DEPARTAMENTO DE GENÉTICA

ANÁLISIS DE LA FUNCIÓN MOLECULAR DE LAS PROTEÍNAS MUSCLEBLIND DE DROSOPHILA

MARTA VICENTE CRESPO

UNIVERSITAT DE VALENCIA
Servei de Publicacions
2008

Aquesta Tesi Doctoral va ser presentada a Valencia el dia 9 de Novembre de 2007 davant un tribunal format per:

- Dª. Margarita Cervera Jover
- Dª. Mª Isabel Palacios de Castro
- D. Ibo Galindo Orozco
- D. Darren Monckton
- Dª. Nuria Paricio Ortiz

Va ser dirigida per:
D. Rubén Artero Allepúz

Depòsit legal:
I.S.B.N.:978-84-370-7063-6
Edita: Universitat de València
Servei de Publicacions
C/ Artes Gráficas, 13 bajo
46010 València
Spain
Telèfon: 963864115

VNIVERSITAT DE VALÈNCIA

Análisis de la función molecular de las proteínas Muscleblind de *Drosophila*.

Analysis of *Drosophila* Muscleblind molecular function.

MEMORIA PRESENTADA POR Dña. MARTA VICENTE CRESPO
PARA OPTAR AL TÍTULO DE DOCTORA EN CIENCIAS BIOLÓGICAS
POR LA UNIVERSIDAD DE VALENCIA

VALENCIA, 12 DE JULIO DE 2007

Dr. D. RUBÉN D. ARTERO ALLEPUZ, Profesor Contratado Doctor del Departamento de Genética de la Facultad de Ciencias Biológicas de la Universidad de Valencia

CERTIFICA: que Dña. MARTA VICENTE CRESPO, licenciada en Ciencias Biológicas, ha realizado bajo mi dirección el trabajo de investigación recogido en esta memoria bajo el título "Análisis de la función molecular de las proteínas Muscleblind de *Drosophila*", para optar al grado de Doctora en Ciencias Biológicas y

AUTORIZA: su presentación en el Departamento de Genética de la facultad de Ciencias Biológicas para tal fin.
Para que conste, y en cumplimiento de la legislación vigente, firma el presente certificado en Burjassot a 12 de julio de 2007.

Fdo. Dr. D. Rubén D. Artero Allepuz

A mis padres

y a mis abuelos

La única lucha que se pierde es la que se abandona

-autor desconocido-

"Leonardo wrote that a painter should begin every canvas with a wash of black, because all things in nature are dark except where exposed by the light.

[Paul] understands the value of starting with the shadows. The only things people can ever know about you are the ones you let them see."

"Leonardo escribió que un pintor debería comenzar cada lienzo con una capa de negro, porque todo en la naturaleza es oscuro excepto allí donde está expuesto a la luz.

[Paul] entiende el valor de empezar con las sombras. Las únicas cosas que la gente puede saber de ti son aquéllas que tú les dejas ver."

I. Cadwell & D. Thomason
The rule of four

AGRADECIMIENTOS

Por si no vuelvo a escribir un libro, intentaré no dejarme a nadie aquí. Gracias…

A Nuria y Manuel, por acogerme como colaboradora cuando no era más que una pipiola. A Rubén, por presentarme un gran proyecto, porque sin ti no habría tesis que escribir. A Maury y a Chris, por dejarme ir de okupa a sus laboratorios y hacer posible gran parte de esta tesis; a Lei, por acogernos a mí y a mis moscas. A Clare, por enseñarme en tres meses lo que no aprendí en tres años. A Margarita Cervera y Jose, a Carlos García y Ana, y a Joaquim Culí, por su ayuda con la tnt, el tres híbridos y las células S2. A Isabel, por mostrarme un camino hacia delante, mi carril bici sobre el puente.

A mis compañeros de laboratorio. A Lidón, por hacer de nuestro equipo una extraña familia, por llevarme a dar una vuelta por el jardín y ser mi guía en el depósito; a Zaida, Maya y Silvia, por abrir camino en un laboratorio que apenas empezaba; a Vero, por estar ahí desde el principio y aguantar hasta el final; a Use, por ese laboratorio que nos creamos en el espacio que nos dejaron; a Amparo, por ser un encanto de mujer que se merece todo en esta vida; a Fabrice, por recomendarme que fuera a Churchill; a Juanma, por transmitir tanta alegría y sacar partido a mis construcciones, y por hacer posible que mi tesis llegara "a tiempo"; a Yaiza, por hacerme un hueco en su casa; a los que llegaron después, Irma, MªCarmen, Ana los colaboradores…¡buena suerte! Al resto de sufridores habitantes del departamento de genética, sobretodo a los moleculares que por cercanía, por número o por bulliciosos me hicisteis mucha compañía. A los que desde secretaría o cocina nos hacéis la vida más fácil a los demás, porque a todos os debo algo. A la gente del IVIA, por mi primer verano entre pipetas en tan buena compañía. A mis nuevos compañeros, por darme una nueva "segunda casa", en especial a Kat, por su ayuda con el inglés de la tesis.

A mis compañeros precarios. A todo Joves, por los logros con Consellería y con el CIPF; a mis queridas hormiguitas de la comisión de documentación, por trabajar en la sombra; a la Federación al completo, por seguir en la lucha.

A mi familia. Al núcleo duro, por soportarme tanto tiempo. A mis padres, por ser los mejores padres del mundo, porque para dejarme ser como soy hay que tener mucho aguante. A mi hermano, por seguir siendo él mismo; a Laura, por hacer que yo no sea la única que trae de cabeza a mis padres; a mi cuñada, por ser una más en la familia, por traer a mi sobri al mundo. A los no tan cercanos, que por cosas de la vida vivís a muchos kilómetros pero que os mantenéis en mi mundo, en especial a Ana y Miguel, por ser amigos

además de primos; y a Maite, Ana y mi tia Rosalía, porque les encantan mis historias.

A Ruth, por sacarme de mí misma, por mostrarme otra cara de la vida, por venir aquella tarde. A Almu, Laura, Nieves, Lidón, Dani, a tanta gente, por salvarme de mí misma, por insistir, por acudir, por todo.

A mis compis de carrera que siguen en mi vida después de tanto tiempo, San, Ni, Ana, Ale, Raquel. Por todos los apuntes compartidos, todos los exámenes sufridos, por Monfragüe, los espaguetis carbonara y por disfrutar de mi pueblo como si fuera vuestro, ¡CODOS EXISTE! A Almu, por tantas horas en la resi y fuera de ella, por no preguntar, por cuidarme como una madre, por la noche antes de Cambridge. A Laura, por volver a mi vida después de tantos años, por abrirme los ojos. A Asun, por ahorrarme un dineral en psicólogos, por adivinarme y hacerme hablar. A Virgilio, por tantas noches de ciber-consejo cuando estaba lejos de casa, por ser mi hotel en Londres.

A mis amigos de siempre. A Andrés, por seguir siendo raro, mantenerse fiel a nuestra cita en nochevieja y poner banda sonora a mi viaje a Italia; a Maxi, por hacerme reír y bailar conmigo; a Raquel, Vero, Marta, Mireia, Lucía, Neus y Ramón, por aquel junior de oro, sin duda, de lo mejor que me ha pasado en la vida; a Jorge, por ser un pedazo de pan; a Toni, por hacerse adicto a nuestras nocheviejas.

A mi familia en Cambridge. A Rafa, por darme un techo cuando no tenía dónde caer en Cambridge y acogerme entre sus amigos sin conocerme. A Silvia, por sacarme de casa cuando no tenía ganas y leerse mi tesis sin tener por qué. A Roxani, por mirarme a los ojos y acompañarme a ver a Peter Pan. A James, por invitarme a una fiesta en su casa apenas unas horas después de conocerme. A la otra Silvia, por ser tal como es y transmitir exactamente eso. A Miriam, por hacer que no tuviera que salir del lab para encontrar una amiga. A mis chicas de Firenze. A Silvia, por encontrarme casa y acompañarme en tierras lejanas. A Valentina, por adoptarme más que hospedarme. A todas las demás, por hacerme un hueco en Careggi. A mis compañeros en Florida. A Yuan, Jihae, Pam y, cómo no, a CP, por acompañarme mientras descubría que en Florida llueve todos los días en verano.

A todos los que me encontré dando vueltas por el mundo, por llenar mi memoria de momentos inolvidables y mis cajones de fotos increíbles.

GRACIAS A TODOS (y perdón por todos los ratos que os he hecho sufrir)

hoy es siempre todavía,
toda la vida es ahora,
y ahora,
ahora es el momento de cumplir las promesas que nos hicimos
porque ayer no lo hicimos,
porque mañana es tarde,
AHORA

Antonio Machado

Según lo exigido en la convocatoria, se hace constar que la autora de este trabajo "disfrutó" de una beca predoctoral del programa FPU/MEC, beca que limitó sus derechos sociales y le negó los derechos laborales de los que disfrutan los investigadores contratados.

Por la dignidad en la investigación, Federación de jóvenes investigadores-Precarios, **FJI**. www.precarios.org

Índice

Ojalá me hubiera dado cuenta antes;
no siempre lo urgente es lo importante
-Fito-

Abbreviations and gene symbols

3-AT = 3-aminotriazol
A = absorbance
α–actn = α-actinin
BDGP = Berkeley Drosophila Genome Project
BSA = bovine serum albumin
CUGBP = CUG binding protein
DMPK = Dystrophia Myotonica Protein Kinase
ETR-3 = Embryonically lethal abnormal vision Type RNA binding protein 3
IPTG = Isopropyl β-D-1-thiogalactopyranoside
GFP = Green Fluorescent Protein
mbl = muscleblind
Mhc = Myosin Heavy Chain
NSDB = 3-[Benzyldimethyl-ammonio]propanesulfonate]
o/n = over night
ORF = open reading frame
PCR = Polymerase Chain Reaction
PMSF = phenylmethanesulphonylfluoride
PTB = Polypirimidine Tract Binding protein
RT = Reverse Transcription
SDS-PAGE = sodium dodecyl sulfate polyacrylamide gel electrophoresis
ubi = ubiquitous
UTR = untranslated region
w = white
y = yellow

Introduction

Ícaro,
no hay escapatoria real
cuando el cepo es la mente.
-Raúl Heraud-

Human Muscleblind proteins have emerged as a new family of alternative splicing factors, implicated in many developmental processes and diseases. *Drosophila muscleblind* shows good signs of being a useful model to study Muscleblind protein function in development and the diseases in which the human homologues are implicated.

1. **Regulation and developmental consequences of splicing.**

Pre-mRNA splicing involves the removal of introns and ligation of the flanking exons. Through alternative splicing, the exonic sequences of a pre-mRNA are combined generating different transcripts from a single precursor. Alternative splicing can generate more transcripts from a single gene than the number of genes in an entire genome. A dramatic example is *Drosophila Down Syndrome Cell Adhesion Molecule* (*Dscam*), which potentially encodes 38016 proteins [1]. 40-60% of human genes are thought to undergo alternative splicing [2].

The final exon composition of a transcript is determined by both cis sequence elements and factors in trans assembled in a complex called spliceosome. The excision of the introns from a pre-mRNA and the joining of the exons are directed by special sequences at the intron/exon junctions called splice sites [3]. Components of the basal splicing machinery bind to the classical splice-site sequences and promote assembly of the spliceosome, which performs both the recognition of the intron/exon boundaries and the catalysis of the cut-and-paste reactions that remove introns and join exons. Changes in the assembly of the spliceosome generate changes in splice site choice. The splice site consensus sequences generally do not carry enough information to determine whether

a site will assemble a spliceosome and function in splicing. Additional information and interactions come from many non–splice site regulatory sequences that strongly affect spliceosome assembly (Fig.I.1A,B).

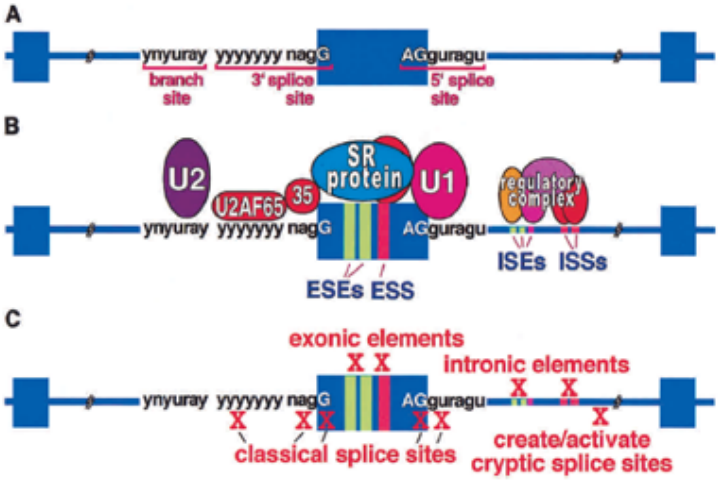


Figure I.1. Classical and auxiliary splicing signals A) Classical splicing signals found in the major class (>99%) of human introns are required for recognition of all exons. **B)** Classical and auxiliary splicing sequence elements and binding factors. Splicing enhancers and silencers in exons (ESEs and ESSs) and introns (ISEs and ISSs) are commonly required for efficient splicing of constitutive and alternative exons. Intronic elements also serve to modulate cell-specific use of alternative exons by binding multicomponent regulatory complexes. **C)** *Cis*-acting splicing mutations. Mutations that disrupt *cis*-acting elements required for pre-mRNA splicing can result in defective splicing that causes disease. (figure from [4]) (n = G, A, U, or C; y = pyrimidine; r = purine).

The splicing machinery, known as the spliceosome, is one of the most complex machineries in the cell. This complexity is a response not only to the necessity of accuracy but also to the connection with previous and subsequent steps of RNA metabolism. Splicing is intimately coupled to transcription, RNA export, nonsense mediated decay and translation [5-7]. Such a broad spectrum of effects requires accurate control. Indeed,

alternative splicing is spatially and temporally regulated during development, regulating protein expression and generating a complicated network of isoform-dependent protein interactions that establish cell functionality [8-12]. Furthermore, the alternative splicing pattern of a cell can be altered by cell activity [13, 14] and also in response to diverse stimuli [15, 16]. Whereas some splicing decisions are regulated by small variations in general splicing factors, others require specific factors whose expression is highly restricted during development (reviewed in [17-19]). The equilibrium between the levels and activity of general and cell-specific splicing factors decides which isoforms are present in every cell type at each developmental stage. With so much energy dedicated to controlling alternative splicing, it is not surprising that unprogrammed changes in isoform ratio affect cellular functions, frequently leading to disease. Mutations disrupting cis-acting splice sequences (Fig. I.1C), the basal splicing machinery or alternative splicing factors have been linked to cellular transformation, metastasis and various hereditary diseases (reviewed in [4, 20-22]).

2. Muscleblind family of proteins: RNA binding proteins with different functions in RNA metabolism.

Muscleblind proteins are characterized by the presence of Cys_3His (CCCH) zinc finger domains. These motifs have been described to perform very diverse functions including DNA recognition, transcriptional activation and repression, RNA packaging and turnover, protein folding and assembly and lipid binding [23] [24] [25]. TIS11d is a vertebrate CCCH protein that binds *TNF-α* mRNA [26] and this interaction has been studied in depth. The two zinc fingers of TIS11d bind in a symmetrical fashion to adjacent 5'-

UAUU-3' sub-sites on the ARE region [27]. RNA bases and conserved aromatic residues get intercalated in a stack and both electrostatic and hydrogen-bonding contribute to binding the protein to the RNA.

The characterization of *Muscleblind-Like1* (*Mbnl1*) knockdown (*Mbnl1*$^{\Delta E3/\Delta E3}$) mice [28] showed that *muscleblind* function is required for proper alternative splicing regulation. *Mbnl1*$^{\Delta E3/\Delta E3}$ mice developed myotonia, histological muscle defects, cataracts and impairment of splicing of several muscular transcripts (Table I.1) [28-30]. A set of exons that undergo a synchronized switch between post-natal day 2 and 20 in wild type, were misregulated in *Mbnl1*$^{\Delta E3/\Delta E3}$ mice [30]. Interestingly, Mbnl1 was translocated from a predominantly cytoplasmic to nuclear distribution during this post-natal interval. In contrast to the *Mbnl1* mutant, no myotonia or splicing defects were detected in *Mbnl2* knockdown mice and muscle histology was normal.

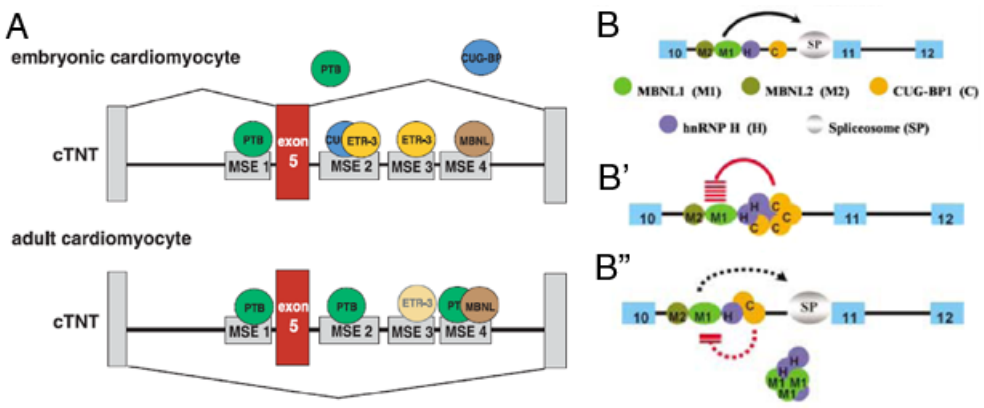


Figure I.2. Mechanisms of splicing regulation by MBNL1. A) Vertical boxes represent exons, lines introns and MSE boxes muscular splicing enhancer sequences. Dynamic balance model proposed by Ladd et al. [31] in which maintenance of exon E5 inclusion repressors MBNL1 and PTB levels in adult cardiomyocyte, together with the reduction of E5

inclusion activators CUG-BP1 and ETR-3 levels, originates a switch in cTNT splicing. E5 is only present in adult cardiomyocytes (figure taken from [31]). **B-B")** Model of coordinate regulation of *IR* splicing in myoblasts. Blue boxes are exons (numbers indicate nomenclature used); lines are introns; circles are proteins. (**B)** MBNL proteins are required facilitators of *IR* E11 splicing but a limiting factor might be needed as elevated levels of MBNL proteins do not alter the *IR* splice equilibrium. **(B')** hnRNP H and CUG-BP1 form an RNA-dependent suppressor complex that is required for the maximum repression of *IR* E11. **(B")** Increased levels of MBNL1 can partially dampen the inhibitory activity of hnRNP H by physical interaction with it.

Human Muscleblind proteins (MBNL1-3) regulate alternative splicing of *cardiac troponin T* (*cTNT* or *TNNT2*) and *insulin receptor* (*IR*) minigenes in HEK293 cell culture [32], although the effects of MBNL proteins on *IR* mRNA were not reproduced in myoblasts [33]. Over-expression of MBNL1, 2 and 3 promoted the exclusion of human and chicken *cTNT* exon 5 (E5) and inclusion of human *IR* exon 11 (E11) in mature transcripts, whereas elimination of MBNL1-3 function by small interfering RNA generated the opposite effect in exon choice. Studies on human and chicken *cTNT* mRNA led to the description of a consensus sequence YGCU(U/G)Y (Y is a pyrimidine) for MBNL1 binding [32]. Two of these binding sites were identified close to the 5' end of *cTNT* E5, which is only included in foetal cardiomyocytes. These MBNL1 binding sequences were required for repression of E5 inclusion by MBNL1. A dynamic balance between activators and repressors of splicing has been proposed to regulate the developmental switch in E5 inclusion in chicken and mouse heart development [31] (Fig. I.2A). The expression levels of MBNL1 and polypirimidine tract binding protein (PTB), previously described to repress E5 inclusion in skeletal muscle [34], are maintained during heart development. In contrast, levels of CUG binding protein 1 (CUG-BP1) and embryonic lethal abnormal vision type RNA binding protein 3 (ETR-3), both

activators of *cTNT* E5 inclusion [34, 35], are down-regulated in adult cardiomyocytes. Thus, during the foetal stage, CUG-BP1 and ETR-3 preferentially bind the regulatory regions promoting exon inclusion but, when their levels decrease in adult cardiomyocytes, MBNL1 and PTB repress E5 inclusion. Coordinated physical and functional interactions between hnRNP H, CUG-BP1 and MBNL1 dictate *IR* splicing in myoblasts [33]. In this case, MBNL1 has been described to activate E11 inclusion by repressing hnRNP H activity (Fig. I.2B). CUG-BP1 and hnRNP H form a repressor complex which is inhibited by binding of MBNL1 to hnRNP H. Hence, the mechanisms by which MBNL1 regulates *cTNT* and *IR* splicing seem to be different.

Although MBNL2 and 3 were also able to regulate splicing of *cTNT* and *IR* minigenes in cell culture, they have not been shown to regulate any physiological splicing event. Indeed, MBNL2 is thought to be involved in the RNA-dependent localisation of Integrin α3 protein to focal adhesions by binding to a specific ACACCC motif in the 3' untranslated region (UTR) of the mRNA [36]. Furthermore, *MBNL3* expression inhibits muscle differentiation in cell culture [37], whereas *Mbnl1* knockdown mice show impaired muscle development pointing to Mbnl1 as a pro-myogenic factor [28]. Thus, although the high level of structural and functional similarity of these proteins could explain their similar behaviour in minigene splicing assays in cell culture, the physiological functions that each protein performs might be regulated by tissue-specific and developmentally regulated expression, post-transcriptional modification, sub-cellular localisation and other post-translational processes.

3. **Muscleblind expression is regulated both transcriptional and post-transcriptionaly.**

Human *MBNL1* and *MBNL2* expression was detected in brain, kidney, liver, pancreas and muscle tissues [37, 38]. While *MBNL2* was found in similar levels in all tissues, *MBNL1* was more abundant in skeletal muscle and the heart. *MBNL3* transcripts, however, were expressed at much lower levels, with peak expression in the placenta and no expression in skeletal muscle. In contrast to humans, the expression of mouse *Mbnl1*, *Mbnl2* and *Mbnl3* genes in adults was more uniform across tissues, although *Mbnl1* transcript levels were higher in heart and lower in skeletal muscles [39]. During embryonic development, all three *Mbnl* genes were prominently expressed in the developing head region and forelimb bud by 9.5 days post conception (dpc). Later in development, the expression patterns of *Mbnl1* and *Mbnl2* were similar (detected in tongue, mandibular and maxillary regions, lips, thymus, lung and intestines among other tissues) while *Mbnl3* expression was stronger in the thymus, lung and intestines. *Muscleblind-like2* and *3* (*tmbnl2a* and *tmbnl3*) expression patterns have been also characterized in *Takifugu rubripes* [40]. No differences between juvenile and adult stages were found. Expression of *tmbnl2a* was ubiquitous but predominant in the heart and brain and *mbnl3*, in contrast to its vertebrate homolog, was also detected in all analyzed tissues including skeletal muscle. Expression of *Drosophila muscleblind* mRNA is mainly detected in muscle and the nervous system ([41], see below)

In addition to these tissue-specific expressions, different post-translational modification sites were detected in human MBNL protein sequences [42]. No functional studies have been performed, but alternative splicing

generates transcripts with different motifs in the final protein, which probably confer diverse functional abilities. Nine splicing variants of *MBNL1* transcripts, three of *MBNL2* and six of *MBNL3* have been reported (reviewed in [42]). The C-terminal region was the most variable but some of the isoforms generated also lacked zinc fingers or the linker between them, which are required to interact with RNAs [43]. Alternative splicing also generates four mature transcripts from *Drosophila muscleblind*, which show different developmental expression patterns and give rise to protein isoforms differing in their C-terminal regions ([44, 45]; see below).

4. **Muscleblind proteins in development.**

A number of studies have related Muscleblind proteins to various developmental processes. Analysis of *Drosophila muscleblind* mutants showed that *mbl* is required for proper terminal differentiation of the peripheral nervous system, photoreceptors and muscles in flies [41, 44, 46]. Mouse *Mbnl1* loss-of-function led to the maintenance of the foetal splicing pattern in skeletal muscle transcripts [28, 29]. MBNL1 is implicated in splicing regulation in chicken and mouse cardiomyocytes, participating in a developmental switch from foetal to adult splicing pattern [31]. In contrast, human MBNL3 has been described as an anti-myogenic factor [37]. Transcripts expressed in murine 3T3-F442A pre-adipocytes before and after differentiation induction by growth hormone (GH) were compared [47]. *Mbnl2* messenger resulted rapidly up-regulated by GH signalling suggesting that, similarly to MBNL1 implication in cardiomyocytes differentiation, *Mbnl2* could have a role in adipocyte differentiation in mouse. A study performed with murine foetal liver stem cells identified *Mbnl3* transcripts being expressed in hematopoietic stem cells [48].

Perhaps murine *Mbnl3* is inhibiting the differentiation of hematopoietic stem cells, as it is in the case of human MBNL3 inhibiting myogenesis. Finally, human MBNL3 was upregulated in H295R cells after stimulation with angiotensin. [49].

