 19: 4439-48.
- [17] Philips AV, Timchenko LT, Cooper TA: Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* 1998; 280: 737-41.
- [18] Reddy S, Smith DB, Rich MM, Leferovich JM, Reilly P, Davis BM, Tran K, Rayburn H, Bronson R, Cros D, Balice-Gordon RJ, Housman D: Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet* 1996; 13: 325-35.
- [19] Sarkar PS, Appukuttan B, Han J, Ito Y, Ai C, Tsai W, Chai Y, Stout JT, Reddy S: Heterozygous loss of *Six5* in mice is sufficient to cause ocular cataracts. *Nat Genet* 2000; 25: 110-4.
- [20] Savkur RS, Philips AV, Cooper TA: Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet*. 2001; 29: 40-7.

EXP Interacts with CCUG Repeats

- [21] SenGupta DJ, Zhang B, Kraemer B, Pochart P, Fields S, Wickens M: A three-hybrid system to detect RNA-protein interactions *in vivo*. *Proc Natl Acad Sci U S A* 1996; 93: 8496-501.
- [22] Takahashi N, Sasagawa N, Suzuki K, Ishiura S: Synthesis of long trinucleotide repeats in vitro. *Neurosci Lett* 1999; 262: 45-8.
- [23] Takahashi N, Sasagawa N, Suzuki K, Ishiura S: The CUG-binding protein binds specifically to UG dinucleotide repeats in a yeast three-hybrid system. *Biochem Biophys Res Commun* 2000; 277: 518-23.
- [24] Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH: Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* 1995; 128: 995-1002.
- [25] Timchenko LT, Miller JW, Timchenko NA, DeVore DR, Datar KV, Lin L, Roberts R, Caskey CT, Swanson MS: Protein Identification of a (CUG)_n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res* 1996; 24: 4407-14.
- [26] Timchenko NA, Cai ZJ, Welm AL, Reddy S, Ashizawa T, Timchenko LT: RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J Biol Chem* 2001; 276: 7820-6.
- [27] Timchenko NA, Iakova P, Cai ZJ, Smith JR, Timchenko LT: Molecular basis for impaired muscle differentiation in myotonic dystrophy. *Mol Cell Biol* 2001; 21: 6927-38. Meola G: Clinical and genetic heterogeneity in myotonic dystrophies. *Muscle Nerve*. 2000; 23: 1789-99.