5. **Muscleblind proteins in disease.**

Most of the functional data about Muscleblind proteins were generated because of the implication of the MBNL proteins in human diseases. In particular, MBNL proteins are key factors in myotonic dystrophy (DM1 and DM2) [28, 50, 51] and genetic data implicate *muscleblind* in the Spinocerebellar Ataxia 8 (SCA8) disorder [52]. More recently, MBNL proteins have also been implicated in the pathogenic mechanism of Huntington Disease-Like 2 (HDL2) [53]. Finally, altered levels of *MBNL* expression have been described in several tumours, squizophrenia, and sporadic idiopathic pulmonary arterial hypertension [36, 54, 55].

5.1. **Myotonic dystrophy.**

Myotonic dystrophy is a complex multisystemic disease involving musculature, and the nervous and endocrine systems [56]. Dynamic mutations in untranslated regions of two genes form the molecular basis of DM. In DM1, a CTG trinucleotide repeat located in the *Myotonic Dystrophy Protein Kinase* (*DMPK*) 3'UTR is expanded from its normal range (5-37) to more than 1000 repeats, with a correlation between symptom severity and expansion length [57-62]. In DM2, a CCTG tetranucleotide is expanded in the first intron of ZNF9 [63], giving a very similar, although milder, clinical

manifestation. The main symptoms of myotonic dystrophies are muscular weakness, myotonia, cataracts, cardiac problems, insulin resistance, male infertility and neurological disorders such as excessive daytime sleepiness. At the molecular level, DM is characterized by a general impairment of alternative splicing regulation, which leads to the presence of foetal transcript isoforms in adult tissues.

5.1.1. **Contribution of a mutant RNA gain of function and sequestration of MBNL proteins to DM pathogenesis.**

The inability of *Dmpk* knockout mice to reproduce DM1 symptoms [64, 65], make it unlikely that the disease is a consequence of a lack of *DMPK* function in the heterozygous condition (haploinsuficiency). Also, CTG repeats have been shown to alter chromatin structure locally, thereby affecting the expression of nearby genes [66]. Extensive similarities between DM1 and DM2, despite the unrelated mutations that cause the diseases, suggested a common mechanism independent of the mutated gene. The observation of nuclear accumulation of mutant RNA into aggregates in the cells of myotonic dystrophy patients led to the RNA gain of function hypothesis: expanded sequences are not translated into a toxic protein but form a secondary structure that stabilizes the mRNA; this mRNA is retained in the nucleus and interferes with the function of diverse nuclear factors, thus leading to disease [67, 68].

The RNA gain of function hypothesis is strongly supported by murine models in which the expression of an expanded CUG containing RNA in an unrelated genomic context reproduced most of the symptoms of myotonic dystrophy (*HSA*^{LR} mice; [69]; table I.1). RNA molecules carrying expanded

repeats fold into a stable hairpin structure that generates the nuclear retention of the mRNAs [70]. Several nuclear factors have been described to be retained by these RNAs, including transcription and alternative splicing factors [51, 71, 72]. MBNL proteins co-localise with RNA foci in the myoblasts, fibroblasts, cardiomyocytes, neuromuscular junctions and neurons of DM patients [73-77]. MBNL proteins are the only nuclear factors described to bind CUG repeats in a length-dependent manner, an observation that suggested an explanation for the correlation between the length of the expansion and the severity of the disease: the larger the expansion, the stronger the loss of MBNL function [51]. Two main results demonstrate that *MBNL* loss-of-function contributes to DM pathogenesis. First, *Mbnl1* knock-down mice reproduce the main features of myotonic dystrophy: myotonia, muscle histopathology, cataracts, and mis-splicing of *Chloride Channel1*, *TnnT2* and *3* and other transcripts (Table I.1) [28]. Secondly, *MBNL1* over-expression in mice expressing expanded CUG repeat RNA reverses the splicing defects observed in this DM model [29].

Gene	Accession number	Alt. spliced exon number	Neonatal isoform[a]	Post-natal splicing switch[b]	Aberrant splicing in DM			
					Mouse[c]		Human[d]	
					HSA^{LR}	*Mbnl1*$^{-/-}$	DM1	DM2
Sercal	NM_007504	22	exon22−	+	+	+	+	+
ZASP	AY206013	11	exon11+	+	+	+	+	+
z-Ttn	NT_039207.3	Zr4	Zr4+5+	+	+	+	+	+
z-Ttn	XM_130322	Zr5	Zr4+5+	+	+	+	+	+
m-Ttn	XM_130312	Mex5	Mex5+	+	+	+	+	+
Nrap	AY177622	12	exon12	+	+	+	+	+
Capn3	X92523	16	exon16−	+	+	+	+	+
Alp	AF002283	5a	exon5a+5b−	+	+	+	+	+
Alp	NT_039460.3	5b	exon5a+5b−	+	+	+	+	+
Fhos	NT_078575.3	11a	exon11a−	+	+	+	+	+
Gfat1	AF334736	10	exon10−	+	+	+	+	−
Mbnl1	BC060031	7	exon7+	+	+	+	+	+
Mbnl2	NM_175341	7	exon7+	+	+	+	+	+

Table I.1. Failure of MBNL1-dependent postnatal splicing transitions in DM1 and DM2. Misregulated alternative splicing in two mouse models of DM1 is concordant with human DM1 and DM2 molecular defects. Mice over-expressing long CUG repeat tract (HSA^{LR}) also reproduce these defects. [a] Indicates the isoform that is preferentially expressed in neonatal muscle at post-natal day 2 (P2) when compared with adult WT muscle. [B] Denotes exons that show post-natal splicing transition between P2 and P20 in WT hind limb muscle. [C] Denotes exons that show miss-regulated alternative splicing in adult (6 month) HSA^{LR} transgenic or *Mbnl*$^{-/-}$ mice when compared with WT mice of appropriate background strain. [d] Denotes exons that show misregulated alternative splicing in quadriceps muscle from DM1 or DM2 patients compared to healthy individuals (table from [30]).

5.1.2. **DM: a network of intricate molecular defects.**

In spite of strong evidence supporting the contribution of MBNL sequestration to the disease, the array of molecular alterations in DM has not been completely elucidated yet. *Mbnl1* loss-of-function mice show a phenotype very similar to that observed in DM patients, but no muscular weakness. Another molecular alteration characteristic of the myoblasts of patients, the increase in CUG-BP1 activity, is not reproduced in these mice [28]. Moreover, experimental data showed that sequestration of MBNL by CUG containing RNA could be separated from splicing misregulation [78] thus suggesting that CUG repeat RNA would trigger additional molecular

changes. One of them seems to involve CUG-BP1. CUG-BP1 is an alternative splicing factor that antagonizes MBNL1 in several RNA targets misregulated in DM [35, 79, 80]. Transgenic mice over-expressing CUG-BP1 also reproduce the typical splicing impairment of DM [81]. Thus, although the mechanism that up-regulates CUG-BP1 activity in DM is not known, it might be contributing to the final systemic effects.

Gene expression is also misregulated in DM. Another factor that is mislocalised in DM1 cells, Sp1, has been implicated in the transcriptional regulation of *muscular specific Chloride Channel 1* (*CLC1*), an RNA that is also mis-spliced in Myotonic Dystrophy. The absence of CLC1 protein is the cause of the myotonia, one of the most severe symptoms of DM patients. Mice over-expressing CUG-BP1 and *Mbnl1* knock-down mice also showed strong reduction in CLC1 levels due to an alteration in *CLC1* RNA splicing [28, 81] similar to that found in DM patients. This means that at least three factors, MBNL1, CUG-BP1 and Sp1, could be contributing to the absence of CLC1 protein that leads to myotonia in DM. Furthermore, CUG-containing RNA hairpins were described as targets of the RNA interference machinery, which could lead to the misregulation of gene expression by a different mechanism [82].

In addition to all these intricate molecular events, basic questions about DM aetiology have not yet been solved. Although considerable effort has been devoted to understanding whether foci formation in DM1 cells is pathogenic, protective, or perhaps irrelevant, the answer is not clear. The expression of 162 CUG repeat containing RNA in *Drosophila* formed ribonuclear foci only in some cell types [83]. Muscleblind protein co-localised well with those foci, but *muscleblind* expression was shown to be

neither required nor sufficient for foci formation. Moreover, those flies did not show any apparent defect, indicating that RNA foci were not toxic in *Drosophila*. In contrast, flies expressing 480 CUG repeat RNA showed muscle wasting and degeneration, which were reduced when co-expressing human MBNL1 [84, 85].

It is also very difficult to distinguish the primary effects of expanded RNA nuclear retention and secondary consequences of those primary defects. Fibre type transitions involving changes in gene expression and splicing regulation occur during development and in response to innervation, neuromuscular activity and hormone signalling [86]. In particular, myotonia causes fast muscles to become slower generating characteristic changes in muscular histopathology [87-90]. Therefore, myotonia could be the cause of some of the histological defects and splicing alterations found in DM muscles. Furthermore, it was found that the administration of sera from mothers of children with congenital myotonic dystrophy impairs muscle maturation in rats, suggesting there is a circulatory factor maintaining the immature skeletal muscle found in these patients [91].

6. ***Drosophila muscleblind.***

Drosophila muscleblind (*mbl*) was identified as a gene required for the morphogenesis of the *Drosophila* peripheral nervous system [46]. *mbl* mutants have neuronal abnormalities at the chordotonal organs, which are embryonic propioreceptors situated at the body wall. The *muscleblind* gene was independently isolated as a suppressor of the *sev-svp2* eye phenotype [44]. s*evenup* (*svp*) acts autonomously to positively regulate the neuronal fate of photoreceptor precursors [92, 93]. Analysis of *mbl* mutant clones

generated during eye development showed that mutant cells fail to differentiate properly. Rhabdomeres were smaller than in wild type tissue and vertical sections showed that they failed to extent into the basal retina [44]. Recently, mutations in the *Drosophila muscleblind* gene were found to increase fly resistance to starvation [77]. The characterization of *mbl* null mutant embryos showed disruption of Z bands (Fig. I.3A-B) [41]. Expression of the Z-band specific protein *kettin* was tested, showing no differences in expression levels between *mbl* mutants and wild type controls, but the protein failed to assemble properly into the Z bands. *mbl* mutants also show severe reduction of extracellular tendon matrix at the indirect muscle attachments to the epidermis (Fig. I.3C-F). However, β_P*S integrin* expression and localisation of the tendon matrix component Tiggrin appeared to be normal. These mutants die at a late embryonic stage; in fact, they die as larvae unable to hatch. When the chorion is artificially removed, the larvae show a severe abdominal contraction. With all these results it was concluded that *Drosophila muscleblind* is implicated in the terminal differentiation of nervous and muscular tissue.

Consistent with this phenotype, characterization of Muscleblind expression by anti-Mbl antibody staining showed expression both in ectodermic and mesodermic derivatives [41] (Fig. I.4). In the ectoderm, *mbl* is detected at Bolwig's organs (containing larval photoreceptor precursors) and segmentally repeated groups of cells in the central nervous system. In the mesoderm, Mbl protein is first detected at late stage 11 and becomes more abundant during germ band retraction. At the end of this process it is expressed throughout the somatic mesoderm, becoming restricted to a subset of mesodermal derivatives by stage 16. Then *muscleblind* is detected in alary, pharyngeal, visceral and somatic musculature, a pattern

that is typical of genes involved in terminal muscle differentiation as *myosin* or *PS2 integrin* [94, 95].

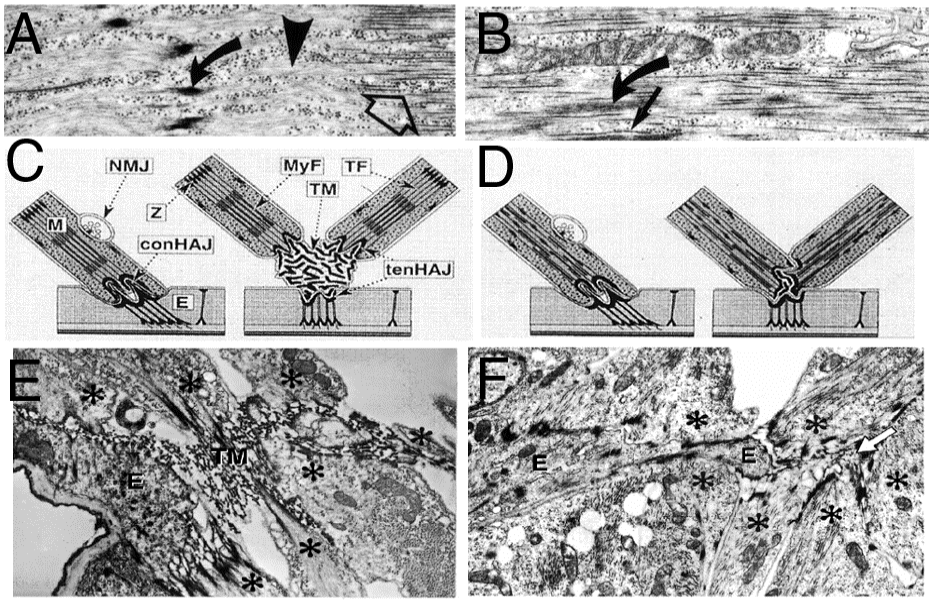


Figure I.3. *muscleblind* mutants show absence of Z-bands and reduction of extracellular matrix at indirect muscle attachments. Muscles (*), epidermal cells (E), and tendon matrix (TM).**A)** Relaxed wild-type muscle shows A-band with thick filaments (open arrow), I-band with only thin filaments (arrowhead), and Z-band in the form of electron dense spindles (bent arrow). **B)** Accumulations of electron dense thin filaments (bent arrows) most likely represent components of the Z-band; the I-band is absent and thick filaments occur adjacent to potential Z-band structures (small arrows). **C-D)** Schematic representation of direct (left) and indirect (right) muscle attachments in wild type **(C)** and *muscleblind* mutant embryos **(D)**. At indirect muscle attachments numerous muscle tips are connected to few epidermal cells via extracellular tendon matrix (E). At *mbl* mutant indirect muscle attachments, TM (white arrow) is severely reduced, and epidermal cells bulge into the body cavity, staying in close contact with muscles (F). (Pictures from [41, 96])

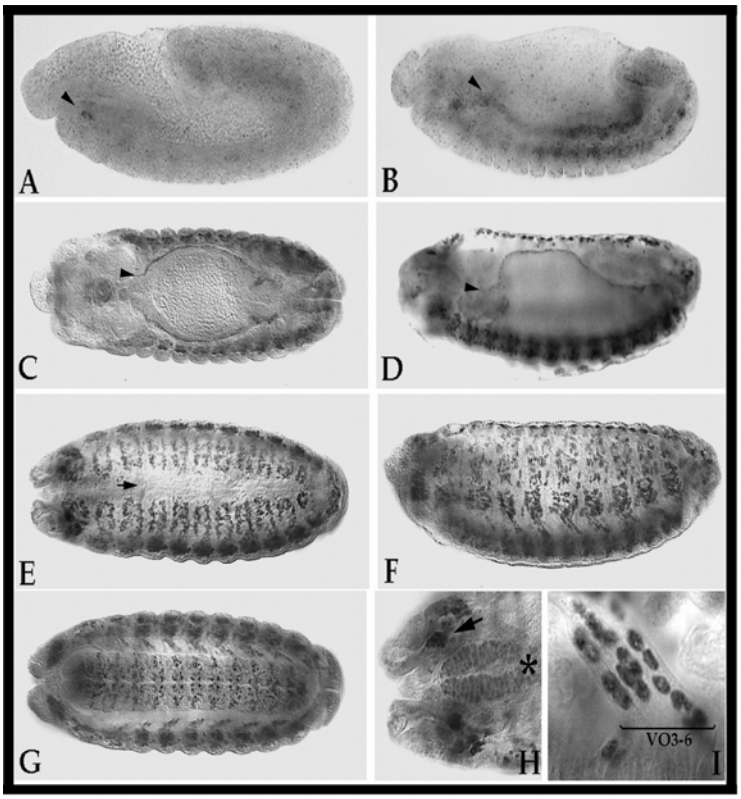


Figure I.4. Muscleblind protein is present in somatic, visceral, and pharyngeal musculature. A) Lateral view, stage 11 embryo with Mbl expression in the cephalic mesoderm (arrow) and a barely detectable signal in the remainder of mesoderm. **B, C)** Mbl expression restricted to visceral (arrow) and somatic mesoderm in a lateral view of a late germ-band retracting embryo **(B)** and a dorsal view of a late stage 13 embryo **(C)**. **D)** Ventral–lateral view of an embryo at stage 15 showing expression in visceral and somatic musculature and CNS. **E)** Dorsal view, stage 16 embryo. Mbl expression is observed in repeating nuclear clusters of fused somatic mesodermal cells but not in the heart precursor cells (arrow). **F)** Lateral view, stage 16 embryo. Mbl-positive clusters of nuclei in locations corresponding to all syncitial fibers of differentiating somatic muscles are shown. **G)** Ventral view, stage 16 embryo. Mbl expression is also detected in the CNS in repeated clusters of cells. **H)** High magnification of the embryo in E showing the expression of Mbl in pharyngeal muscles (asterisk) and Bolwig's organ (arrow). **I)** High magnification of VO3-6 muscles showing individual nuclei stained by α-Mbl antibody. (Figure from [41])

The detection of Muscleblind protein in *Dmef2* mutant flies showed a marked reduction of *mbl* expression in somatic, visceral and pharyngeal musculature thus pointing to *Dmef2* as a transcriptional activator of *mbl* [41]. This activation was confirmed driving expression of *Dmef2* in the epidermis, which led to ectopic detection of Muscleblind. Ectodermic *mbl* signal remains unaffected in *Dmef2* mutants and also some signal persists at mesodermal derivatives indicating that other factors might regulate *mbl* expression. Consistent with these results, chromatin immunoprecipitation experiments isolated four fragments bound by Dmef2 surrounding the *mbl* start site [97].

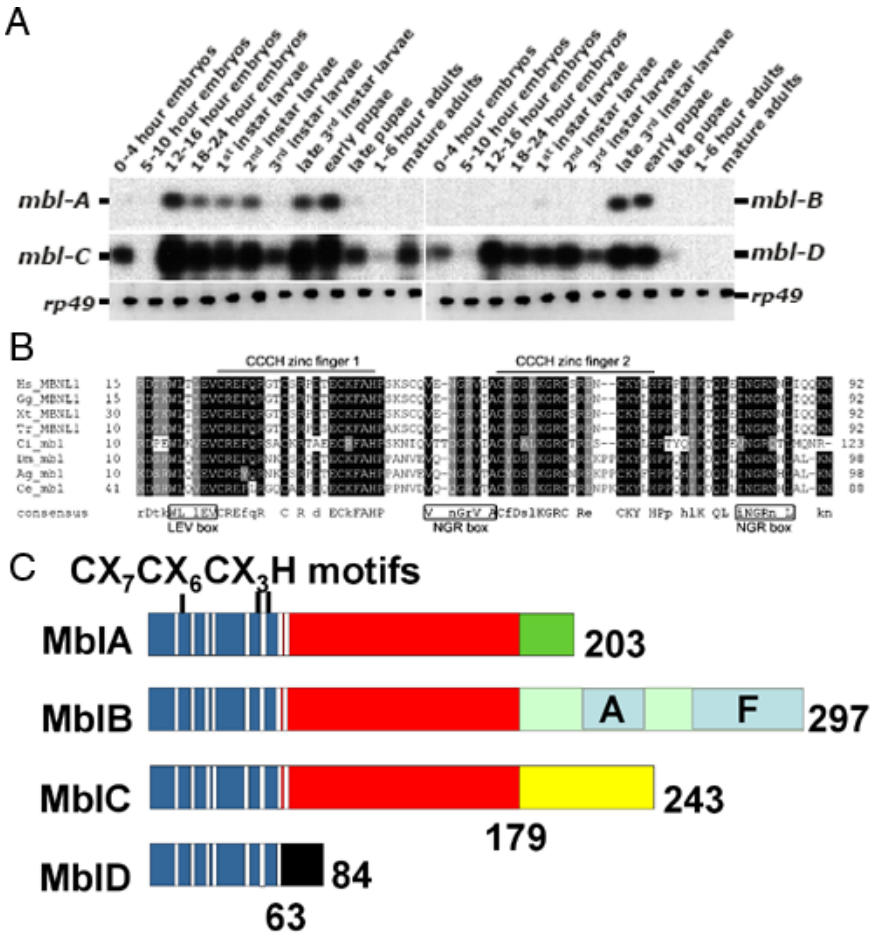


Figure I.5. Alternative splicing generates four *mbl* transcripts coding four CCCH proteins with specific C-terminal regions. A) The analysis of *muscleblind* transcript expression by RT-PCR showed isoform-specific developmental regulation (Figure from J. Houseley from the Institute of Biomedical and Life Sciences, University of Glasgow, UK). **B)** Amino acid alignment within the first two CCCH zinc fingers from flies (Dm), *Anopheles gambiae* (Ag), *Caenorhabditis elegans* (Ce), human (*Homo sapiens*; Hs), *Gallus gallus* (Gg), *Xenopus tropicalis* (Xt), *Takifugu rubripes* (Tr) and *Ciona intestinalis* (Ci). Besides the CCCH zinc finger domain, Muscleblind proteins show highly conserved sequences, such as the LEV and NGR boxes (boxed in the consensus sequence) (Figure from [42]). **C)** Schematic representation of Muscleblind proteins. Size in amino acids is given on the right, shared amino acids are given below. Zinc finger motifs (I, II) and lower complexity regions in MblB (Alanine- (A) and Phenylalanine (F)-rich regions are denoted.

muscleblind transcripts undergo alternative splicing, giving rise to four transcripts with specific 3' end sequences [44]. Isoform-specific RT-PCR analysis showed that *mbl* splicing is regulated during development, as different amounts of the transcript isoforms are detected at different stages of *Drosophila* life cycle [45] (Fig. I.5A). These transcripts encode four CCCH zinc finger proteins (MblA-D; Fig. I5B-C) that share their amino terminal region [44]. MblA, B and C have two zinc fingers of the CCCH type whereas MblD has just one, as the second is truncated before the last histidine. Several putative sites for post-translational modification were identified recently [42]. Also, two low complexity regions are present in MblB (alanine-rich and phenylalanine-rich) but no functional studies have been performed that demonstrate functionality for any of these domains. Interestingly, rescuing the *muscleblind* mutant phenotype by expressing *UAS-mblA*, *B* and *C* transgenes showed that the different isoforms rescued to different extent, supporting the idea that Muscleblind isoforms are not functionally redundant [45]. Over-expression of the UAS-*mblC* transgene in a general embryonic pattern resulted in over 80% rescue of embryonic lethality, whereas *mblB* and *A* rescued to a much lower extent.

In summary, *Drosophila muscleblind* function has been implicated in the terminal differentiation of several tissues, but the only information we have about its molecular function comes from its vertebrate homologues. When this thesis project was initiated, human MBNL proteins had recently been related to the DM pathogenic pathway. Since then, relevant functional data have been published showing the role of MBNL in RNA metabolism, particularly in splicing regulation. Our work has generated important data about the conservation of Muscleblind protein function and the DM pathogenic mechanism between vertebrates and *Drosophila*. In particular,

we have shown that Muscleblind proteins are alternative splicing factors that regulate the splicing of specific transcripts. Interestingly, we found these *Drosophila muscleblind* targets to be altered by CUG repeat-containing RNA toxic effect. We have also shown that *Drosophila* Muscleblind proteins are functionally diverse and we have given some insights into the possible causes of this diversity. Furthermore, we have demonstrated a possible role for *Drosophila muscleblind* in a crucial cellular process, the cell death.

Objectives

La utopía es como el horizonte: cuanto más te acercas a ella, más se aleja... pero te hace andar.
-Gloria G.-

Since their description in 1997 our laboratory has maintained a continued interest in characterizing the molecular functions of Muscleblind proteins in *Drosophila* and, more recently, their relevance to the mechanism of pathogenesis of myotonic dystrophy. It is within this far-reaching goal that we began this thesis project. We hypothesized that *Drosophila* could serve as a biomedical model and, in particular, that human and *Drosophila* Muscleblind proteins performed similar functions *in vivo*. Therefore, the overall objective of this thesis project was demonstrating the functional conservation between *Drosophila* and human Muscleblind proteins. We also wondered whether the binding of Muscleblind proteins to physiological targets and to CUG repeat containing RNA had different properties in a way that we could find mutations that inhibited the binding to CUG repeat RNA but left their physiological function uncompromised. This overall aim required addressing the following specific research objectives:

1. Assessment of conservation of the myotonic dystrophy pathogenesis pathway in *Drosophila*.

2. Assessment of conservation of the alternative splicing regulatory function described for human MBNL proteins in *Drosophila* Muscleblind.

3. Evaluation of functional diversification of *Drosophila* Muscleblind protein isoforms as naturally occurring protein variants.

4. Analysis of Muscleblind binding to physiological RNA targets and CUG containing RNA.

5. Characterization of functional domains in *Drosophila* Muscleblind proteins.

Materials and methods

Dicen que no nos obligan, pero
tenemos que cruzar el río y
sólo hay un puente.
-Desconocido-

MATERIALS

I. *Drosophila melanogaster* strains.

Oregon-R	Depto. Genética, Facultad de Biología, U.V.
y^1 w^{1118}	Depto. Genética, Facultad de Biología, U.V.
w; mbl^{E27}/CyO	[44]
mbl^{k7103}/CyO	[46]
dgo/CyO, ubiquitous (*ubi*)*-GFP*	available in our laboratory from Muñoz, S.
y^1w^{1118}; UAS-(CTG)480_1.1	[85]
y^1w^{1118}; UAS-(CTG)480_2.1	[85]
w; daughterless-Gal4	Bloomington Stock Center (Indiana)
Mhc-Gal4; ry^-	[98]
mbl^{k7103}/CyO, ubiquitous-GFP	generated in this work
mbl^{E27}/CyO, ubiquitous-GFP	generated in this work
w,mbl^{E16}/CyO, ubiquitous-GFP	available in our laboratory from Monferrer, L.

y^1w^{1118}; mbl^{E27}/CyO, y+; P[w+mC=UAS-mblC:GFP]T15.3 available in our laboratory from Pascual, M.

II. *Escherichia coli* strains.

XLBlue1: recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB laclqZΔM15 Tn10 (Tetr)]. Depto. Genética, Facultad de Biología, U.V.

electro competent DH5α: F-, φ80dlacZΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(rk-, mk+), *phoA*, *supE*44, λ-, *thi*-1,*gyrA*96,

*relA*1. (obtained from Swanson, M., Dept. of Molecular Genetics and Microbiology, University of Florida, USA)

SURE: e14–(McrA–) Δ(mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5 (Kan^r) uvrC [F´ proAB lacIqZΔM15 Tn10 (Tet^r) Amy Camr] (Stratagene; obtained from Swanson, M., Dept. of Molecular Genetics and Microbiology, University of Florida, USA)

BL21: F– ompT hsdSB(rB– mB–) gal dcm (obtained from Smith, C., Dept. Biochemistry, University of Cambridge)

BL21 star (DE3): F- ompT hsdSB (rB-mB-) gal dcm rne131 (DE3)
(obtained from Smith, C., Dept. Biochemistry, University of Cambridge)

Rosetta (DE3): Δ(ara–leu)7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL (DE3) F'[lac+ lacI q pro] gor522::Tn10 trxB pRARE2 (CamR, KanR, StrR, TetR) (obtained from Smith, C., Dept. Biochemistry, University of Cambridge)

III. *Sacharomyces cerevisiae* strains.

L40 coat: MATa, ura3-52, leu2-3,112, his3Δ200, trpΔ1, ade2, LYS::(LexA op)-HIS3, ura3::(LexA op)-LacZ, LexA-MS2 coat (TRP1) [99] (obtained from García, C., Depto. Bioquímica, Facultad biología, U.V.)

YBZ1: MATa, ura3-52, leu2-3, 112, his3-200, trp1-1, ade2, LYS2<(LexAop)-HIS3, ura3<(lexA-op)-lacZ, LexA-MS2 MS2 coat (N55K) [100] (obtained from García, C., Depto. Bioquímica, Facultad biología, U.V.)

IV. Cell lines.

COS: Simian fibroblasts (CV-1 cells) transformed by SV40 that is deficient in the origin of replication region (obtained from Swanson, M., Dept. of Molecular Genetics and Microbiology, University of Florida, USA, and Smith, C., Dept. Biochemistry, University of Cambridge).

HEK293T: Embryonic human kidney containing SV40 large T antigen (obtained from Swanson, M., Dept. of Molecular Genetics and Microbiology, University of Florida, USA, and Smith, C., Dept. Biochemistry, University of Cambridge).

S2: Schneider's Line S2 obtained from dissociated embryos (obtained from Zhou, L. Dept. of Molecular Genetics and Microbiology, University of Florida, USA, and Llorens, J., Depto. Genética, Facultad de Biología, Universidad de Valencia).

V. *Drosophila melanogaster* cDNAs.

muscleblind (*mbl*) transcript cDNAs in pBluescript were available in our laboratory [44]. Accession numbers are the following:

- *mblA* AF001625
- *mblB* AF001626
- *mblC* AF001536
- *mblD* AF001422

bruno cDNAs and putative *CLC1* homologues were obtained from the Berkeley Drosophila Genome Project (BDGP):

- *bruno1* (*bru1, aret*) LD29068
- *bruno2* (*bru2*) LD19052
- *bruno3* (*bru3*) LD31834

- *CG6942* GH23529
- *CG8594* RE63672
- *CG5284* LD07266

VI. Vectors.

pEGFP-N3: expresses a protein of interest fused in N-t to GFP in eukaryotic cells (Clontech)

pEGFP-C1: expresses a protein of interest fused in C-t to GFP in eukaryotic cells (Clontech)

pACT2: expresses the protein of interest fused to the Gal4 activation domain in *S. cerevisiae* [99]

pIIIAMS2.2: expresses RNAs fused to two MS2 binding sequences in *S. cerevisiae* [99]

pFP105 (modified from pFP98): expresses GST and His_6-tagged proteins in *E.coli* [44]

pGEX-6P-1: expresses GST-fused proteins in *E.coli* (Pharmacia)

pIEI-4: expression vector for *Drosophila* S2 cells (Novagen)

pSG5: vector for minigene expression in eukaryotic cells (Stratagene)

pSP72: vector for expanded DMPK 3'UTR expression in eukaryotic cells (Promega)

pGEM4T3: control DNA in cell transfection assays (Amersham)

VII. Constructs.

pIEI-Diap1 obtained from Zhou, L. (Dept. of Molecular Genetics and Microbiology, University of Florida, USA)

pIEI-LacZ1 obtained from Zhou, L. (Dept. of Molecular Genetics and Microbiology, University of Florida, USA)

MBNL1-GFP obtained from Swanson, M. (Dept. of Molecular Genetics and Microbiology, University of Florida, USA)

pSP72-3'DMPK$_{300}$ obtained from Swanson, M.

pSG5-tnnt3 obtained from Swanson, M.[29]

pEGFP-N3-MblC$^{\Delta sumo}$ available in our laboratory from Pascual M.

Table M.1. presents a summary of constructs involving Muscleblind isoforms generated in this work

	pEGFP-N3 (GFP)	**pMV1** (myc)	**pMV2** (myc-GFP)	**pIEI4** (myc)	**pFP105** (GST/His$_6$)	**pGEX6** (GST)	**pACT2** (G4AD)
MblA	+	+	+	+	+		+
MblB	+	+	+	+	*	#	+
MblC	+	+	+	+	+	#	+
MblD	+	+	+	+			+
MblC$^{\Delta SUMO}$	\	+		+			
ZC							+
CCCH							+

Table M.1. Epitope tagging of Muscleblind coding sequences generated in this work. First lane shows vectors used. The epitope to which Mbl is fused to is represented in brackets (GFP = Green Fluorescent Protein; G4AD = Gal4 activation domain). ZC is a fragment containing the common region of MblA, B and C. CCCH is a fragment containing the two zinc fingers. MblC$^{\Delta SUMO}$ is a mutated variant of MblC. pMV1 vector is a derivative of pEGFP-N3 where GFP was substituted by the myc epitope. pMV2 is a derivative of pEGFP-N3 where the myc epitope was inserted between Mbl protein and GFP. A + indicates that the construct was generated in this work; * designates that the C-t sequence was not confirmed by sequencing; # designates that the Mbl isoform sequence is correct, but transferring to a fresh vector is needed as they come from PCR; \ means that was available in our laboratory.

VIII. Primers.

name	sequence	use	t ºC
TNTE2	CGACGATGAAGAGTACAC	*Drosophila troponin T*	55
TNTE6	CTCTGGATCGCCCTCTCC		
TNT4	TTTTCAACCCTAAACCGTAGC	*Drosophila troponin T*	56
TNT200	ACTCGGTGATGTATTCTTTCAG		
6942/E12	GAAGTCCACCACGCTCCAGG	*Drosophila CG6942*	
6942/E17	GCTCCCATTGCTTCTGATCCTCCG		
6942/E1	CCGAAGATCAAGTCGCACGCC	*Drosophila CG6942*	
6942/E4	GAGCTTTCTGGCTTCATCC		
5284/E7	GGGTGTTATTGGCACATTC	*Drosophila CG5284*	60
5284/E9	GAAGGTTCTCAACATCGTCC		
ACTE5	TTATCCATCGCCATCGTC	*Drosophila α-actinin*	65
ACTE10	TTGAAGTTGGTCTCCAGC		
RP49for	ATGACCATCCGCCCAGCATAC	*Rp49*	65
RP49rev	ATGTGGCGGGTGCGCTTGTTC		
Xho-Mbl	CCGCTCGAGATGGCCAACGTTGTCAATATG	5'Mbl ORF in pEGFP-N3	See reverse primer
MblA-gc-E	CGGAATTCCGAATTGACTTCATTGGATAC	3'MblA ORF in pEGFP-N3	55
MblB-gc-E	CGGAATTCCGGCATGCAACAAAAAAGGC	3'MblB ORF in pEGFP-N3	60
MblC-gc-E	CGGAATTCCGTCTTGGCACACCGGGAGGG	3'MblC ORF in pEGFP-N3	65
MblD-gc-E	CGGAATTCCGGCAGATTAATTTTTTACTTAC	3'MblD ORF in pEGFP-N3	60
Eco-Mbl	GAATTCCGGCCAACGTTGTCAATATGAACAGCC	5'Mbl ORF in pACT2	See reverse primer
MblA-Xho	CCGCTCGAGCAATTGACTTCATTGGATAC	3'MblA ORF in pACT2	55
MblB-Xho	CCGCTCGAGCGCATGCAACAAAAAAGGC	3'MblB ORF in pACT2	60
MblC-Xho	CCGCTCGAGCTCTTGGCACACCGGGAGGG	3'MblC ORF in pACT2	65
MblD-Xho	CCGCTCGAGCGCAGATTAATTTTTTACTTAC	3'MblD ORF in pACT2	60
ZC-Xho	CCGCTCGAGCCTCTAATCTGTCGGAACGTGG	3'Common Region in pACT2	55

CCCH-Xho	CCGCTCGAGCCTTGAGGGCCAAATGATTGCG	3'Zinc Fingers in pACT2	55
Sma-EcoA	GGGTATACGCATGCTTCG	3 hybrid adaptor	
Sma-EcoB	AATTCGAAGCATGCGTATACCC		
Xba-SphA	CTAGACCATGGATATCCCGGGCATG	3 hybrid adaptor	
Xba-SphB	CCCGGGATATCCATGGT		
ACTE7f	CCCACAAGACAATACACC	Actn1 in PIIIA/MS2.2	55
ACTI7r1	ACATGCATGCCTTAGAACGAGGAAGGC		
ACTI7f1	GGGCCTTCCTCGTTCTAAG	Actn2 in PIIIA/MS2.2	55
ACTI7r2	ACATGCATGCGTACCGCTGCGACCTTG		
Nde-Mbl	GGAATCCATATGATGGCCAACGTTGTCAATATG	5'Mbl ORF in pFP105	See reverse primer
MblB-Eco	CGGAATTCGCATGCAACAAAAAAGGC	3'MblB ORF in pFP105	65
MblC-Eco	CGGAATTCTGGCACACCGGGAGGG	3'MblC ORF in pFP105	65
MblD-Eco	CGGAATTCGCAGATTAATTTTTTAC	3'MblD ORF in pFP105	65
ZF-Eco	GCATCAGTGAATTCTTGAGGG	3'Zinc Fingers in pFP105	60
T7actn1D	CGTAATACGACTCACTATAGGGAACCCACAAGACAATACACC	Actn1 *in vitro* transcription template	63
T7actn1R	ATACTTAGAACGAGGAAGGC		
T7actn2D	CGTAATACGACTCACTATAGGTATGCCTTCCTCGTTCTAAG	Actn2 *in vitro* transcription template	57
T7actn2R	CTGCGACCTTGGCAGTG		
T7actn5R	CTGTGAACGTGTGCGTGTTG	Intron 6 *in vitro* transcription template	65
T7actn6D	CGTAATACGACTCACTATAGGGCCCAAAAGGTCTGTTATACG	Intron 7 *in vitro* transcription template	65
T7actn6R	CTTGAGCACCTTGCAGATCC		
T7CUGD	CGTAATACGACTCACTATAGGTCCTTGTAGCCGGGAATG	CUG repeats *in vitro* transcription template	
T7CUGR	AATGGTCTGTGATCCCCC		

T7CAGD	CGTAATACGACTCACTATAGGAATGG TCTGTGATCCCCC	CAG repeats *in vitro* transcription template	
T7CAGR	TCCTTGTAGCCGGGAATG		
myc-dir	TCGACGAGCAGAAGCTGATCAGCGAG GAGGACCTGG	myc tag into pEGFP-N3	
myc-rev	GATCCCAGGTCCTCCTCGCTGATCAG CTTCTGCTCG		
5ACTMGE5	GGCGAATTCTTCTGCGCCCTTATCCA TCGC	*α-actinin* minigene	60
3ACTMGE9	GGCGAATTCCCAGACGCTCGTACTCC TCCATG		
1938	GCTGCAATAAACAAGTTCTGCTTT	pSG5	56
1956	AGAATTGTAATACGACTCACTATAGG GC		
S-myc-Nf	TCGACGAGCAGAAGCTGATCAGCGAG GAGGACCTGGGAGCGGGCCCATAGTA GGC	myc tag to build pMV	
S-myc-Nr	GGCCGCCTACTATGGGCCCGCTCCCA GGTCCTCCTCGCTGATCAGCTTCTGC TCG		
Xho29	GGCCTCGAGATGTTCACCAGCCGCGC TTC	Bru1 ORF in pEGFP-N3	55
29Eco	GGCGAATTCCAGTAGGGCTTCGAGTC CTTGG		
Xho19	GGCCTCGAGATGATGTTGCAATCCTT GAG	Bru2 ORF in pEGFP-N3	60
19Bam	GGCGGATCCTAAAAATTGCAAGTCGG AAAATGG		
Xho31	GGCCTCGAGATGGTTCATATTATTGA ATTG	Bru3 ORF in pEGFP-N3	55
31Eco	GGCGAATTCCAATAGGGTCGACTGGC ATC		

Table M.2. Primers used in this work. Primer name, sequence, application (use) and annealing temperature if used in PCR are given.

METHODS

I. Drosophila work.

I.1. Sequence homology searches.

Human nucleotide sequences were used to search homologous sequences in the *Drosophila* genome with the Washington University Blast software package at http://blast.wustl.edu

I.2. Assessment of mexiletine effect.

I.2.1. Mexiletine administration.

The mexiletine dose for humans is 8.6 mg/ kg·per day. Assuming an average larval weight of 1 mg, that supposes a dose of $3.6*10^{-5}$ μg/h for larvae. Considering the food ingested equivalent to the weight gain from embryo to pupae, larvae eat approximately 21.7 μg /h. That means 0.36% mexiletine in the fly food for a dose equivalent to the human dose. With 126 mg of excipient (starch) per 326 mg of pharmacological presentation of mexiletine, 0.035 mg of medicine were added per ml of food. 0.013 mg of excipient were added per ml of food as control. The appropriate amount of mexiletine or excipient powder was added to previously mixed and cooled fly food. Everything was then vigorously mixed and aliquoted into culture tubes.

I.2.2. Fly viability.

w; mbl^{E27}/CyO and *mbl^{k7103}/CyO* flies were crossed with *dgo/CyO, ubi-GFP* to obtain *mbl^{E27}/CyO, ubi-GFP* and *mbl^{k7103}/CyO, ubi-GFP* flies. *mbl^{E27}/CyO, ubi-GFP* and *mbl^{k7103}/CyO, ubi-GFP* flies were crossed to obtain the transheterozygous *mbl^{E27}/ mbl^{k7103}*, which behaved as a hypomorphic mutation and reduced adult fly viability. *Oregon-R* (*OrR*) flies were used as wild type control for the potential toxic effects of mexiletine. 60 virgin females and 36 males were crossed in culture bottles. Emerged adults were counted daily. *mbl^{E27}/ mbl^{k7103}* flies were identified as GFP negative using a GFP fluorescence module mounted on a Leica MZ APO stereo microscope.

Two crosses with 1x and 10x doses of commercial mexiletine were set up with *mbl* strains in both directions and OrR. Excipient was added at 10x dose as control.

I.2.3. Lifespan.

Viable *mbl^{E27}/ mbl^{k7103}* flies from the viability study were kept in tubes with the same dose of mexiletine and the number of dead flies was counted every day when possible. Fly food was changed twice a week. At least 90 flies, in seven to eight samples, were analysed in this assay from each genotype/mexiletine dose combination.

The longevity of flies expressing 480 CUG repeat RNA was registered in parallel to control flies carrying the *UAS-(CTG)480* transgene or the *Gal4* insertion alone. *y^{1}w^{1118}; UAS-(CTG)480_1.1* virgin females were crossed to

w; daughterless-Gal4 and *Myosin heavy chain-Gal4; ry*$^{-}$ males to express the CUG repeat containing RNA in a general or muscle-specific pattern, respectively. Flies carrying the UAS and the Gal4 insertions alone were used as controls. Twenty flies were introduced in each vial and a minimum of two replicas per genotype were made. For those days where dead flies were not counted, the last number registered is taken as an approximation to calculate the percentage of surviving flies that day.

I.2.4. Climbing assay.

mbl^{E27}*/CyO, ubi-GFP* and *mbl*^{k7103}*/CyO, ubi-GFP* flies were crossed in both directions using 15 virgin females and 8 males in vials containing a 1x or 2x dose of mexiletine or excipient. 30 synchronized flies (collected within 48 h periods) were tested for their climbing ability by counting the number of flies that were above a line at eight cm height after 18 s. Two tubes per cross direction were made at each mexiletine dose. Five replicates per assay were made. The day of the assay flies were changed to a new tube to increase their climbing activity. As the assay tube was larger than the habitual culture tube, flies were allowed ten minutes to get used to it. The experiment was carried out at 25ºC and measurements were around noon.

I.3. *In situ* detection of transcripts in *Drosophila* embryos.

For whole-mount *in situ* hybridization OrR embryos were fixed and hybridized as described [101] with minor modifications. Antisense and control (sense) RNA probes were DIG-labelled using the DIG RNA labelling kit following the recommendations from the manufacturer (Roche). Table

M.2. summarizes the labelling reactions performed. Embryos were analyzed with a DM LB2 Leica microscope using bright field microscopy.

Gene	cDNA	Probe	Restriction enzyme	RNA polymerase
CG6942	GH23529	Antisense	*Eco*RI	Sp6
CG6942	GH23529	Sense	*Xho*I	T7
CG8594	RE63672	Antisense	*Not*I	T3
CG8594	RE63672	Sense	*Nru*I	T7
CG5284	LD07266	Antisense	*Xba*I	T7
CG5284	LD07266	Sense	*Kpn*I	T3

Table M.3. Reactives for the generation of RNA probes for in situ hybridization. The table shows cDNAs used as template, the restriction enzyme that linearized the plasmid and the RNA polymerase DNA dependent that was used to carry out the *in vitro* transcription.

I.4. Analysis of alternative splicing defects.

I.4.1. Fly collection.

mbl^{E27}/CyO, ubi-GFP and *mbl^{E16}/CyO, ubi-GFP* flies were crossed *en masse* during 2 h in laying pots to obtain synchronized mutant embryos. Egg collection plates were then incubated at 25 or 18ºC and embryos were collected after 16 h of development. *mbl^{E27}/ mbl^{E16}* embryos were identified by the lack of fluorescence and handpicked. Late pupae (two days after puparium formation) and non-staged adults were obtained from crosses between *mbl^{E27}/CyO, ubi-GFP* and *mbl^{k7103}/CyO, ubi-GFP* flies. *y^1w^{1118};UAS-(CTG)480_1.1* virgin females were crossed to *Mhc-Gal4; ry^-* males to drive expression of CUG repeat RNA to the musculature of the fly. F1 flies expressing CUG repeats with Myosin Heavy Chain pattern were collected at the same stages than mutants. Embryos expressing CUG repeats in a broad pattern were also collected from crosses between

y^1w^{1118};UAS-(CTG)480_2.1 females and *w; daughterless-Gal4* males. *OrR* and *y,w* flies were collected as controls.

To further characterise *tnT* alternative splicing, *mbl^{E27}/CyO, ubi-GFP* and *mbl^{k7103}/CyO, ubi-GFP* flies were crossed *en masse* in culture bottles. 10 to 22 h after egg laying embryos, 1 to 2 day-old pupae and adults of more than six hours expressing CUGs (1.1 strain) with the *Myosin heavy chain (Mhc)* pattern were also collected. *OrR*, heterozygous *mbl^{E27}/ CyO, ubi-GFP* and *mbl^{k7103}/ CyO, ubi-GFP*, *y^1w^{1118};UAS-(CTG)480_1.1* and *Mhc-Gal4; ry^-* flies were collected at the same stages as controls.

I.4.2. RNA extraction.

Total RNA was extracted with Tri-reagent (Sigma) following the recommendations from manufacturer and treated with DNase I (Invitrogen). RNA concentration was measured by a spectrophotometer (Eppendorf). Only RNA extractions with A_{260}/A_{280} and A_{260}/A_{230} ratios over 1.8 were saved.

I.4.3. RNA quality on gel.

RNA quality was checked loading samples with 50 % formamide and 17.5 % formaldehyde on 1.2 % agarose, 5.5 % formaldehyde gels. 20 mM MOPS, 5 mM NaAc, 1 mM EDTA pH 7 buffer was used to prepare the gel and the samples, and to run the electrophoresis.

I.4.4. RT-PCR.

Reverse transcription (RT) was performed with 1 µg of total RNA, Superscript II RNase H- and random hexamers following instructions from the provider (Invitrogen). cDNA was used in a standard 20 µl PCR reaction of 25 cycles with Taq DNA polymerase (Need) following the recommendations from the manufacturer to amplify endogenous *Drosophila* transcripts. Products were separated on 2% agarose gels.

TNT4/ TNT200 primers were used to amplify *troponinT* (*tnT*) mRNA from 1 µl of RT reaction. To increase specificity we used a nested PCR approach with the internal TNTE2/TNTE6 primer pair and 1 µl of a 1:100 (or 1:1000 dilution) of the first PCR as template.

Primers 5284/E7 and 5284/E9 were used to amplify alternatively spliced region of *CG5284*. 2 µl of cDNA were necessary to detect any RT-PCR product. Primer pairs 6942/E12-6942/E17 and 6942/E1-6942/E4 were used to amplify *CG6942* mRNA splicing products.

To characterize *Drosophila* α-*actinin* (α-*actn*) splicing pattern, 4 µl of the RT reaction were used as template in a standard 40 µl PCR reaction of 40 cycles with primers ACTE5 and ACTE10. PCR products were purified by NH_4Ac precipitation and half of the final volume was digested with *Sac*I to distinguish α-*actn* isoforms.

1 µl of a 1:100 dilution of cDNA was used in a 25 cycle PCR with primers Rp49 for/rev as RT efficiency and RNA input control. These primers were designed to encompass an intronic region of *Rp49* to detect contamination by genomic DNA in the RNA extraction. No reverse transcriptase controls

were also amplified with the same primer pairs used to amplify *tnT* and α-*actn* cDNA sequences and showed no amplification.

I.5. UV crosslinking and immunoprecipitation (CLIP) of RNA bound to MblC *in vivo*.

I.5.1. Sample collection and UV crosslink.

Virgin y^1w^{1118}; mbl^{E27}/CyO, *y+;* P{w+mC=UAS-*mblC:GFP}T15.3* homozygous female flies were crossed *en masse* with *w; daughterless-Gal4* homozygous males in large laying pots with agar-sucrose petri dishes daubed with fresh yeast paste. Flies were allowed to mate for at least 12 h. Laid eggs, in plates, were incubated 10 additional hours prior to collection. *da-Gal4; UAS-mblC:GFP* embryos were bleach decorionated and extensively washed with running tap water. Embryos were transferred to a 15 ml conical tube with cold Hank's Balanced Salt Solution (HBSS; GIBCO). Embryos were gently centrifuged at 800 rcf for 10 min at 4ºC to remove HBSS buffer. Supernatant was decanted and 1 ml of fresh HBSS was used to resuspend embryos prior to crosslink. In order to obtain a cell suspension, in a second experiment embryos were transferred to eppendorf tubes and homogenized with plastic pestles. A third experiment was carried out using a 5 ml homogenizer. GFP signal from individual cells was checked under fluorescence microscope after homogenization. The same washes were done with the cell suspensions. The final ml was placed in 1 cm tissue culture plates on ice/water bath in Stratalinker at same height as sensor. Three 400 mJ/ cm^2 irradiations were performed, which were followed by mixing of cell suspension after each exposure. Suspension was

collected and spinned at 800 rcf for 10 min at 4ºC. HBSS was removed and the samples frozen at -80ºC until used.

I.5.2. Immunoprecipitation and cDNA synthesis.

The following steps of the CLIP protocol were done by Dr. Dan Tutle (Dept. of Molecular Genetics and Microbiology, University of Florida, USA), while I followed his work. The CLIP protocol was performed as previously described [102] with minor modifications. Immunoprecipitation was carried out using the anti-GFP antibody from Roche as it was previously shown in our laboratory that achieved higher efficiency than other anti-GFP antibodies. The immunoprecipitation worked efficiently and an adequate quantity of radiolabelled RNA was obtained, but ligation of linkers required to synthesise the cDNA failed thus precluding from reaching a final result.

II. Yeast methods.

II.1. Yeast three hybrid assay.

II.1.1. Generation of hybrid RNA expressing constructs.

480 interrupted CAG repeats cloned into pUAST were obtained from L. Monferrer [35, 85]. Sma-EcoA and Sma-EcoB oligos were annealed by standard procedures [103] to generate a short DNA adaptor with *Sma*I and *Eco*RI cohesive ends. The adaptor also included an internal *Sph*I site. Xba-SphA and Xba-SphB oligos were used to generate a DNA adaptor with *Xba*I and *Sph*I cohesive ends including an internal *Sma*I site. *Eco*RI/*Xba*I digested $(CAG)_{480}$, *Sma*I/*Sph*I digested pIIIA/MS2 plasmid and the two

adaptors were included in a 20 μl ligation reaction. *E.coli* [*XLBlue1* or *DH5α* strain] transformation and plasmid DNA extraction were carried out by standard protocols [103]. After isolating a $(CAG)_{480}$ pIIIA/MS2 clone, which expresses CAG-repeat RNA fused to the MS2 RNA, the orientation of the insert was reversed in order to express CUG repeat RNA fused to MS2 RNA. $(CAG)_{480}$ pIIIA/MS2 DNA and pIIIA/MS2 empty plasmid were *Sma*I/*Sph*I-digested and ligated in a standard 20 μl ligation reaction. The resulting constructs showed a distinctive restriction pattern that allowed the quick verification of CTG and CAG repeat DNA-containing plasmids.

The Actn1 fragment was amplified from *Drosophila* genomic DNA with primers ACTE7f and ACTI7r1 by high fidelity PCR (Pwo DNA polymerase, Roche). PCR product was digested with *Sph*I and ligated to *Sma*I/*Sph*I digested pIIIA/MS2 plasmid. This ligation does not reconstitute the *Sma*I site. Actn2 fragment encompassing a cluster of seven MBNL1 consensus binding sequences was also obtained by high fidelity PCR (Pyrobest, Takara) with ACTI7f1 and ACTI7r2 primers. Cohesive ends ligation did not work and phosphorylated PCR product (standard polynucleotide kinase reaction following instructions from provider; Roche) was blunt-end ligated to *Sma*I digested pIIIA/MS2 plasmid. The two fragments were sequenced to confirm both sequence and orientation.

II.1.2. RNA structure prediction.

RNA folding prediction software Mfold was used to predict the secondary structure of the hybrid RNAs [104].

II.1.3. Generation of hybrid protein expressing constructs.

High fidelity PCR reactions following the instructions from the manufacturer (Triple Master, Eppendorf) were performed to amplify Muscleblind open reading frames (ORFs). A common 5'end primer (Eco-Mbl) was used in combination with specific 3' end primers MblA-Xho, MblB-Xho, MblC-Xho, MblD-Xho, ZC-Xho and CCCH-Xho to amplify MblA, B, C and D, the common region of MblA, B and C, and a fragment containing the two zinc fingers respectively. PCR products were cloned in frame with the Gal4 activation domain of pACT2. All constructs were sequence confirmed.

II.1.4. LiAc yeast transformation.

- Pick fresh colony (a big one, no older than two days) in 5 ml YAPD 2x and grow o/n
- Next morning, inoculate 100 ml of fresh YAPD 2x with the o/n culture
- Grow until A_{600}= 0.5-0.7
- Spin 5000 rpm, 5 min, room temperature
- Wash with 25 ml of sterile water: 5 min spin 5000 rpm
- Resuspend in 0.5 ml LiAc/TE 1x (10 mM LiAc; 10 mM Tris HCl; 1 mM EDTA; pH 7.5)
- Prepare DNA in eppendorfs; add 10 μl of 10 mg/ ml carrier DNA (salmon sperm) per tube
- Add 50 μl of cell mix per tube
- Add 300 μl of 40% PEG/ 1x TE/ 1x LiAc freshly prepared from 50% PEG, 10x TE, and 10x LiAc solutions
- Incubate 30 min at 30ºC with agitation

- Heat shock 15 min 42ºC
- Spin 1 min 6000 rpm
- Take off supernatant and resuspend in 100 μl of sterile H_2O
- Plate 50 μl in selective medium (SD + amino acids)
- Incubate at 30ºC until transformants grow (2-5 days)

This protocol, modified from [105], gave better results than others found in TRAFO website (R. Daniel Gietz, University of Manitoba) or Current Protocols in Molecular Biology [106].

Yeasts were co-transformed with pIIIA/MS2 and pACT2 constructs in all combinations. pACT2-IRP and pIIIA/MS2-IRE were used as positive control interaction [99] and as heterologous protein and RNA for negative controls in our experiments. Empty vectors were also used as negative controls.

II.1.5. Confirmation of interaction specificity.

L40 coat yeast strain was transformed with the pACT2 and pIIIA/MS2 constructs in co-transformation experiments or experiments in which RNA expression plasmid was transformed first and protein expression plasmid was transformed second, or vice-versa. Yeasts carrying both plasmids were platted onto synthetic dextrose medium without tryptophan, leucine, histidine and adenine. After a few days, four white grown colonies of each co-transformation combination were platted onto the same media with increasing concentrations of the *HIS3* inhibitor 3-aminotriazol (3-AT) to check specificity of the interaction.

III. Cell culture assays.

III.1. Subcellular localisation assays in mammalian cells.

III.1.1. Cloning.

pSP72 vector carrying human DMPK 3'UTR with an expanded region of 300 CTG repeats was transformed into SURE cells following provider instructions (Stratagene). The lack of recombination in this strain favours the stability of long repeat DNAs. DNA extracted from individual clones was digested with *Bsa*HI to control for CTG repeat length. A fragment around 950 nt encompassing the repeats was obtained from the clone chosen, which corresponded to approximately 160-200 CTG repeats (for convenience we will refer to this clone as CTG_{197}).

ORFs of *Drosophila muscleblind* splice forms were amplified from full length cDNAs by high fidelity PCR (Pwo DNA polymerase) using isoform-specific primer combinations. A common 5' primer Xho-Mbl was used in combination with 3´ primers MblA-gc-E, MblB-gc-E, MblC-gc-E, or MblD-gc-E, to amplify the corresponding ORFs. PCR products were purified, digested, and cloned into the *Xho*I/*EcoR*I sites of pEGFP-N3 plasmid in frame with the downstream eGFP reporter gene. pEGFP-N3-MblC$^{\Delta SUMO}$ was available in our laboratory.

In order to facilitate future experiments, myc epitope was included in between Mbl proteins and the GFP tag. myc coding sequence, according to the human codon usage tables [103], was generated by annealing oligos myc-dir and myc-rev. pEGFP-N3 vector was *Sal*I/*BamH*I digested and

ligated to the annealed oligos. The presence of the epitope was detected by standard colony PCR [103]. This vector (referred to as pMV2) was *Xho*I/*Sal*I digested and ligated to MblA, B, C, D and $C^{\Delta SUMO}$ coding regions (*Xho*I/*Sal*I digested) obtained from pEGFP-N3 vector.

III.1.2. Cell transfections.

COS and HEK293T cells grown in Dulbecco´s modified Eagle´s medium (DMEM; GIBCO) supplemented with 10% FBS (Foetal Bovine Serum; GIBCO) were seeded in two-well dishes to a density of 80000 cells per well. The transfection mix contained 6 μl of Fugene (Roche), 2 μg of plasmid DNA, and 180 μl of DMEM. After 15 min at room temperature, cells were transfected with 1/6 of the transfection mixture.

In order to analyse the co-localisation of Muscleblind proteins with CUG repeat containing RNA, COS cell transfection mix included 1 μg of CTG_{197} plasmid and 1 μg of plasmid pEGFP-N3, which expressed Muscleblind protein isoforms fused to GFP. HEK293T cells were transfected with a mix containing 300 ng of GFP-tagged Mbl protein plasmid, 300 ng of $C\underline{T}G_{197}$ and 1.4 μg of carrier DNA.

III.1.3. Protein and RNA sub-cellular localisation detection.

Fluorescent *in situ* hybridization (FISH) with Cy3 labelled $(CAG)_{10}$ probes (Operon Technologies) was performed as described previously [67] except that salmon sperm was omitted from the hybridization solution. Muscleblind proteins fused to GFP were detected by GFP fluorescence.

Cells were analyzed under epifluorescence microscopy using a Zeiss Axioskop2 mot plus microscope.

III.2. *α-actinin* minigene splicing assay.

III.2.1. Cloning.

Construction of α-*actinin* minigene for alternative splicing assays was performed as follows: DNA encompassing *Drosophila* α-*actinin* cassette exons 6, 7 and 8 was amplified from wild type genomic DNA using the Triple Master kit and oligonucleotide primers 5ACTMGE5 and 3ACTMGE9. The single PCR product was purified with PCR-product cleaning kit (Qiagen) and digested with *Eco*RI. pSG5 vector (Stratagene) was linearized with *EcoR*I. Digested vector and PCR product were phenol/chloroform extracted, quantified by agarose gel electrophoresis and mixed in a standard ligation reaction.

In order to generate GFP fusions of Bruno proteins, *bruno* ORFs from cDNAs LD29068, LD19052 and LD31834 were amplified by high fidelity PCR using primers carrying adaptors for enzymatic restriction. Xho29/ 29Eco and Xho31/31Eco primer combinations amplified *bru1* and *bru3* ORFs, which additionally introduced XhoI and EcoRI restriction sites. Purified and digested PCR products were cloned into *Xho*I/*Eco*RI digested pEGFP-N3. PCR product obtained with Xho19 and 19Bam primer pair was *Xho*I/*Bam*HI digested to clone *bru2* ORF in pEGFP-N3 digested with the same enzymes.

myc tagged Muscleblind and MblC$^{\Delta SUMO}$ proteins were also generated to perform this assay. The myc epitope was generated by annealing oligonucleotides S-myc-Nf and S-myc-Nr. The hybridization left 3' and 5' cohesive ends that were ligated to purified *Sal*I/ *Not*I digested pEGFP-N3/MblC (see section II.3.1.1.) in order to replace the GFP-encoding DNA with the myc epitope. We refer to this derivative as pMV1 vector. pMV1 carrying *mblC* ORF was digested with *Eco*RI/*Xho*I to release the *mblC* ORF and the myc-encoding vector was purified. The ORFs of the remaining isoforms were obtained from GFP-tagged constructs in pEGFP-N3 (see section II.3.1.1.) by *Eco*RI/*Xho*I digestion and ligation to pMV1. GFP tagged MBNL1 was obtained from Prof. Maurice S. Swanson (Dept. of Molecular Genetics and Microbiology, University of Florida, USA).

myc-tagged Muscleblind proteins were *Bgl*II and *Not*I digested from pMV1 vector and cloned into *Bam*HI/*Not*I digested pIEI4 for protein expression in *Drosophila* S2 cells. Notice that *Bam*HI site is not recovered after ligation.

III.2.2. Cell culture and transfection of cell lines.

HEK293T cells were grown to 40-60% confluency in DMEM. HEK293T cells were co-transfected with 2 µg of *Drosophila* α*-actinin* minigene and 100 ng of pEGFP-N3 expressing Mbl isoforms, or pMV1 vector, using 2 µl of Lipofectamine (Invitrogen) per well and 200 µl Optimem media (Life Technologies). Four hours after transfection the media was changed to antibiotic-free DMEM media supplemented with 10% FBS.

For the analysis of the co-localisation with the expanded CUG repeat containing RNA expression, HEK293T cells were co-transfected with 0.5 µg

of *Drosophila* α*-actinin* minigene, 0.5 μg of CUG_{197} and 0.5 μg of pEGFP-N3 (control) or GFP-tagged Muscleblind protein constructs. Empty pGEM42 vector was used to reach a final 1.5 μg of transfected DNA whenever necessary. 1.5 μl of Lipofectamine (Invitrogen) were used per well.

Drosophila S2 cells were grown in Schneider's medium supplemented with 10% FBS and 0.1% Streptomycin and Penicillin. For transfection, cells were seeded with a density of 1.8* 10^6 cells/ well in serum-free medium. 24 h after seeding, cells were transfected with 2 μg of *Drosophila* α*-actinin* minigene and 100 ng of Mbl isoforms in pMV1 vector using 8 μl of Cellfectine reagent (Invitrogen).

For Bruno protein activities assessment, COS cells grown in DMEM with 10 % FBS were transfected with 0.5 μg of α*-actinin* minigene and 0.25 μg of GFP-tagged proteins.

III.2.3. RNA and protein extraction.

Total RNA was extracted as described in section II.1.4.2. Proteins were extracted using 150 μl of lysis buffer ([50 mM TrisHCl pH 7.1; 150 mM NaCl] + 0.5% triton X + freshly added protein inhibitors) per well. Protein quantification was performed with coomasie blue reagent (PIERCE). To check protein expression, 50 μg of total protein extract were run in 12.5% SDS-PAGE, blotted, and immunodetected with anti-GFP (mouse; Roche) or anti-myc (mouse; Roche) antibodies following standard procedures [103].

Total protein from HEK cells co-expressing GFP-tagged Muscleblind proteins and CUG repeat RNA was extracted with 150 μl of hot 2xSDS loading buffer with freshly added 10% β-mercaptoethanol. Two cycles of hot (>95ºC) and dry ice incubation were made to lysate cells. 20 μl of protein extract were loaded on a 15 % SDS-PAGE gel and scanned for densitometry. Equal amounts of protein were then loaded on a 15 % SDS-PAGE gel and treated as described in the previous paragraph.

III.2.4. RT-PCR.

5 μg of total RNA were used in an RT reaction performed as described in section I.4.4. To detect transcripts arising from the *Drosophila* α-*actinin* minigene we used 4 μl of cDNA (7 μl when working with S2 cells) in combination with primers 1938 and 1956 in a 40-cycle PCR reaction (for primer sequences and annealing temperatures see table M.2). PCR products were purified by NH_4Ac precipitation and half of the final volume was digested with *Sac*I to differentiate α-*actinin* isoforms A and B, which otherwise have the same size. The remaining PCR products and entire digestions were loaded in a 2% agarose gel. 1DAdvanced Phoretix software from Nonlinear dynamics was used to quantify the bands.

Alternatively, PCR products were radiolabelled by including 32.5 μCi of (α^{32}P)-dCTP (PerkinElmer Life Sciences) in the 27-cycle PCR. Amplification used the same primer combination as above. PCR products were purified by PCR-product cleaning kit (Qiagen) and half of the final volume was digested with *Sac*I. Remaining PCR products and their corresponding digestions were resolved in an 8% polyacrylamide gel and

visualized by autoradiography using Biomax MS film (Eastman Kodak). Bands were quantified with a Molecular Dynamics Phosphorimager.

III.3. mouse *Tnnt3* minigene splicing assay in human cells.

HEK293T cells grown and seed as above (section II.3.2.2.), were transfected with 0.5 μg of mouse *Tnnt3* minigene and 0.5 μg of GFP-tagged Muscleblind and Bruno proteins, or empty pEGFP-C1 vector, using 1 μl of Lipofectamine in presence of Optimem Media. Proteins were extracted with 2xSDS loading buffer (see section II.3.2.3) and total RNA was extracted with Tri reagent (see section II.1.4.2). As *Tnnt3* minigene is also in pSG5 vector, detection of minigene products was performed with non-radioactive PCR as described in section II.3.2.4.

III.4. Cell death assay.

III.4.1. Cell culture and transfection of S2 cells.

Drosophila S2 cells were grown in Schneider's medium with 10% FBS in 75 cm^2 flasks. 24 h prior to transfection, cells were seeded to 4.5 * 10^5 cells/ well in 24-well plates.

Prior to transfection, two 1.5 ml sterile tubes were UV irradiated per sample; the tube containing the expression vectors will be called A and the one containing the transfection reagent, B.

1) Prepare transfection mixture by mixing 60 μl of Schneider's serum free medium and DNA (tube A) and 60 μl of medium and Cellfectin (tube B) for each DNA sample (will be applied to 2 wells). The amounts of DNA used are described in detail in table M.4 (pIEI-Diap1 and pIEI-LacZ were

obtained from Dr. Lei Zhou, Dept. Molecular Genetics and Microbiology, University of Florida).

- Dilute 0.1 μg of pIEI-LacZ per sample with serum-free medium in tube A
- Add appropriate amount of DNA sample to each tube A.
- Vortex the cellfectin tube briefly before pipetting. Mix 8 μl of Cellfectin per sample with the x-free medium needed for tubes B.
- Pipet 60 μl from tube B to each A tube. Mix by tapping the bottom of the tube and let the mixture stand at room temperature for 15-30 min.

A)	vector	MblA	MblB	MblC	MblD	MblC$^{\Delta sumo}$
Sample	0.9 μg	0.9 μg	0.9 μg	0.9 μg	0.9 μg	0.9 μg
Total DNA	**1 μg**	**1 μg**	**1 μg**	**1 μg**	**1 μg**	**1 μg**
B)	vector	MblB	MblD	MblB + Diap	MblB + MblD	MblB + caspase inhibitor
Sample	0.3 μg	0.3 μg	0.9 μg	0.3 μg MblB 0.6 μg Diap	0.3 μg MblB 0.6 MblD	0.3 μl MblB
Total DNA	**0.4 μg**	**0.4 μg**	**1 μg**	**1 μg**	**1 μg**	**0.4 μg**
C)	vector	MblB	MblB + casp inh			
Sample	**0.9 μg**	**0.9 μg**	**0.9 μg**			
Total DNA	**1 μg**	**1 μg**	**1 μg**			
Sample	MblB + vector	Diap + vector	MblD + vector	MblB + MblD	MblB + caspase inhibitor	
Sample	0.3 μg MblB 0.6 μg vect	0.6 μg Diap 0.3 μg vect	0.6 μg MblD 0.3 μg vect	0.3 μg MblB 0.6 μg Diap	0.9 μg MblB	
Total DNA	**1 μg**	**1 μg**	**1 μg**	**1 μg**	**1 μg**	

Table M.4. Amounts of DNA used in the different cell death experiments. A) Capacity of Mbl isoforms to induce cell death was evaluated transfecting 0.9 μg of expression vector per each two wells **B, C)** Experiments to test MblB interaction with apoptosis pathway. In the first experiment, the total amount of DNA was not the same for all samples **(B)**. The second one **(C)**, with the same total DNA amounts, gave an unmanageable number of cells per well.

2) Starting 15 min after pipetting B into A, aspirate the medium from the plates, wash once with serum-free medium and aspirate off again.
3) Add 480 μl serum-free medium to each sample tube. Mix by pipetting up and down. Apply 300 μl to each 24-well (2 well per sample).
4) After 5 h at room temperature, aspirate off the transfection mixture and replace with 0.5 ml of medium supplemented with 10% FBS.

III.4.2. Cell viability measurements.

Cells were fixed in 1% paraformaldehyde in PBS, washed twice with PBS and stained for *lac-Z* expression by standard procedures [103]. Number of beta-galactosidase-positive cells (blue) along the diameter of the well was counted under bright field microscopy.

IV. *in vitro* techniques.

IV.1. *in vitro* UV crosslinking.

IV.1.1. Cloning.

myc-tagged Mbl isoforms were *Bam*HI/*Not*I digested from pMV vector and ligated to pGEX-6P-1 digested with *Bgl*II and *Not*I. This introduced a one nucleotide frameshift that was corrected by site-directed mutagenesis (QuikChange Site-Directed Mutagenesis kit from Stratagene). This correction shifts myc epitope out of frame. *Bam*HI site was also reconstituted by site-directed mutagenesis in the same reaction. MblB and MblC fusions to GST were confirmed by sequencing. No positive clones were obtained for MblA and MblD. Since work carried out in parallel with

human MBNL proteins revealed protein solubility and stability problems, we decided to tag *Drosophila* proteins at both ends.

MblA ORF in pFP105 (modified from pFP98), double tagged with GST at 5' end and His_6 at 3' end, was obtained from N. Paricio [44]. 5' primer Nde-Mbl was used in combination with 3' primers MblB-Eco, MblC-Eco, MblD-Eco and ZF-Eco to amplify MblB, MblC and MblD ORFs, and the fragment containing the zinc fingers (ZF) by high fidelity PCR (Pfu DNA polymerase; Promega). PCR products and pFP105/MblA were digested with *Nde*I and *EcoR*I and ligated. MblC sequence showed no mutation. 3' end of MblB could not be confirmed by sequencing. Double tagged MblD and ZF could not be obtained following this strategy.

IV.1.2. Protein expression and extraction.

BL21, BL21star, and Rosetta *E.coli* cells were used to express MblA construct in presence of 0.1 mM IPTG. Influence of temperature during growth, and expression induction, in protein solubility was tested.

MblA and MblC small scale extractions were made by resuspending the pellets with Novagen Bugbuster followed by addition of Novagen Benzonase nuclease. Large scale extracts were lysated using a French press in MTPBS or SB buffer in presence of protease inhibitors (PMSF and Roche Complete EDTA-free protease inhibitor cocktail tablet). Insoluble fraction was extracted with either MTPBS or SB buffer adding 1 M NDSB.

MTPBS: 150 mM NaCl; 16 mM Na_2HPO_4; 4 mM $NaH_2PO_4 \cdot 2H_2O$ + freshly added 1 mM DTT

SB buffer: 50 mM $NaPO_4$ pH 8; 300 mM NaCl; 5% glycerol

IV.1.3. Protein purification.

Proteins were affinity purified using GST, His_6 or both tags. Once lysates were recovered from the French press, all the following steps were made in the cold room (4ºC) to prevent protein degradation. Samples from all extraction steps were loaded on 15 % SDS-PAGE protein gels to check the efficiency and to estimate the amount of protein present, and the volume of beads to be used in the purification procedure (referred to as column volume). Extracts from two 400 ml cultures were purified together.

Appropriate volume (following manufacturer recommendations) of the 50% slurry preparation of Glutathione Sepharose 4B beads (GE Healthcare) was washed with H_2O, MTPBS and finally MTPBS + 1 % Triton-X100. 1 % Triton-X100 was also added to the protein extract before loading it to the beads and the mixture was incubated 1 to 3 h at 4ºC with rotation. Beads were transferred to hydrophobic 1.7 ml tubes as soon as the volume fitted to prevent proteins from sticking to the walls. Four washes (resuspension and 1'-1000 rpm spin) with 2.5 column volumes of MTPBS/Triton were made. Five elutions (resuspension, 5 min rotation and 1'-1000 rpm spin) with 1 column volume were performed with 50 mM Glutathion in 100 mM Hepes pH 8. To ensure the recovery of the remaining protein, a last elution was performed with 2 column volumes. When His purification followed GST purification, final elutions were diluted with 2x SB buffer prior to loading into Ni-NTA beads (Qiagen).

Extracts in SB buffer coming from French press or GST purification were loaded in the appropriate volume of washed (water and SB buffer) Ni-NTA beads. After 1- 2 h rotation at 4ºC, matrix was spinned and washed

(resuspend + 1'-1000 rpm spin) four times with SB buffer. Six elution steps were performed using one column volume of 500 mM imidazole in SB buffer.

After purification, proteins were dialysed o/n at 4ºC against buffer E (20 mM Hepes pH 7.9; 100 mM KCl; 0.2 mM EDTA; 0.5 mM DTT; 10 % glycerol). As the concentration was very low, I used PD-10 columns (Amersham) in subsequent experiments in order to obtain a protein solution in buffer E. Samples from the final fractions were run in 15% polyacrylamide gel with SDS and the fractions showing the highest concentrations, and low degradation, were pooled together, aliquoted and frozen in dry ice.

IV.1.4. PCR template for *in vitro* transcription.

5' primers were designed to include a T7 transcription promoter sequence, plus a few extra nucleotides, to facilitate the binding of T7 RNA polymerase. Primers T7actn1D/R amplified Actn1; primers T7actn2D/R gave Actn2 PCR product; oligos T7actn1D/R and T7actn6D/R were used to amplify the entire intron 6 and 7, respectively; T7CUGD/R and T7CAGD/R were used to amplify CTG and CAG repeat sequences, respectively. PCR amplification failed with both CTG and CAG trinucleotide repeat templates.
For each primer combination, two 50 μl PCR reactions using Pfu DNA polymerase were set up in thin-walled tubes following instructions from the manufacturer.

IV.1.5. *in vitro* transcription.

The transcription reaction was set up as follows (assembled at room temperature to avoid DNA precipitation):

- 5x Transcription Buffer 2 µl (0.2 M Tris pH 7.5; 30 mM $MgCl_2$; 10 mM Spermidine)
- 100 mM DTT 1 µl
- 10x rNTP 1 µl (5 mM CTP; 5 mM ATP; 5 mM GTP; 0.4 mM UTP)
- H_2O 1 µl
- RNasin 0.5 µl
- 10 mM cap 0.5 µl
- α^{32}P UTP 2 µl
- 1 mg/ml linear DNA 1 µl
- T7 RNA polymerase 1 µl

- Incubate 1 – 2 hours at 40˚C
- Add 40 µl of H_2O, take 1 µl and spot onto glass fibre disc for total counts
- Phenol extract (use water-saturated phenol)
- Load phenol-extracted RNA onto a pre-washed G-50 column and spin. Save eluted liquid, measure volume and transfer to an eppendorf tube. Take 1 µl of the elution and perform trichloroacetic acid precipitation to measure the incorporated radioactivity.
- Adjust concentration of in vitro transcribed RNA to 20 fmol/µl

Transcription of intron 7 fragment was inadequate for subsequent experimental procedures. RNA encompassing DNA fragments Actn1 and Actn2 and the whole intron 6, on the other hand, were transcribed efficiently.

Competitor RNAs were generated from large scale *in vitro* transcription reactions as follows:

- 5x Transcription Buffer 10 μl
- 1 M $MgCl_2$ 2 μl
- 100 mM DTT 5 μl
- H_2O 0.75 μl
- RNasin 2 μl
- 10 mM cap 2.5 μl
- 100 mM ATP 5 μl
- 100 mM GTP 5 μl
- 100 mM CTP 5 μl
- 100 mM UTP 5 μl
- α^{32}P UTP 0.25 μl
- 1 mg/ml linear DNA 5 μl
- T7 RNA polymerase 2.5 μl

The procedure was the same as above but 50 ml of 10 mM Tris-10 mM EDTA was added instead of H_2O prior to phenol extraction.

IV.1.6. UV crosslinking.

The following 10 μl reaction was set up in a microtitre plate (Falcon 353911) (all reagents are included in this volume except heparin; when possible, the plate was kept on ice to avoid protein degradation):

- 20 fmol *in vitro* transcribed RNA
- 2 μl 5x Binding Buffer (50 mM Hepes pH 7- 7.2; 15 mM $MgCl_2$; 25% glycerol; 5 mM DTT)
- 1 μl 0.5 - 2 mg/ ml rRNA made in H_2O

- 1 M KCl up to 100 mM final concentration
- 0.5 μg of BSA
- protein extract up to 10 μl

The amount of protein extract added to the reaction varied between 2 - 6 μl depending on the volume left by the other reagents. To test the specificity of the reaction, we titrated with increasing amounts of rRNA in the reaction. For the competition assay, 200, 300 and 400 molar excess of competitor RNA was added to the reaction mix.

- Wrap plate in plastic film and incubate at 30°C for 30 min. To increase the stringency of the assay and test the specificity, add 1 μl 6.81 –55 mg/ml heparin 5 min before the end of the incubation step.
- UV crosslink on ice. Use a Stratalinker 2x 960 mJ placing the microtitre plate on a tray with ice on the floor of the Stratalinker.
- Add 4 μl per well of RNase mix per well, wrap plate in plastic film and incubate for a further 15 min at 37ºC.

  RNase Mix (for 500 μl):

  50 μl 10x RNase dilution buffer (100mM Tris-HCl pH7.5; 10mM $MgCl_2$; 1M KCl)

  50 μl 10 mg/ml RNase A

  14 μl 100 u/μl RNase T1

- Add 4 μl of 4x SDS sample loading buffer with bromophenol blue per well and heat to 80ºC for 5 min including protein markers.
- Load onto 15% protein gel and run until all the bromophenol blue has run out.
- Dry the gel and expose film or phosphorimager screen.

Results

Por más que apagues las otras
velas, la tuya no brillará más

-Gandhi-

I. Evaluation of viable *muscleblind* mutants and flies expressing CUG repeat RNA as myotonic dystrophy fly models.

Human Muscleblind proteins (MBNL1-3) have been described as key factors in DM pathogenesis. The RNA gain of function hypothesis explains defining symptoms of Myotonic Dystrophy by sequestration of nuclear factors, mainly human MBNL. Mutant *DMPK* transcripts sequester MBNL proteins in ribonuclear foci preventing a significant fraction of them from raising their functional localisation [51]. *MBNL* loss-of-function phenotypes are therefore predicted to include DM-like defects. This is the case of *Mbnl1* knockdown mice, which exhibit the main symptoms of the disease [28]. *MBNL* genes are homologues to the *Drosophila muscleblind* (*mbl*) gene and it has been previously shown that *Drosophila* and human MBNL1 proteins are functionally exchangeable [107]. If the pathway of pathogenic RNAs is conserved in flies, and the function of Muscleblind proteins impaired in DM is conserved between flies and humans, one expects to find DM-like defects in *muscleblind* mutants and in flies expressing CUG repeat containing-RNA. Since DM patients show muscle histopathology including absence of Z bands [108], *Drosophila muscleblind* muscular defects (absence of I and Z bands and reduction of extracellular matrix) are in concordance with the RNA gain of function hypothesis [41],

Different *muscleblind* loss-of-function flies and flies expressing expanded CUG repeat containing RNA, which are genetically sensitive to *muscleblind* function [85], were available in our laboratory. My interest in this point was to validate these flies as models to study DM pathogenesis, showing the similarities between the fly models and the individuals affected by the disease. *muscleblind* is essential for *Drosophila* development and mbl^{E27}

and *mbl*^{k7103} *muscleblind* alleles are lethal in homozygosis. However, *mbl*^{E27}/*mbl*^{k7103} was reported as a hypomorph allelic combination that allows some individuals to reach adulthood [109]. *mbl*^{E27} is lethal in combination with *mbl*^{E16}, another null allele of *muscleblind*, whereas *mbl*^{E16}/*mbl*^{k7103} also behaves as a hypomorph allelic combination, thus suggesting that *mbl*^{k7103} is a hypomorph allele itself but it has accumulated a lethal mutation in the chromosome.

In order to find DM-like defects in muscleblind hypomorph mutants and flies expressing expanded CUG containing RNA, we characterised these flies at different levels. In a external analysis, we detected a reduced viability and defects in wing and leg development in mutant flies and a shortened lifespan in flies expressing the expanded RNA. We did not detect any significant effect of the anti-myotonic drug mexiletine on viability, lifespan or climbing activity of mutant flies, which we used as an indirect test for the existence of one of the main symptoms of DM, the myotonia. Finally, we analyse these flies at a molecular level looking for defects in the regulation of alternative splicing. We found interesting defects in specific transcripts.

I.1. Adult phenotype of *mbl* ^{E27}/*mbl* ^{k7103} mutants.

An initial quantification of the number of escapers resulting from crossing *mbl*^{k7103}/*CyO, ubi-GFP* with *mbl*^{E27}/*CyO, ubi-GFP* maintained at 25 ºC gave 52% of the expected adults (considering the 33% of the total offspring expected to be homozygous mutant as 100%). In a second counting, a more accurate scoring of the number of viable adult flies (scoring emerged adults every few hours to avoid underestimating the transheterozygous class), gave a percentage of around 70% ranging

between 62.9 and 79.3 % when the females in the cross were heterozygous for mbl^{E27} or mbl^{k7103}, respectively.

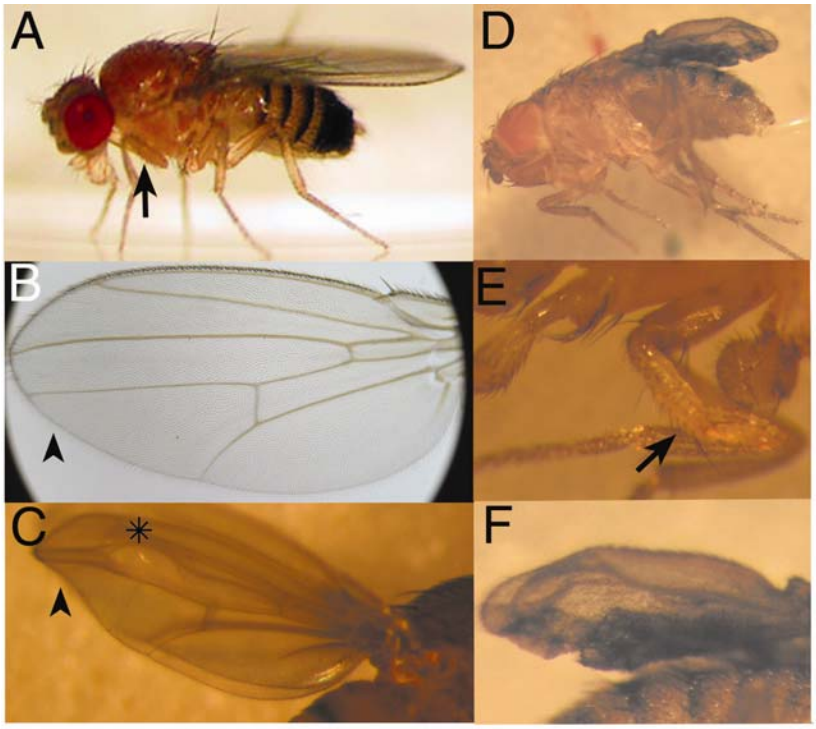


Figure R.1. mbl^{E27}/mbl^{k7103} flies show wing defects and leg deformation. A, **B)** Wild type fly and wing to compare with mutant phenotype. **C)** mbl^{E27}/mbl^{k7103} class II flies showed mild defects like small wing blisters (asterisk) and lack of tissue in the posterior wing compartment (arrowhead) **D)** Example of class I acute phenotype with twisted femur (magnified in **(E)**; arrow) and unexpanded wings in which the two lamina do not adhere to each other normaly (magnified in **(F)**).

mbl^{E27}/mbl^{k7103} adult flies showed a variable phenotype. Most of them showed no apparent morphological defects, others had only mild wing defects (Fig. R.1A-C), and some presented an acute phenotype with severe

wing defects and leg deformations (Fig. R.1A, D-F). Transheterozygous mutant offspring were divided into four phenotypic classes:

- class 0 included those flies that gave no viable adults or died in the first hours of adult stage;
- class I included flies with acute phenotype;
- class II, those individuals with mild wing defects;
- class III, individuals with no apparent external defects.

I.2. Lifespan of flies over-expressing CUG repeat RNA.

Alterations in the insulin pathway resulting in lower insulin signalling have been related to an increased lifespan in *Drosophila*, *C. elegans* and mouse [110, 111]. Alternative splicing of human insulin receptor (*IR*) pre-mRNA gives rise to two receptors with different insulin responsiveness. Transcripts present in those tissues responsible for glucose homeostasis (e.g. muscle) include exon 11 (E11; *IR-B*) and code for a receptor with less affinity for insulin but with higher signalling activity, whereas transcripts expressed in foetal tissues lack E11 (*IR-A*) [112]. DM1 skeletal muscle presents an abnormal abundance of the *IR-A* isoform [80], which causes the insulin resistance found in patients. Although insulin receptor structure and processing in *Drosophila* are different to those described in humans [113], flies expressing 162 CUG repeat RNA presented a mean lifespan that was 15% longer than controls [83], suggesting that the general insulin signalling alteration could be conserved despite the different molecular mechanisms that regulate IR activity. Thus we analysed flies expressing 480 CUG repeat RNA in order to see if longevity was altered by CUG repeat containing RNA.

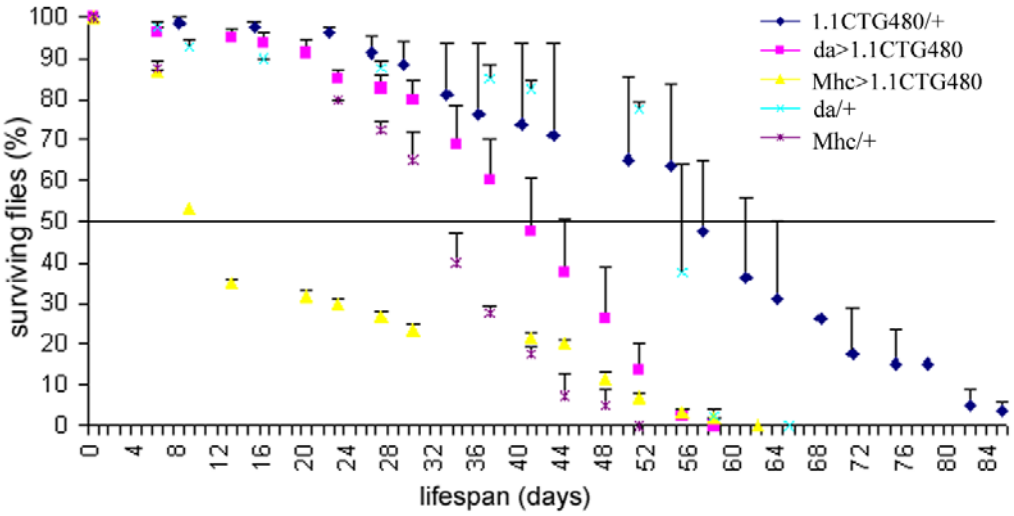


Figure R.2. Flies expressing CUG repeats show shorter lifespan. Average percentages of live flies along the days of the different strains used are represented. Transgene *UAS-CTG$_{480}$ 1.1* (1.1CTG480) was expressed with both *da-gal4* (*da*; n. replicas = 4) and *Mhc-gal4* (*Mhc*; n. replicas: 3) drivers. UAS transgene and Gal4 driver strains were crossed to *white* flies to generate control flies 1.1CTG480/+ (n. replicas = 4), da/+ (n. replicas = 4); Mhc/+ (n. replicas = 2). CUG repeat expressing flies showed reduced longevity when compared with their respective controls.

Contrary to previous observations [83], a significant decrement of mean lifespan was observed when expressing CUG repeat RNA either in a muscular (*Myosin Heavy Chain*; *Mhc*) or general embryonic pattern (*daughterless; da*) with the Gal4/UAS system compared to control flies carrying the UAS or the Gal4 transgene (Fig. R.2). The average percentage of live flies was reduced to 50% after ten days when expressing repeats in *Mhc* pattern, whereas in flies carrying the UAS transgene it was after 58 days and in the *Mhc* Gal4 driver strain, after 32 days. When expressing repeats in a general embryonic pattern, flies showed a milder reduction in lifespan as 50% of flies died after 41 days compared to the 54 days for the *da* driver strain and the 58 days of the transgene strain. Flies expressing

480 CUG repeats show phenotypic defects as muscle degeneration and flightlessness [85], which could have an effect on fly fitness, reducing longevity. Flies expressing 162 CUG repeat RNA do not show any phenotypic defects, so, no reduction in general fitness could allow these flies to reflect other aspects affected by CUG repeat expression. No further studies have been carried out in these flies to elucidate if their extended lifespan is related to insulin pathway.

I.3. Effect of anti-myotonic drug mexiletine in *mbl*^{E27}/*mbl*^{k7103} mutants.

Myotonia is one of the main symptoms of DM. Myotonia consists of the hyper excitability of the sarcolemma and delayed muscle relaxation after voluntary contraction, resulting from the defective function of voltage-gated sodium or chloride channels. In the case of DM, the origin of the myotonia is a severe reduction in CLC1 protein on the myocyte surface of DM patients [79, 114]. *Mbnl1* knockdown mice showed myotonic electromyographic recordings [28]

Electromyographic recordings are very difficult to obtain in *Drosophila* and we could not find a laboratory to do either whole fly electromyography or cell patch-clamp recordings to directly analyse cell membrane conductance to chloride ions. Because of these technical limitations, we decided to analyse whether the anti-myotonic drug mexiletine used in DM patients had any effect on *muscleblind* mutant flies. Mexiletine is a lidocaine derivative that blocks voltage-gated sodium channels [115]. If *muscleblind* mutant flies had myotonia similar to DM patients, they could also respond to mexiletine.

We performed preliminary study of the influence of the anti-myotonic drug on the viability, lifespan and climbing ability of *mbl*^{E27}/*mbl*^{k7103} flies.

I.3.1. **Viability analysis.**

Number of viable flies emerged when crossing *mbl*^{E27}/*CyO, ubi-GFP* and *mbl*^{k7103}/*CyO, ubi-GFP* flies in presence of excipient or 1x or 10x dose of mexiletine was analysed (Fig. R3A). Wild type *OrR* flies were exposed to the drug in order to control drug toxicity. No increase in the average of total number of viable flies was detected at any mexiletine dose. Contrary, a certain susceptibility to the drug was observed when the females crossed carried the *mbl*^{E27} allele although no toxicity was detected in *OrR* flies. The viability of heterozygous flies was used as an internal control for mexiletine toxicity (Fig. R.3B). A significantly reduced number of heterozygous flies emerged at 1x mexiletine in crosses with *mbl*^{k7103} females compared to the crosses made on medium containing excipient. No difference was found for crosses made with *mbl*^{E27} females. This suggests that the previous reduction in the total number of flies emerging from crosses with females heterozygous for *mbl*^{E27} was due to a decrease in mutant viability and not to a general reduction of fly viability under those conditions.

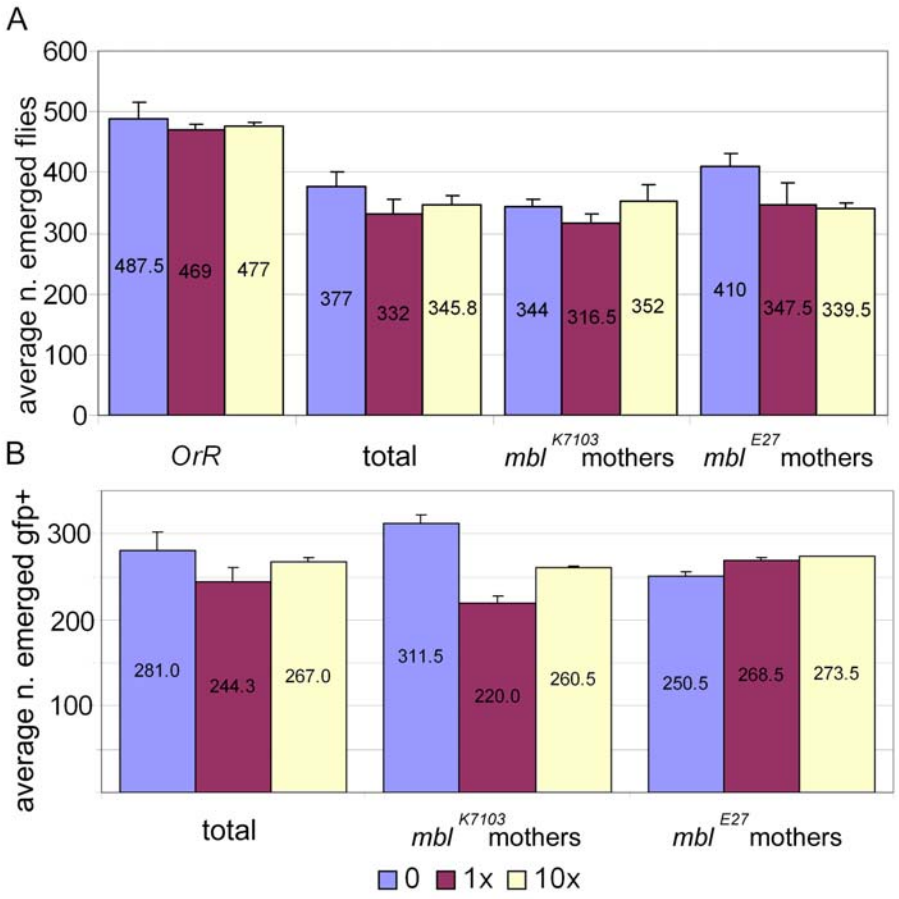


Figure R.3. Controls of mexiletine toxicity. A) Average number of flies that emerged from crosses made in presence of 0, 1x and 10x dose of mexiletine (bars) and standard deviations (error line) are shown. "total" correspond to the data from all the crosses between *mbl*^{E27}*/CyO, ubi-GFP* and *mbl*^{k7103}*/CyO, ubi-GFP* flies independently of the strain where females came from. Whereas no mexiletine dose altered the number of *OregonR* flies emerged, a reduction was found when crossing females heterozygous for *mbl*^{E27} with *mbl*^{k7103}*/CyO, ubi-GFP* males (1x dose: P-value = 0.09; 10x dose: P-value = 0.07). **B)** Average number of heterozygous flies (GFP-positive) emerging from crosses between *mbl*^{E27}*/CyO, ubi-GFP* and *mbl*^{k7103}*/CyO, ubi-GFP* flies (bars) and standard deviations (error line). 1x dose of mexiletine significantly reduced GFP-positive flies when heterozygous *mbl*^{k7103} flies were used as mothers compared to flies developing in food with excipient (P-value < 0.05).

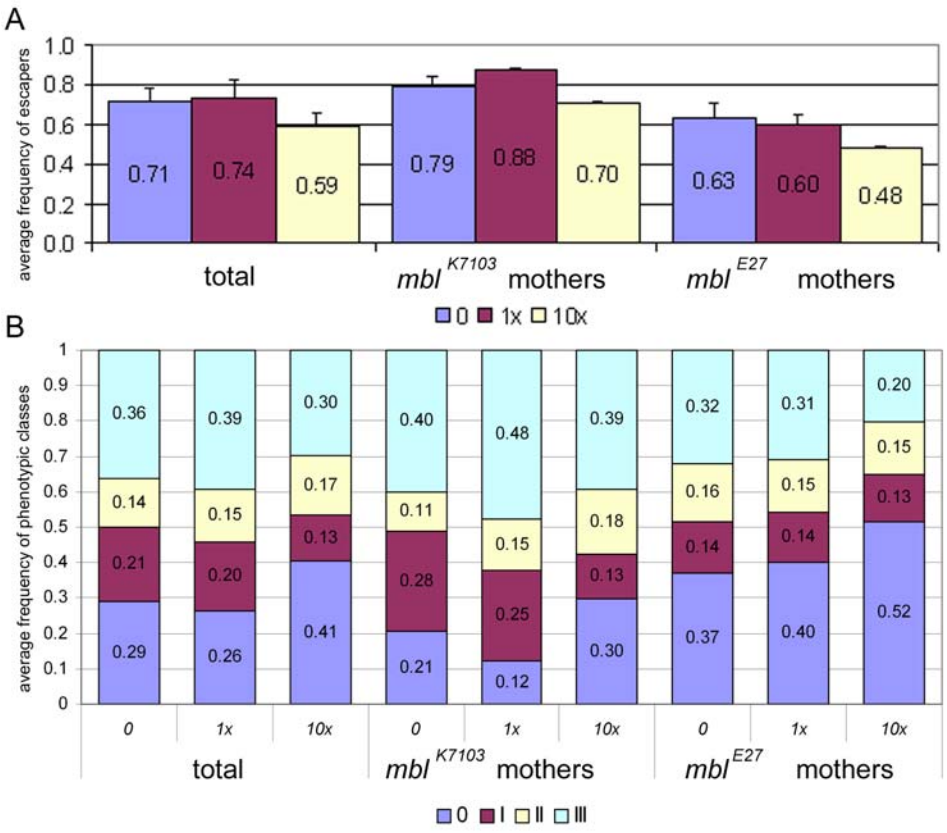


Fig. R.4. Frequency of *mbl^{E27}/mbl^{k7103}* viable flies was reduced by high dose of mexiletine. Average frequencies of mutant flies emerging from crosses between *mbl^{E27}/CyO, ubi-GFP and mbl^{k7103}/CyO, ubi-GFP flies* are represented. **A)** Average frequency of escapers (viable *mbl^{E27}/mbl^{k7103}* flies) emerging. Administration of 1x dose of mexiletine generates a marginal increment in viable mutant flies when mothers are heterozygous for *mbl^{k7103}* allele (see Fig. R.3.). Toxicity of mexiletine at 10x dose detected with the total numbers is mainly due to sensitivity of offspring from crosses made with *mbl^{E27}/CyO, ubi-GFP* mothers. **B)** Phenotypic class average frequencies. Increase of Class 0 in crosses treated with 10x mexiletine comes from the reduction of class I in crosses with *mbl^{k7103}/CyO, ubi-GFP* mothers and from class III in crosses with *mbl^{E27}/CyO, ubi-GFP* mothers.

When analysing the effect of mexiletine on *mbl*^{E27}/*mbl*^{k7103} viability, no increase in the proportion of viable mutant flies was observed at any concentration (Fig. R.4A). Actually, a 10x dose of mexiletine reduced the frequency of viable hypomorph flies compared to the same crosses carried out without the drug. Separation of the data into phenotypic classes (Fig. R.4B) showed that intermediate phenotypes remained similar to control crosses when heterozygous *mbl*^{E27} mothers were used. A general deterioration of flies is observed at 10x mexiletine dose compared to the crosses without drug. Contrarily, the frequency of class I was significantly decreased at 10x mexiletine dose in crosses with heterozygous *mbl*^{k7103} mothers but the other classes remained similar. So, the offspring of this cross with acute phenotype appeared to be more sensitive to the drug but those with mild or no defects did not present increased sensitivity to mexiletine. A marginal improvement on phenotypic class distribution was observed in crosses at 1x dose of mexiletine with *mbl*^{k7103}/*CyO, ubi-GFP* mothers but differences were not statistically significant (p-value > 0.2).

I.3.2. **Lifespan analysis.**

mbl^{E27}/*mbl*^{k7103} flies that emerged from the viability study were kept in vials with the same dose of anti-myotonic drug and their longevity recorded. The lifespan analysis showed no general improvement of mexiletine-treated flies at any concentration. However, a 1x mexiletine dose significantly increased the percentage of surviving flies between days 6 and 13. A slight increase in the survival of 1x mexiletine-treated flies persists after day 13 but was not statistically significant.

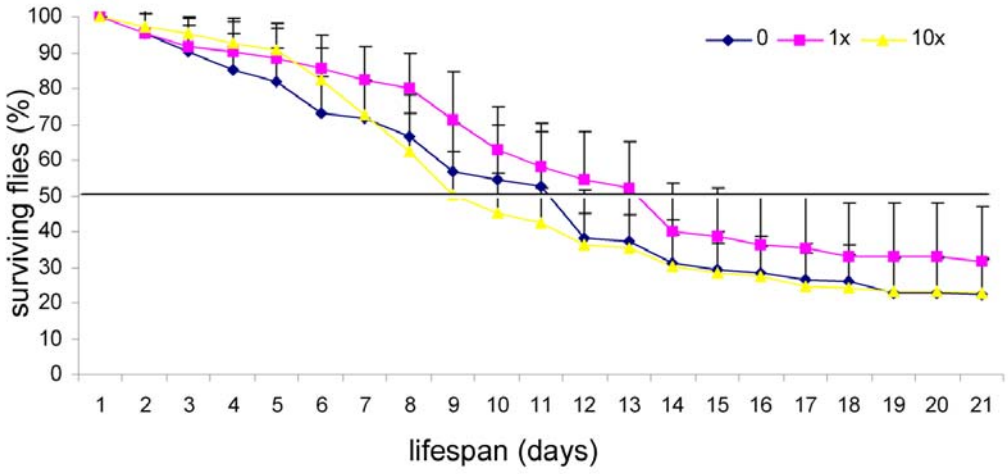


Fig. R.5. Average percentage of *mbl^{E27}/mbl^{k7103}* flies surviving marginally increases by mexiletine treatment. The average percentage of flies surviving at the ages indicated in presence of mexiletine concentrations 0, 1x and 10x, are represented together with standard deviation (bars). Differences were statistically significant between days 6 and 13 (days 6, 8 and 12, P-value <0.05; days 7, 9 and 13, P-value< 0.1).

When considering the data from crosses with *mbl^{E27}/CyO, ubi-GFP* and *mbl^{k7103}/CyO, ubi-GFP* females separately, we detected (Fig. R.6A) a general slight trend of increased survival of *mbl^{E27}/mbl^{k7103}* flies treated with 1x dose of mexiletine in the offspring of crosses where mothers carried the milder *mbl^{k7103}* allele. An increase in mutant survival between the fifth and 13th day was also observed when the mothers carried *mbl^{E27}* but the effect completely disappeared after the thirteenth day.

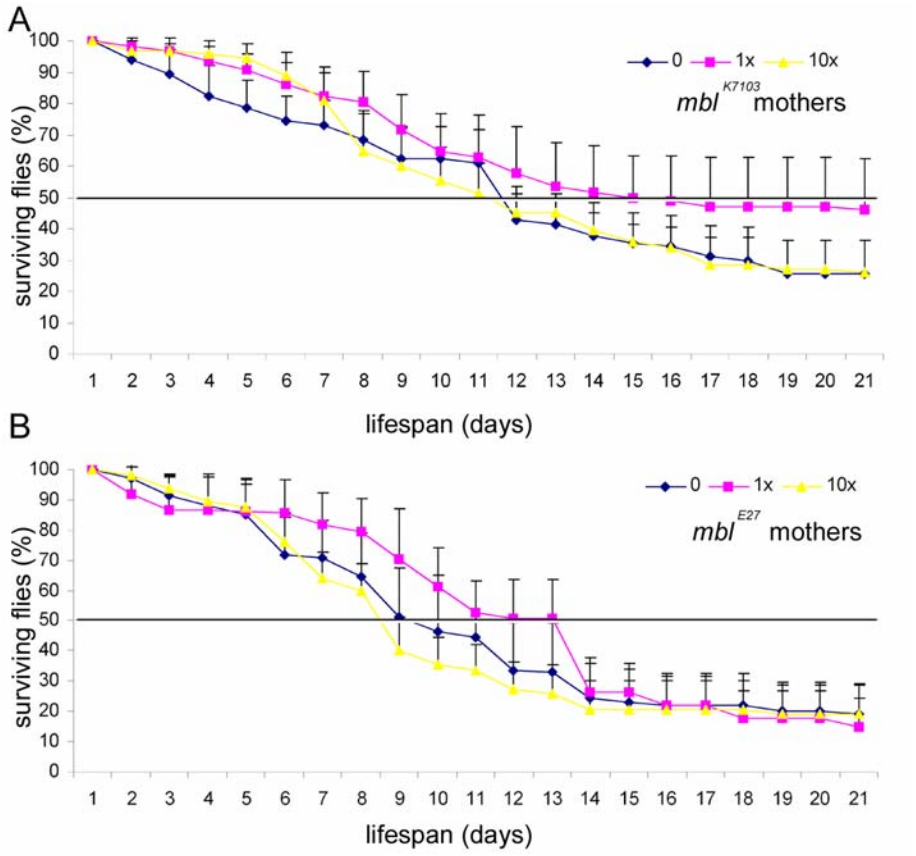


Fig. R.6. Survival of *mbl*^{E27}/*mbl*^{k7103} offspring depends on the genotype of the mothers. Average percentages of offspring surviving are represented versus time. Crosses were made in presence of different amounts of mexiletine (0, 1x, 10x). Number of flies in the assay (n) was, from crosses made with *mbl*^{k7103}*/CyO, ubi-GFP* mothers: 0: n = 169, 1x: n = 153, 10x: n = 111; and from crosses with *mbl*^{E27}*/CyO, ubi-GFP* mothers: 0: n = 90, 1x: n = 128, 10x: n = 95. **A)** Offspring of crosses using *mbl*^{k7103}*/CyO, ubi-GFP* females show a weak, although maintained, increase of survival when treated with 1x dose of mexiletine. Differences between *mbl*^{E27}/*mbl*^{k7103} alive flies in vials with 1x dose of mexiletine and excipient were significant at days 4, 5, 6, 19, 20 and 21 (P-value < 0.1). An initial significant improvement is also observed with 10x dose in the crosses with *mbl*^{k7103} females encompassing days 4 (P-value 0.06), 5 and 6 (both with P-value< 0.05) **B)** Apparent positive effect on survival of offspring of crosses made using *mbl*^{E27}*/CyO, ubi-GFP* females in presence of 1x dose of mexiletine is not statistically significant and is not maintained after day 13.

I.3.3. **Climbing ability analysis.**

Climbing assays have been extensively used in *Drosophila* to ascertain motor dysfunction [116]. Because we expected *muscleblind* mutant flies to develop myotonia, impairment of climbing activity in these flies was a possible phenotype. Therefore, the climbing activity of *muscleblind* mutant flies was studied in presence of mexiletine and compared to flies taking excipient as control. The data obtained were quite variable. Number of flies situated over the line (see M&M) gave no clear tendency (Fig. R.7) and no significant improvement was detected at a 1x concentration of mexiletine neither when treating all the data together (Fig. R.7A) nor when considering the different crosses separately (Fig. R.7B).

A small scale assay performed in a 2x concentration of mexiletine did not show any improvement. We also initiated a study administering 2x dose of mexiletine but insufficient flies were obtained to undergo an analysis similar to the previous with 1x. No trend for improvement was observed in the initial recordings.

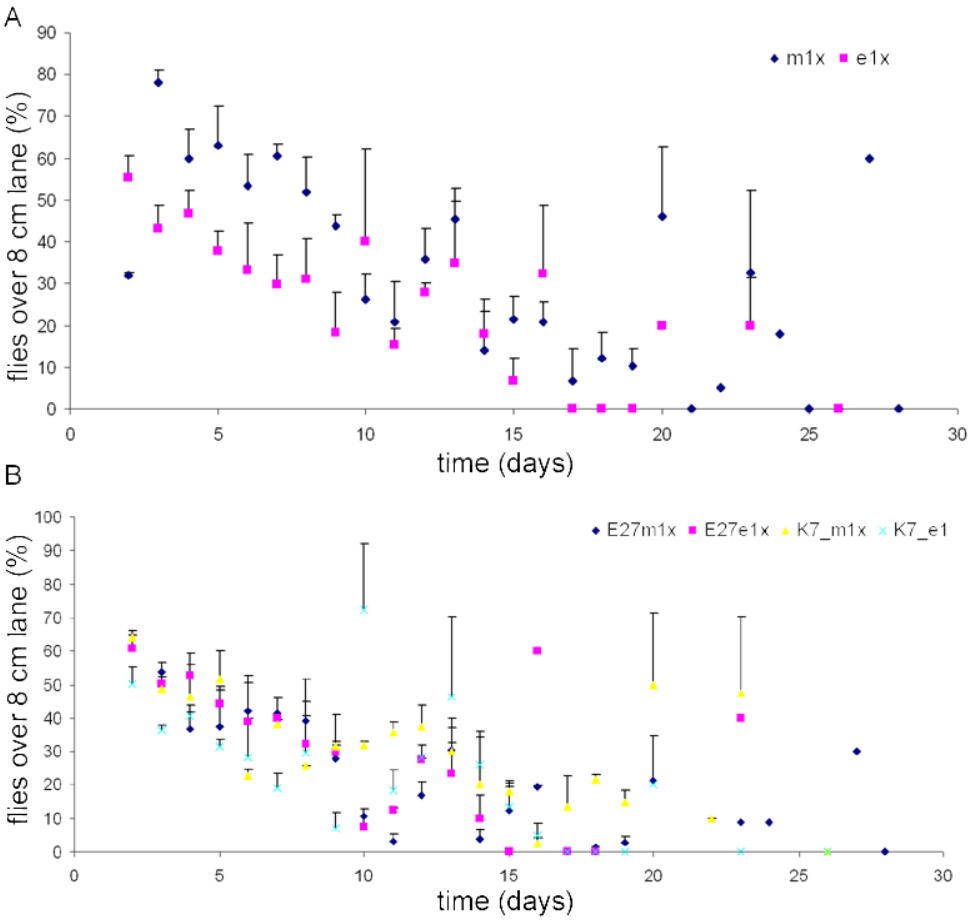


Fig. R.7. Treatment with mexiletine 1x dose did not improve climbing ability of *mbl^{E27}/mbl^{k7103}* flies. The average percentage for all replicas of each mexiletine or excipient concentration is represented. E27 and K7 indicate the genotype of the mothers (*mbl^{E27}/CyO, ubi-GFP* and *mbl^{k7103}/CyO, ubi-GFP* respectively); m and e designate mexiletine and excipient, respectively. Very dispersed data were obtained, with no clear effect of mexiletine on climbing ability.

In summary, no significant positive effect of mexiletine was detected either in survival, lifespan or climbing ability even though a small window of positive effect on mutant lifespan was detected, especially during the first

days of treatment. On the other hand, some sensitivity to the drug was detected as a reduced viability was found when the flies came from crosses with females carrying the strong allele *mbl*^{E27}.

I.4. Analysis of alternative splicing in *Drosophila* DM models.

Mbnl1 knockdown mice show disrupted alternative splicing regulation resulting in the generation of transcripts characteristic of immature cells in adults (Lin et al, 2006, [29, 39]). This foetal arrest has also been described in DM patients and mouse models of DM disease, and it has been shown that is MBNL1-dependent [29, 30]. Some of these splicing alterations have been directly related to DM symptoms [79, 80, 114]. Then, we decided to check *muscleblind* mutant flies and flies expressing CUG repeat RNA for miss-splicing events.

I.4.1. Chloride channel.

Myotonia is caused by a dramatic reduction of CLC1 protein in the membrane of DM1 myocytes. The origin of this defect is the alteration of *muscle specific chloride channel 1* (*CLC1*) splicing that leads to the inclusion of premature stop codons in the mature mRNA [79, 114] generally by maintaining a foetal splicing pattern in the adult tissue of DM1 and DM2 patients. Mice expressing expanded CUG repeat containing-RNA in skeletal muscle (*HSA*^{LR}) present the same splicing alteration and myotonic electromyographic recordings [69, 114]. Flies over-expressing 480 CUG repeats show muscle and behavioural alterations that could reflect a

myotonic condition [85]. Thus, we decided to check if the alteration of chloride channel transcript splicing was conserved in flies.

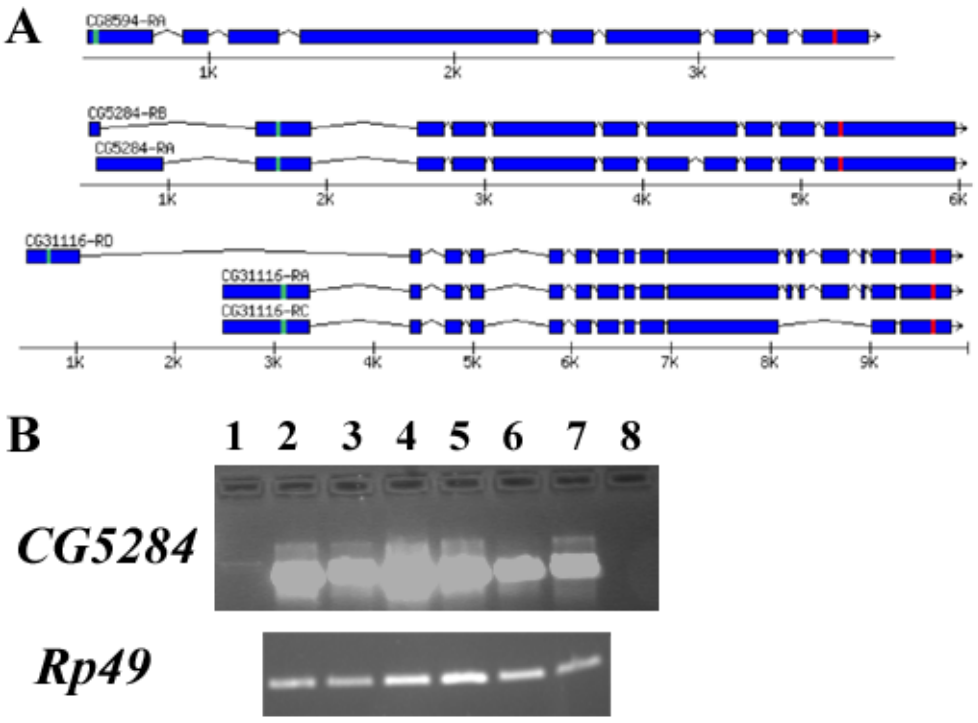


Figure R.8. No splicing defects were found in *Drosophila* chloride channels. A) Described ESTs from *Drosophila* genes which gave highest similarity to human *CLC1*. One, two and three splicing variants were found for *CG8594*, *CG5284* and *CG6942* (synonymous to *CG31116*), respectively **B)** 2% agarose gel electrophoresis of *CG5284* RT-PCR products. 1: No reverse transcriptase control; 2, 3: independent RNA extractions from embryos expressing CUG repeat RNA (*UAS-(CTG)480, 1.1* transgene) in a muscular (Mhc-Gal4) pattern; 4, 5: wild type controls (*OrR* and *yw,* respectively); 6: embryos expressing CUG repeat RNA (*UAS-(CTG)480, 2.1* transgene) in a general (*da-Gal4*) pattern; 7: mbl^{E27}/mbl^{E16} mutant embryos; 8: negative control for PCR (no cDNA template added). Band intensity differences in CG5284 amplification correlated with differences in RNA input denoted by *Rp49* amplification.

No ortologous gene has been described in *Drosophila* for *CLC1*. However, Genbank BLAST searches identified *CG5284*, *CG6942* and *CG8594* as sequences highly homologous to human *CLC1*. No functional studies have been carried out on any of them, but voltage-gated chloride channel activity has been inferred from their sequences (Flybase) and electrophysiological recordings in flies suggest that Cl^- currents are involved in *Drosophila* muscle physiology [117].

EST analysis showed that *CG5284* had two alternative splicing variants while *CG6942* had three. In order to analyse the possible conservation of chloride channel mis-splicing in *Drosophila*, we performed reverse transcription-polymerase chain reactions (RT-PCR) with primers encompassing the cassette exons. When amplifying *CG5284* exons seven to nine, only one band appeared, the same in embryos and adults, and no differences between wild type embryos and those expressing CUG repeats in a general (*daughterless*; *da*) or muscular (*Myosin Heavy chain*; *Mhc*) pattern were found. As no splicing event was detected to be altered, we performed RNA *in situ* hybridization in wild type flies to characterize the expression pattern of the three sequences. *CG5284* and *CG8594* gave a weak general signal and *CG6942* was detected in the Malpighian tubules, nervous system and intestine (not shown).

In summary, no chloride channel mRNA was found to have a muscle-specific pattern similar to C*LC1* expression in vertebrate skeletal muscle. Further optimization for the amplification of other alternatively spliced regions and a deeper analysis of splicing variants by RT-PCR should be performed to elucidate if any *Drosophila* chloride channel is aberrantly regulated in these flies.

I.4.2. *α-actinin.*

DM patients and *muscleblind* loss of function flies show a disruption of the Z-discs [41, 118]. Z-bands appear where the thin filaments of two adjacent sarcomere units overlap. Several proteins compose this connective structure, of which the α-Actinin (α-Actn) is the most abundant. Vertebrate *α-actinin* has tissue-specific isoforms that are developmentally regulated by the activity of Polypirimidine Tract Binding protein (PTB), CUG Binding Protein1 (CUG-BP1), Embryonically lethal abnormal vision Type RNA binding protein 3 (ETR-3) and CUB-BP1 and ETR-3 Like Factor 4 (CELF4) [119].

Drosophila α-actinin mutant flies show disruption of the Z-discs and the muscle attachments to epithelial tendon cells [120], which is a phenotype similar to *muscleblind* mutant embryos [41]. Transcripts from the unique *Drosophila α-actinin* gene undergo alternative splicing to generate both non-muscle (*α-actinin A*) and muscle-specific (*α-actinin B* in adult muscle and *α-actinin C* in larval muscle) isoforms (Fig. R.9A, B). To test the hypothesis that *Drosophila muscleblind* regulates *Drosophila α-actinin* splicing, we analyzed alternative exon usage by RT-PCR both in *muscleblind* mutant and wild type flies. To prevent the interference of second-site mutations that might have accumulated in chromosomes carrying *muscleblind* alleles, and to be able to identify homozygous mutant flies at all stages, we generated *mbl* ^{E27}*/mbl* ^{k7103} flies by crossing *mbl*^{E27}*/CyO, ubi-GFP* and *mbl*^{k7103}*/CyO, ubi-GFP* flies. Primers annealing at exons E5 and E10 were used to amplify alternatively spliced exons. (Fig. R.9B, C). A strong reduction in the levels of *α-actinin C*, the isoform

normally found in larval muscle, was detected in *muscleblind* mutant embryos from 16 to 18 h AEL. This correlated with increased levels of *α-actinin A* and *α-actinin B*. No changes in *α-actinin* isoform ratio were attributable to different amounts of starting material as shown by the *Ribosomal protein 49* (*Rp49*) transcript amplification (Fig. R.9D). No differences in the isoform ratio were found either in late pupae or adults (Fig. R.9E, F). In order to confirm that the defects were due to *muscleblind* loss of function we performed the same RT-PCR from mbl^{E27}/mbl^{k7103} embryos over-expressing MuscleblindC protein fussed to GFP. Unfortunately, no enough RNA was obtained to detect any PCR product.

We also characterized the *Drosophila α-actinin* splicing pattern in flies expressing CUG repeat RNA in a muscular and a general embryonic pattern to check whether the DM model flies presented any defect in the maturation of this mRNA. Non-muscle *α-actinin* isoform *A*, mostly found in adult wild type tissue, was reduced in embryos when expressing CUG repeat containing-RNA in both patterns (Fig. R.9G). Again, changes were not due to differences in the initial starting material and no changes were found in *α-actinin* pre-mRNA splicing in late pupae or adult flies.

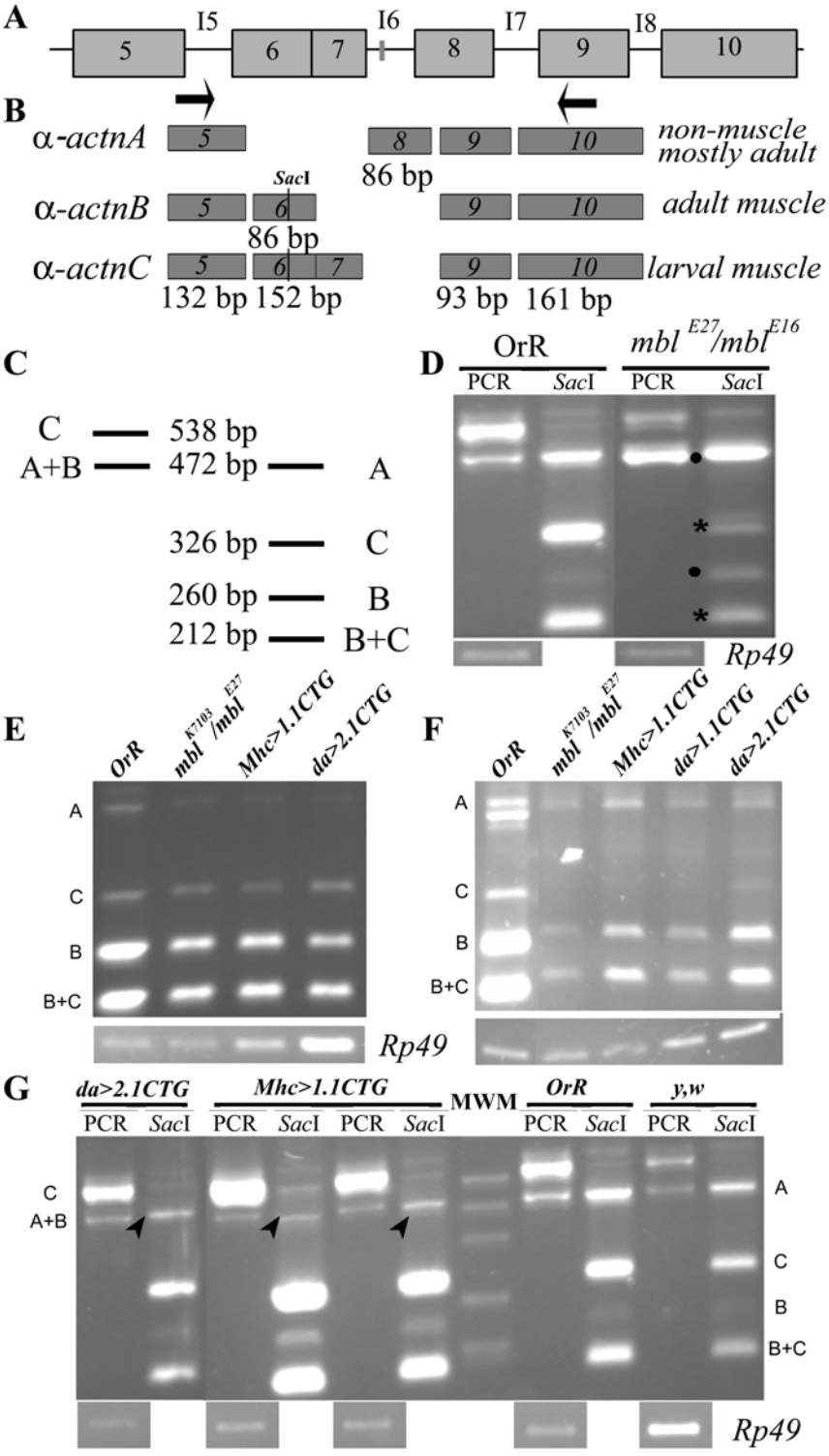

A
I5
I6
I7
I8
5
6
7
8
9
10
B
α-actnA
non-muscle
mostly adult
86 bp
SacI
α-actnB
adult muscle
86 bp
α-actnC
larval muscle
132 bp 152 bp
93 bp 161 bp
C
C
A+B
538 bp
472 bp
A
326 bp
C
260 bp
B
212 bp
B+C
D
OrR
mbl E27/mbl E16
PCR
SacI
PCR
SacI
Rp49
E
OrR
mbl K7103 E27/mbl
Mhc>1.1CTG
da>2.1CTG
A
C
B
B+C
Rp49
F
OrR
mbl K7103 E27/mbl
Mhc>1.1CTG
da>1.1CTG
da>2.1CTG
A
C
B
B+C
G
da>2.1CTG
Mhc>1.1CTG
MWM
OrR
y,w
PCR
SacI
PCR
SacI
PCR
SacI
PCR
SacI
PCR
SacI
C
A+B
A
C
B
B+C
Rp49

Figure R.9. *α-actinin* mis-splicing in *muscleblind* mutant embryos and embryos expressing CUG repeat RNA. A) Schematic representation (not to scale) of the *Drosophila α-actinin* genomic structure showing the intron/exon nomenclature used. Lines represent introns and boxes exons. Thick vertical bar represents a cluster of MBNL1 binding sequences found in intron I6 by *in silico* analysis **B)** Alternative splicing of *α-actinin* pre-mRNA is developmentally regulated giving rise to three mature transcripts referred to as *α-actinin A–C*. Primers used to amplify isoform-specific exons are denoted by arrows. Tissue and stage-specific expression is indicated on the right. Sizes in base pairs (bp) are given below each exon. *Sac*I restriction site used to differentiate isoforms *A* and *C* is denoted by a vertical line. **C)** Schematic representation of electrophoretic mobility pattern of bands of the PCR products and the *Sac*I digestions needed to reveal all *α-actinin* isoforms. Expected sizes are indicated. **(D-G)** Electrophoretic resolution (2% agarose) of RT-PCR products and their corresponding *Sac*I digestions from: **D)** wild-type (*OrR*) and mutant (mbl^{E27}/mbl^{E16}) embryos; **E)** wild type and *muscleblind* mutant pupae and pupae expressing 480 CUG repeat containing RNA in a general (*da>CTG*) and muscular (*Mhc>CTG*) pattern (numbers indicate transgenic strain); **F)** adults of the genotypes: *OrR*, mbl^{E27}/mbl^{E16} *Mhc-gal4/+;UAS(CTG480)1.1/+* (*Mhc>1.1CTG*); *da-gal4/+;UAS(CTG480)2.1/+* (*da>2.1CTG*) **G)** control (*OrR* and *y w*) embryos and embryos expressing repetitive RNA. *muscleblind* mutant embryos show diminished levels of *α-actnC* isoform (asterisks; **D**) and increased levels of *α-actinin B* (black dot) when compared with controls. Embryos expressing repeats show severe reduction of *α-actinin A* (arrowheads; **G**). No differences were observed in pupae or adult stages between mutants or CUG repeat expressing flies and controls **(E**, **F)**. No changes were attributable to RNA input as shown by *Rp49* amplification. Molecular weight marker VI (MWM) bands correspond to 517, 453, 394, 298, 234 and 154 bp.

In summary, *muscleblind* mutant embryos and embryos expressing CUG repeat RNA show mis-splicing of *α-actinin* mRNA. Surprisingly, however, the misregulation found in this case was different. Whereas *muscleblind* mutant embryos showed a reduction in *α-actnC*, CUG expressing embryos reduced *α-actinin* isoform *A*. These defects are specific to embryonic stages as they are not detected in pupae or adults.

I.4.3. ***Troponin T.***

Troponin T, together with Troponin C, Troponin I and Tropomyosin, belongs to a complex involved in the regulation of calcium-mediated contraction. Myofibrils containing the cardiac Troponin T (cTNT or TNNT2) protein isoform, coded by transcripts including the E5, are more sensitive to Ca^{2+} [31]. This exon is only included in transcripts expressed in skeletal and cardiac embryonic muscle fibres. The final isoform ratio of *cTNT* in cardiac cells is controlled by levels of splicing factors MBNL1, CUG-BP1, PTB and ETR-3 [31]. Human *cTNT* alternative splicing is altered in cardiaomyocytes of myotonic dystrophy patients as transcripts in adult tissue retain E5. *fast skeletal Troponin T* (*TNNT3*) splicing is also altered in myocytes of DM patients by abnormal inclusion of a cassette exon that is only present in developing normal muscles [28]. Both *Tnnt2* and *Tnnt3* mRNAs are aberrantly spliced in *Mbnl1* knockdown mice and mice expressing long CUG repeat RNA, showing alterations similar to those found in DM patients [28, 29].

Blast searches for the *Drosophila* homolog of human *cardiac troponin T* (NM_000364) returned *CG7107* (also known as *upheld, DmTROPT* and *troponinT*) as the sequence showing the highest homology. EST analysis detected three splice isoforms of *Drosophila troponin T* (*tnT*), although a forth transcript differing in a three nucleotide exon has also been reported [121]. The four splicing variants differ in the inclusion or exclusion of exons E3, E4 and E5 (Fig. R.10A). Splicing isoforms are muscle-type specific and developmentally regulated.

We performed RT-PCR reactions, amplifying exons E2 to E6 of *tnT* mRNA, from total RNA of *muscleblind* mutant and CUG repeat RNA-expressing flies. PCR product size identifies all described splicing isoforms (Fig. R.10). Regular (Fig. R.10B) and nested (Fig. R.10C) PCR showed no detectable difference in *muscleblind* mutant embryos and no reproducible defects in CUG expressing embryos. Early mutant pupae (Fig. R.10D, F) showed an increase in the *tnT* isoform specific to the tergal depressor of trochanter (TDT) and indirect flight muscles (IFM). This isoform was also clearly detected in mbl^{E27}*/CyO, ubi-GFP* flies, whereas it was barely detectable in mbl^{k7103}*/ CyO, ubi-GFP* ones (Fig. R.10D). This is consistent with a dependence on *mbl* function since *tnT* splicing defects correlated with *mbl* allele strength. CUG expressing flies also showed the same defect but milder, although differences were not statistically significant when comparing with flies carrying the UAS transgene. Furthermore, it was also detected in *Mhc-Gal4* flies. An unspecific band appeared when amplifying cDNA from adults collected at least six hours after eclosion, and all the samples gave the same amplification pattern (Fig. R.10E).

muscleblind mutant pupae show specific splicing defects that are already present in heterozygous individuals carrying a strong loss of function allele (mbl^{E27}) but not a weak one (mbl^{k7103}), thus pointing to a dose effect. The detection of defects in *Mhc-Gal4* flies does not allow us to form any conclusions about CUG-containing RNA over-expressing flies. Indeed, this *Mhc-Gal4* line had reduced viability (see section I.2 of results) and were difficult to maintain and amplify. Thus, *tnT* splicing alterations could be originated by the insertion of the Gal4 P-element into an essential loci.

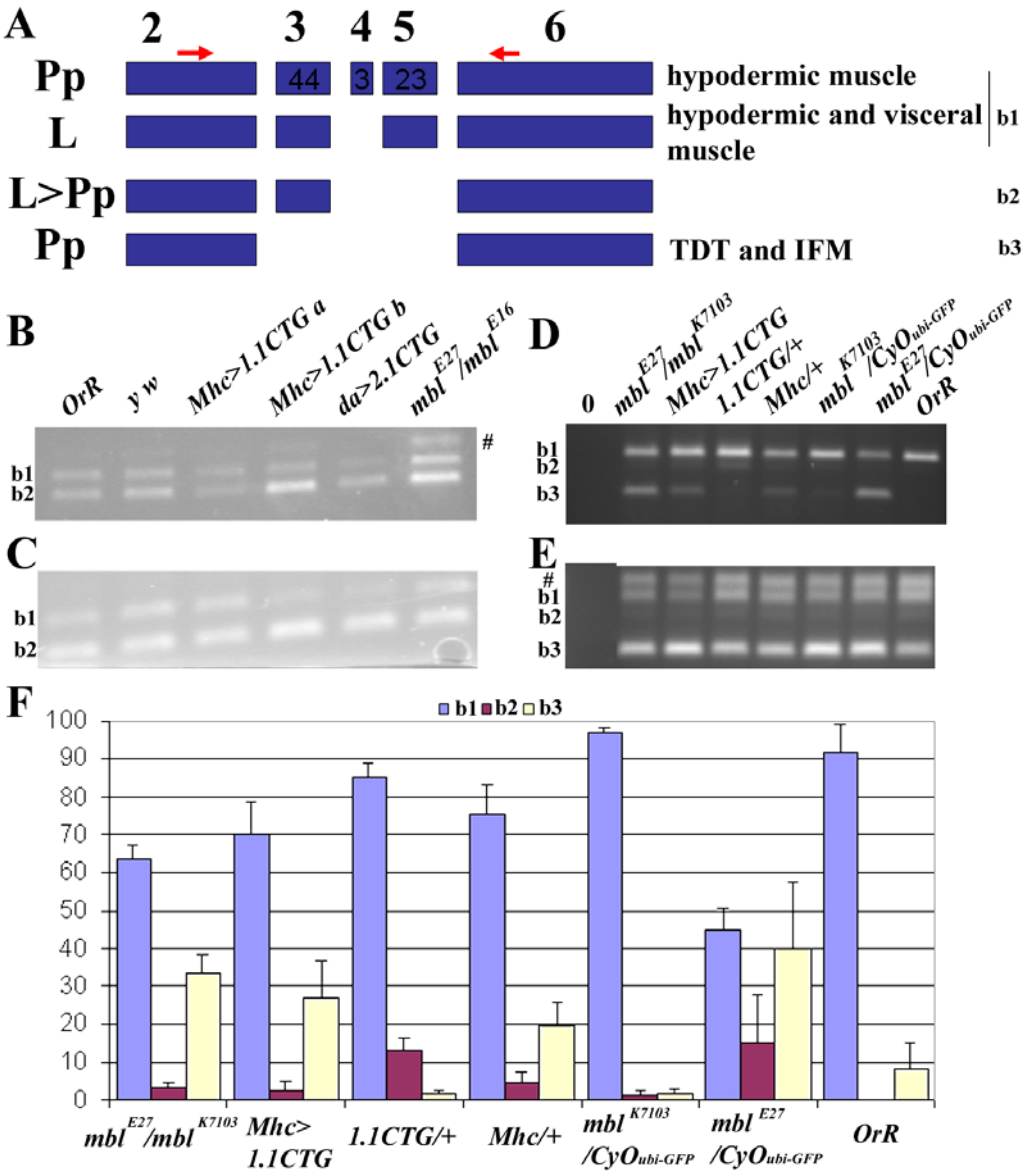


Figure R.10. *muscleblind* mutant pupae show alterations in the alternative splicing of *troponinT* RNA. A) Schematic representation (not to scale) of *Drosophila troponinT* transcripts showing the nomenclature used. Boxes represent exons and their sizes in base pairs are indicated inside. Primers used in the PCR to amplify the alternatively spliced region are denoted by arrows. Isoforms specific, or mainly expressed, in pupae or larvae, are

designated as Pp and L respectively. Tissue-specific expression is described on the right together with the nomenclature used in the rest of the figure. **B-E)** 2% agarose gel electrophoresis of RT-PCR products from embryos **(B, C),** pupae **(D)** and adults **(E)**. RNA extractions used were: **(B, C)** control strains *OrR* and *y w*; two independent extractions from flies expressing repetitive RNA in a muscular (*Mhc-Gal4>UAS-(CTG)480,1.1*; *Mhc>1.1CTG*) or general embryonic (*da-Gal4>UAS-(CTG)480,2.1*; *da>2.1CTG*) pattern; and *muscleblind* mutants (mbl^{E27}/mbl^{K7103}); **(D, E)** negative control for PCR (no cDNA; 0); *mbl* mutants; flies expressing repetitive RNA in a muscular pattern; control flies carrying the UAS (*UAS-(CTG)480,1.1/+*; *1.1CTG/+*) or the Gal4 (*Mhc-Gal4/+*; *Mhc/+*) transgene, or the *mbl* mutant alleles, and wild type *OrR*. # denotes an unspecific band. **F)** Average percentage of RT-PCR products from three independent RNA extractions from pupae.

Taken together the analysis of alternative splicing of a number of pre-mRNAs showed that *muscleblind* function is required for the alternative splicing regulation of *α-actinin* and *troponinT* pre-mRNAs. They also showed that Muscleblind function is transcript-specific as no alteration was found in *CG5284* splicing in *muscleblind* mutants. The absence of defects in *α-actinin* splicing in mutant pupae and adults and the developmental window where we detect defects in *troponinT* mRNA (pupa stage) indicate that *muscleblind* regulates alternative splicing during development, as described for human MBNL1 proteins [31, 32].

II. **Analysis of Muscleblind binding to RNA.**

Drosophila muscleblind encodes protein isoforms with a pair of CCCH zinc finger motifs [44]. Human MBNL proteins, which show two pairs of zinc fingers (except for a few isoforms generated by alternative splicing that lack some [38]), have been described as splicing factors that directly interact with RNA targets [32]. Deletion analysis with a yeast three hybrid assay

showed that zinc fingers are necessary for the interaction with RNAs containing CUG repeats [43].

We also wondered whether the binding of Muscleblind proteins to physiological targets and to CUG repeat containing RNA had different properties. Thus, we characterized the interaction between Muscleblind and CUG repeat RNA and target transcripts. First, we checked the ability of Muscleblind isoforms to bind to expanded CUG-containing transcripts in a human cell context. This assessed the conservation of the co-localisation described to occur in DM cells. Secondly, we tested the ability of the different Muscleblind isoforms to interact with both CUG-containing RNA and fragments of *Drosophila* α*-actinin* transcript in a yeast three hybrid system. Finally, to show that the interaction between Muscleblind proteins and α*-actinin* was direct, we performed *in vitro* UV crosslinking binding assays.

II.1.1. Co-localisation of Muscleblind proteins with expanded CUG containing-RNA in mammalian cells.

Myoblasts from myotonic dystrophy 1 patients show ribonuclear foci containing mutant *DMPK* mRNAs aberrantly retained in the nucleus [67]. Human MBNL proteins have been shown to co-localize with ribonuclear foci in DM myoblasts as well as in DM neurons and cardiomyocytes [73, 74, 76]. *Drosophila* Muscleblind binding activity to CUG repeat RNA has not been described. MblA, B and C contain two zinc fingers similar to those present in human MBNL, whereas MblD contains just one. As an initial test of *Drosophila* Muscleblind ability to bind CUGs, we worked in a mammalian

cell context to analyze the capacity of different *Drosophila* Muscleblind protein isoforms to co-localize with ribonuclear foci. Because there was no anti-Muscleblind antibody available, we decided to fuse Muscleblind proteins to the Green Fluorescent Protein (GFP), in order to detect them in cell culture assays. The activity of the fusion protein was tested *in vivo* in our laboratory with a mutant rescue experiment. The MuscleblindC:GFP coding sequence was cloned downstream of UAS sequences and transgenic flies were generated. Expression of UAS-MuscleblindC:GFP throughout the embryo (*daugtherless-Gal4* driver) rescued the *muscleblind* embryonic lethal phenotype to the same extent as wild type MblC [122].

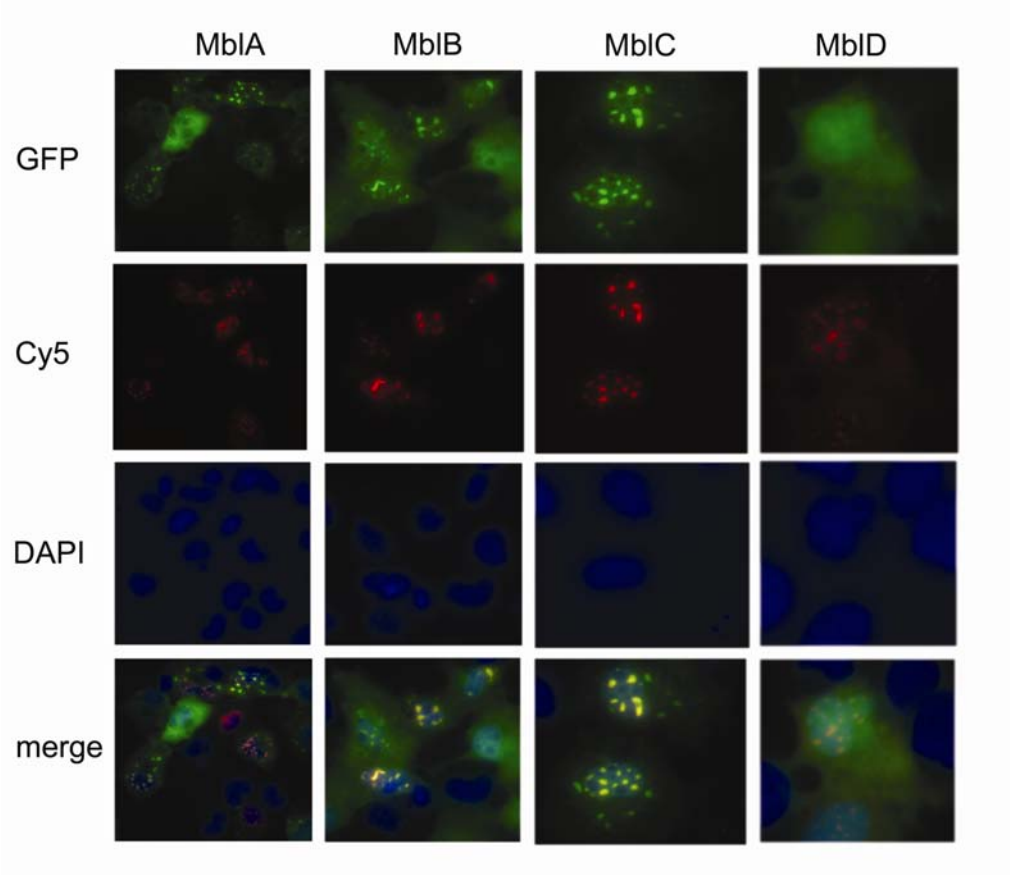


Figure R.11. Muscleblind proteins co-localise with CUG repeat containing RNA in COS cells. GFP-tagged Muscleblind proteins were expressed in COS cells together with *DMPK* 3'UTR (CUG_{197}) RNA. 1 μg of each plasmid was transfected as described in M&M. Mbl-GFP fusion proteins are detected in the green channel (GFP), hybridized CAG probe is in the red channel (Cy5), and nuclear DAPI staining is in the blue channel (DAPI). Images combining all three channels are shown at the bottom (merge). All muscleblind isoforms but MblD co-localise with mutant *DMPK* RNA (yellow signal in merged panels).

We co-transfected GFP-tagged Muscleblind isoforms and the expanded *DMPK* 3'UTR (CUG_{197}) in COS cells. In these experiments we found that MblA, B and C co-localized with CUG expansions in prominent ribonuclear foci, while MblD showed a diffuse signal (Fig. R.11). Foci were clearly nuclear but for MblA some of the nuclear inclusions were so close to the nuclear membrane that they might actually be cytoplasmic. These foci were different to those found when over-expressing the GFP-fussions alone (see section IV.1 and Fig. R.21). In order to test if the prominent protein foci originated from an excess of transfected protein, and to ensure that co-localisation was not cell type-specific, we co-transfected GFP fusions and CUG_{197} RNA in HEK293T cells using one third of the DNA used before. Ribonuclear foci persisted when working with MblA, being more prominent those close to the nuclear membrane; the ability of MblB to aggregate in RNA foci was sharply reduced, whereas MblC still co-localised with RNA foci although the size of the aggregates was smaller (Fig. R.12).

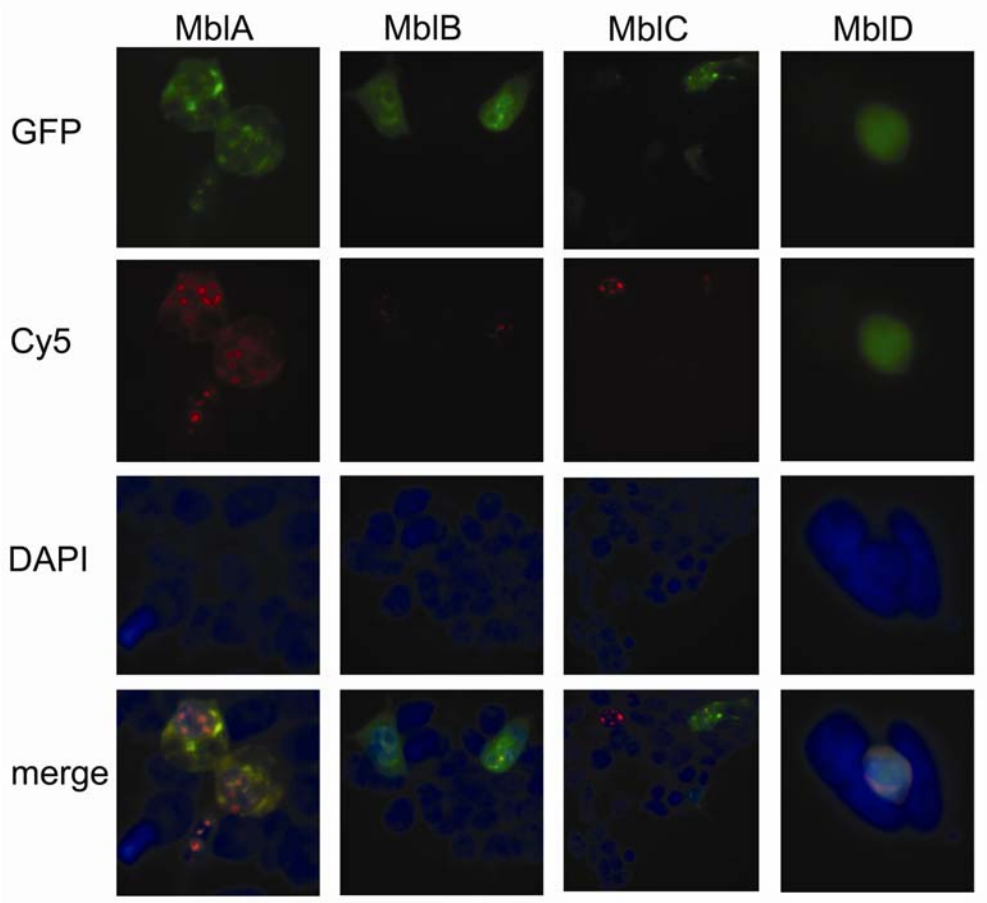


Figure R.12. Muscleblind proteins co-localise with CUG repeat RNA in HEK cells. GFP-tagged Muscleblind proteins were expressed in HEK293T cells together with *DMPK* 3'UTR (CUG_{197}) RNA. 300 ng of each plasmid were transfected as described in M&M. Detection of Mbl-GFP fusion proteins (GFP, green), CAG probe hybridizing to CUG repeat-containing RNA (Cy5, red), and DAPI staining the nuclei (DAPI, blue) is shown. All muscleblind isoforms but MblD co-localise with RNA.

These results were fully consistent with those obtained by Mike Poulos (Dept. of Molecular Genetics and Microbiology, University of Florida, USA) in HeLa cells and demonstrate that *Drosophila* Muscleblind proteins co-localise with CUG trinucleotide expansions in mammalian cells.

II.1.2. **Interaction of Muscleblind proteins with CUG repeat containing-RNA and *α-actinin* transcript in a yeast three hybrid assay.**

The yeast three hybrid system is a useful technique for analysing protein-RNA interactions *in vivo* (reviewed in [100]) (Figure R.13.A). It also allows mutational screens to be carried out to detect both amino-acids of the protein and nucleotides of the RNA that are required in a given interaction. MBNL1 interaction with pathogenic CUG and CCUG repeats has been analyzed in depth using this system [43]. This study showed that both the zinc fingers and the linker region between them were necessary for the interaction with the repetitive RNA. A UV crosslinking assay showed that MBNL1 binds the consensus sequence YGCUU/GY (where Y is a pyrimidine) in *cardiac Troponin T* mRNA [32].

Apart from the zinc fingers, no other functional domains have been described in *Drosophila* Muscleblind proteins. Therefore, we decided to assay the interaction properties of the four natural isoforms (see Fig. I.5C) and two artificial constructs, one containing the region common to MblA, B and C (referred to as ZC) and the other containing just the two zinc fingers (referred to as CCCH), to elucidate which regions of the protein are involved in RNA binding.

To further analyze the interaction between *Drosophila* proteins and DM pathogenic expansions we cloned 480 interrupted CUGs and 480 interrupted CAGs as control of sequence specificity in the expression vector pIIIA/MS2.2 [99]. Given the splicing defects in *Drosophila α-actinin* mRNA in *muscleblind* mutant embryos, we expected Muscleblind proteins

to regulate *α-actinin* splicing by binding the mRNA directly. To test this hypothesis we computationally screened the genomic sequence of the alternatively spliced region of *Drosophila α-actinin* for MBNL1 consensus binding sequences. The *in silico* analysis revealed a cluster of seven overlapping perfect matches of MBNL1 consensus binding sequence 132 nucleotides downstream of the 3' end of exon E7 (Fig. R.9A). The three hybrid system is optimized for RNA fragments around 150 nucleotides [100]. Therefore, we designed Actn1, without binding sequences, to be 142 nucleotides long and Actn2, containing the cluster, to be 161 nucleotides long, and we cloned them in the appropriate expression vector (see section II.2.1 of M&M and Fig. R.13D).

Interrupted CUG repeat RNA generates hairpin structures similar to those formed by pure CUG repeat RNA.

Because we used CTG repeats interrupted every 20 units by the sequence ctcga [35], we first checked that our repeats would fold into double stranded hairpins like pure CTG repeats. We introduced the complete sequence expressed from pIIIA/MS2.2 plasmids into the web-based RNA folding prediction program Mfold (see M&M). The study of the predicted secondary structures of the RNA expressed showed that interrupted 480 CUG repeats generate a hairpin similar to that described for pure repeats (Figure R.13B-C'). Predicted free energies of structures generated by interrupted and pure repeats were also in the same range. A similar structure was obtained when introducing the 480 CAG construct sequence in the RNA folding website. This supported that 480 CUG repeat RNA would faithfully model the behaviour of pure repeats.

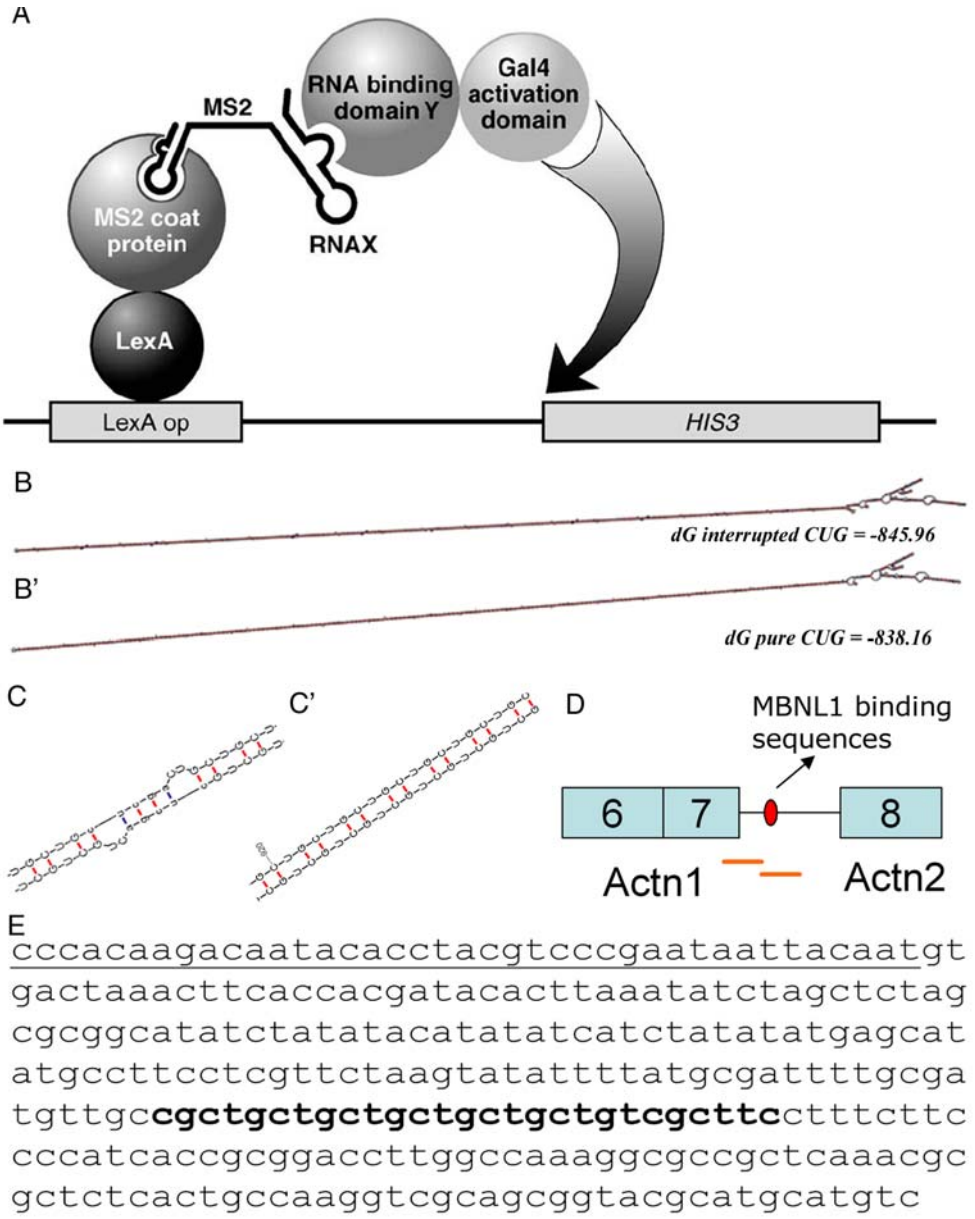


Figure R.13. Schematic representation of the yeast three hybrid system and hybrid RNAs used. A) Overview of the three-hybrid strategy to detect RNA–protein interactions. A hybrid protein containing LexA DNA-binding domain with RNA-binding domain of MS2 coat protein localizes to the promoter of reporter genes *HIS3* and *LacZ*. A second hybrid protein containing *Gal4* transcriptional activation domain with the protein domain that we want to

test for RNA-binding activity will activate transcription of the reporter gene when in close proximity to the upstream regulatory sequences of reporter genes. A hybrid RNA containing sites recognized by MS2 coat protein and the RNA submitted to study links the two hybrid proteins to one another. When the RNA and the protein of interest interact the tripartite complex results in detectable expression of the reporter genes. **B)** RNA folding of the sequence expressed from pIIAMS2-480CTG interrupted repeats gave a similar hairpin structure to that obtained when introducing 480 pure repeats **(B')**. The most frequent conformation and its free energy (dG) are shown. **C)** Zooming into the hairpin the cucga sequence that interrupts the CUG tandems are shown to generate a small loop that does not interfere with the general hairpin structure. **C')** Zooming into the predicted structure for pure CUG repeats is shown. **D)** Schematic representation of *α-actinin* alternatively spliced region showing the cluster of MBNL1 binding sequences found downstream E7 (red dot) and Actn1 and Actn2 fragments used in the three hybrid system and UV-crosslinking assays. **E)** Nucleotide sequence of the region encompassing the 3' end of exon 7 and the beginning of the I6 intron. Protein coding sequence is underlined and the putative MBNL1 binding sequences are highlighted in bold

Muscleblind isoforms show differential binding activity in a yeast three hybrid system.

In order to detect the smallest fragment of Muscleblind protein that can bind RNA we cloned Muscleblind isoforms and two constructions, one encompassing the common region of MblA, B and C, and one containing the two zinc fingers. Initially we transformed L40 coat yeast with each of the pIIIA/MS2.2 constructs for the expression of the hybrid RNA (480 CUG repeats, 480 CAG repeats and Actn1 and Actn2 fragments of *Drosophila α-actinin* mRNA), and grew the yeast on synthetic dextrose medium (SD) without tryptophan and uracil to select L40coat colonies that had taken up the plasmid. After 3 to 4 days of growth, a colony carrying each RNA expression plasmid was grown in YAPD2x (growing in selective medium did not work) and transformed with the pACT2 constructs for hybrid protein

expression. For the initial experiments, constructs of MblB, C and D as well as ZC fused to a Gal4 transcription activation domain were available. Colonies of each protein/RNA combination were plated on increasing concentrations of *HIS3* competitive inhibitor 3-AT (see M&M). Ability to grow and representative pictures are shown in Table R.1. and Fig. R.14A-C, respectively. MblC interacted with both CUG and CAG repeat RNA and also with the Actn2 fragment whereas MblB only interacted with CAG repeats and MblD with the Actn2 fragment. The common region gave no repetitive results with Actn2 RNA in the two initial experiments.

	480 CUG			480 CAG			Actn1				Actn2				IRE		
[3-AT]mM	0	1	5	0	1	5	0	1	2.5	5	0	1	2.5	5	0	1	5
Vector	+	-	-	+	-	-	+ +	- -	- -	- -	+ +	- -	-	- -	+	+	-
IRP	+	-	-	+	-	-	+	-	-	-	+	-	-	-	+	+	+
ZC	+	-	-	+	-	-	+ +	- -	- -	- -	+ +	- +	+	- +	+	-	-
MblB	+	-	-	+	+	+	+	-		-	+ +	- -	-	- -	+	-	-
MblC	+	+	+	+	+	+	+	-		-	+ +	+ +	+	+ +	+	-	-
MblD	+	-	-	+	-	-	+	-	-	-	+ +	+ +	+	+ +	+	-	-

Table R.1. Muscleblind isoforms show differential binding activity in a yeast three hybrid assay. Combinations of pACT2 constructs coding hybrid proteins (first column) and pIIIA/MS2 for the expression of RNAs (first row) that were used to transform yeasts. Growth (positive interaction) is denoted by +; absence of growth by -. Empty cell means no replica was made at that concentration of inhibitor. Iron Regulatory Protein (IRP) was used as positive control of interaction with Iron Response Element (IRE). Empty vector was used as negative control. ZC designates region shared by MblA, B and C.

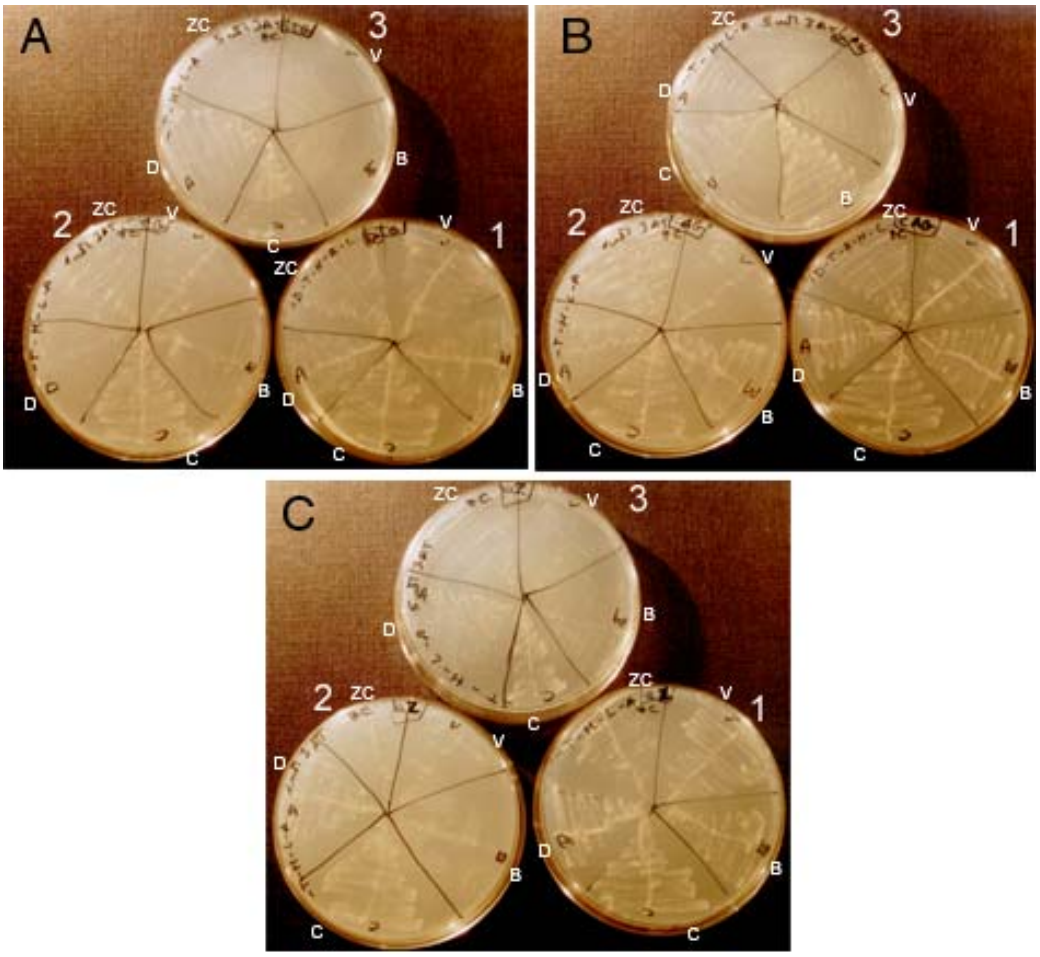


Figure R.14. Ability to grow of yeasts expressing different combinations of Muscleblind isoforms and target RNAs. Colonies transformed with pACT2 vector (v), pACT2-MblB (B), pACT2-MblC (C), pACT2-MblD (D), and pACT2-ZC (ZC) together with pIIIA/MS2-CTG **(A)** pIIIA/MS2-CAG **(B)** pIIIA/MS2-Actn2 **(C)** were grown in SD plates without triptophan, histidine, leucine and adenine (-T –H –L –A) and increasing concentrations of 3-AT: none (1), 1 mM (2) and 5 mM (3).

The experiment was repeated several times because yeasts eventually grew in negative controls as well. It has been shown that the YBZ1 yeast

strain, with a mutant MS2 coat protein, reduces the background of RNA-independent interactions, enabling the detection of weaker interactions [100]. Although this system was more specific (absence of non-specific interactions with empty vectors, heterologous proteins, or RNAs at 3-AT inhibitor concentration of 1 mM), it was also less sensitive. The strength of the control protein-RNA interaction was reduced and previous interactions were not detected with the YBZ1 strain.

As no conclusive results were obtained with yeast, we looked for an alternative method to study the binding properties of Muscleblind proteins.

II.1.3. *in vitro* UV-crosslinking assay for binding to *α-actinin* RNA.

UV crosslinking experiments showed that human MBNL proteins bind *cTNT* mRNA at sequences located 18 and 36 nucleotides upstream of exon E5. Elimination of those binding sites reduces the ability of MBNL proteins to repress E5 inclusion [32]. We decided to test *Drosophila* Muscleblind ability to directly bind *α-actinin* RNA through the cluster of MBNL1 consensus binding sequences located in intron six (Fig. R.9A). To do so, we tagged Muscleblind proteins with GST at N-terminus and His_6 at C-terminus. We expressed constructs in bacteria, purified tagged proteins, and performed *in vitro* UV crosslinking assays.

Muscleblind proteins fused to GST and His_6 are highly insoluble and unstable when expressed in *E.coli*.

Small scale expression of MblA tagged with GST (5') and His_6 (3') showed that most of the protein remained insoluble. Furthermore, protein in the soluble fraction was degraded (Fig. R.15A). A high scale expression at 16ºC using BL21 strain eventually gave enough soluble protein for additional experiments (Fig. R.15B). Glutathione Sepharose purification recovered most of the protein, although a significant proportion remained in the unbound fraction (Fig. R.15C). Subsequent purification with a Ni-NTA column gave a cleaner but still degraded protein preparation. Expression of MblC fusion protein gave a similar degradation pattern and recovery results after GST and His purification (Fig. R.15E). After desalting with PD-10 columns (see M&M), western blotting of the final protein preparation with anti-His and anti-GST antibodies showed that full length protein was present and most of the degradation fragments were N-terminal (Fig. R.15D,F).

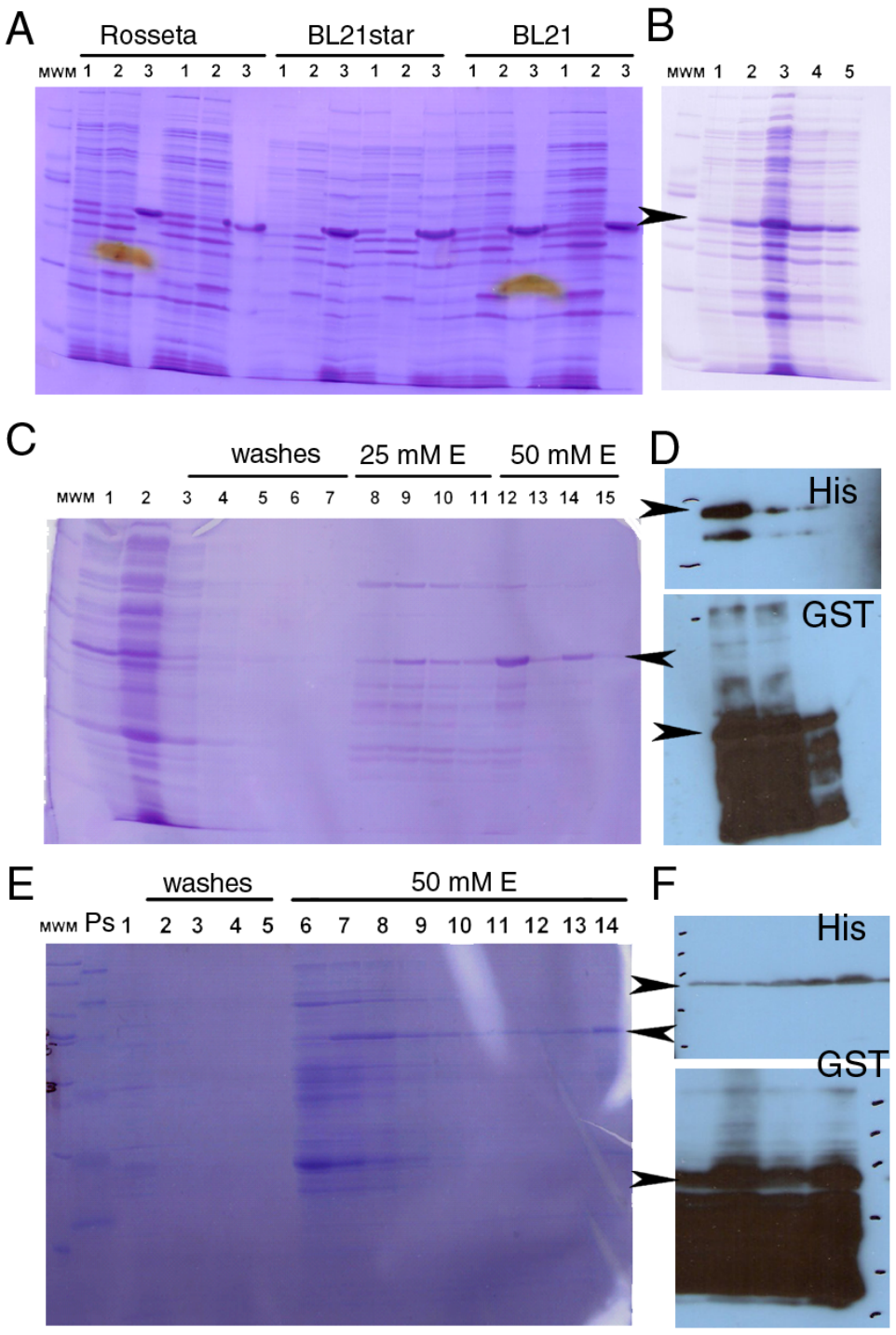

A
Rosseta
BL21star
BL21
MWM 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
B
MWM 1 2 3 4 5
C
washes
25 mM E
50 mM E
MWM 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
D
His
GST
E
washes
50 mM E
MWM Ps 1 2 3 4 5 6 7 8 9 10 11 12 13 14
F
His
GST

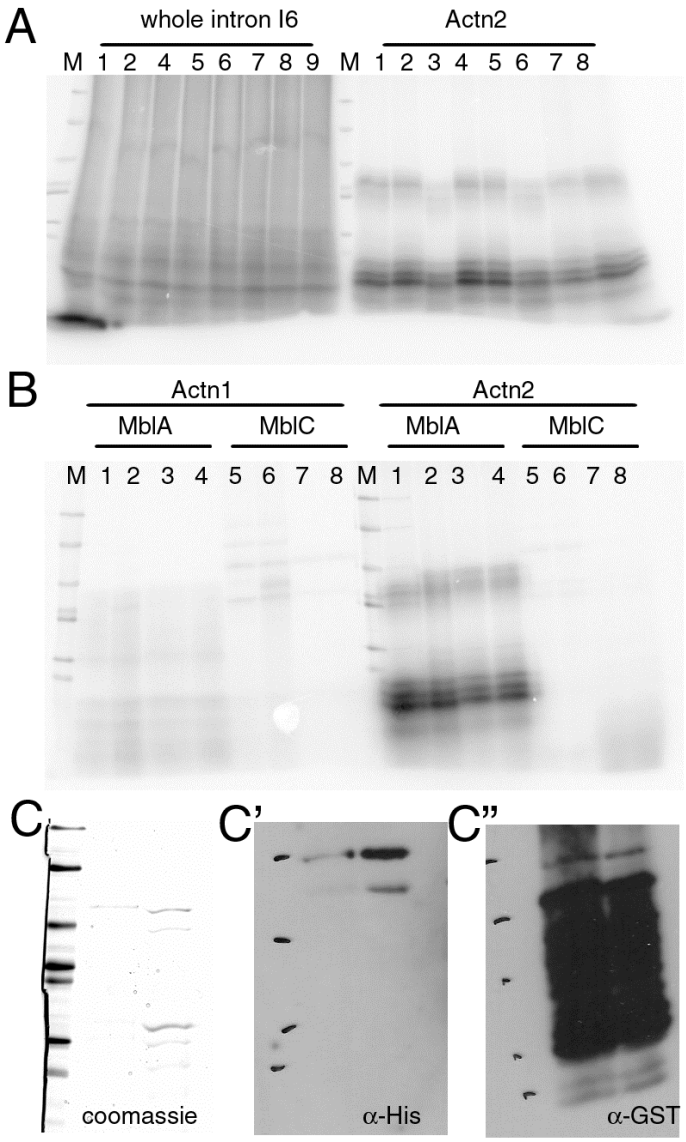
Figure R.15. Muscleblind proteins expressed in bacteria are insoluble and unstable. Arrowheads mark full-length protein migration **A)** Small scale expression of GST-MblA-His_6 protein in different *E.coli* strains (Rosseta, BL21star and BL21). Two colonies where induced in order to reveal colony dependence of expression efficiency. 1: sample 15 min after induction with IPTG; 2: soluble protein; 3: insoluble fraction. **B)** High scale protein expression of MblA in BL21 with 24 h induction at 16ºC. Lanes correspond to the following samples: 1: uninduced; 2: induced; 3: soluble; 4: extracted; 5: insoluble. **C)** Elutions from Glutathion Sepharose purification of MblA. Samples loaded were 1: induced protein pooled; 2: unbound; 3-7: washes with MTPBS buffer; 8-11: elutions with 25 mM Glutathion; 12-15: elutions with 50 mM Glutathion. **D)** Western blotting with α-His and α-GST antibodies of final fractions obtained after PD-10 column elution of MblA. Note the detection of several GST-tagged N-terminal degradation products. **E)** Elutions from Ni-NTA column purification of MblC. **F)** Western blotting with anti-His and anti-GST antibodies of final fractions obtained after PD-10 column elution of MblC. Note that the protein is degraded in N-terminal fragments reactive to anti-GST antibody. Molecular weight marker (MWM) bands of 116, 97, 66, 57, 55, 43, 35, 28 and 13 KDa are shown in (A), (B), (C), and (E). Pre-stained (Ps) molecular weigh marker bands of 47.5 and 32.5 KDa are handwritten in (D) and 83, 62, 47.5, 32.5, 25 and 16.5 KDa bands are in (E) and (F). The 16.5 KDa band is not shown in the film showing the anti-His staining. Note the different mobility of pre-stained molecular weigh marker compared to MWM in (E).

The N-terminal region of MblA interacts with a fragment of *Drosophila α-actinin* containing MBNL1 consensus binding sequences.

UV-crosslinking of MblA or MblC proteins with Actn1 and Actn2 fragments (Fig. R.16) only gave a strong signal of interaction between MblA and Actn2 although the protein was severely degraded during the process (signal around 30 KDa). Just a weak signal appeared when incubating MblC and Actn2 and no signal of interaction was detected between Actn1 RNA and MblA and C proteins in these experiments. Proper size and equivalent amount of RNAs was checked by acrylamide gel electrophoresis by Clare Gooding (Dept. biochemistry, University of Cambridge) confirming

that the same amount of input RNA was used in all conditions. As intron 6 contains other MBNL1 consensus sequences, we transcribed the whole intron and tested the RNA for the interaction with different protein preparations of MblA and C, and also with protein extracts from cells transfected with GFP-tagged Muscleblind proteins (see section II.12 of M&M for constructs and II.18/20 for transfection procedures). Very weak signal around 30 KDa was detected in all samples. To test the specificity of the interactions between MblA with both the whole intron 6 and the Actn2 fragment, we titrated both rRNA and heparine, which make the binding reaction more stringent. Both interactions persisted even under the most stringent conditions assayed (Fig. R.16A).

Figure R.16. MuscleblindA binds a fragment of *α-actinin* that contains MBNL1 binding sequences through the N-terminus. A) 15% SDS-PAGE resolution of UV crosslinked samples. MblA protein was incubated with the entire I6 or the Actn2 fragment, and rRNA and heparine were titrated to test the specificity of the binding. M: Molecular weight marker bands 116, 97, 66, 57, 55, 43, 35 KDa; 1: no rRNA, no heparine; 2: 0.05 μg/μl rRNA; 3: 0.1 μg/μl rRNA; 4: 0.2 μg/μl rRNA; 5: 0.7 μg/μl heparine; 6: 1.4 μg/μl heparine; 7: 2.8 μg/μl heparine; 8: 5.5 μg/μl heparine. **B)** 15% polyacrilamide gel of *in vitro* UV-crosslinking reactions. MblA and MblC were incubated with Actn1 and Actn2 under restrictive conditions. M: Molecular weight marker as (A); 1: no rRNA, no heparine; 2: 0.2 μg/μl rRNA; 3: 5.5 μg/μl heparine; 4: 0.2 μg/μl rRNA and 5.5 μg/μl heparine. Coomassie staining (**C)** and antibody detection with anti-His antibody **(C')** and anti -GST antibody of a sample of MblA protein going through the UV-crosslinking assay (left lane) and a freshly thawed sample (right). Note N-terminus degradation products. Molecular weight marker bands are in C as in (A); in C': 47.5, 32.5, 25 and 16.5; and in C": 47.5, 32.5, 25, 16.5 and 6.5.



To confirm the results we incubated both MblA and MblC preparations with Actn1 and Actn2 fragments under more restrictive conditions, i.e. including heparine, rRNA or both in the incubation reaction, in order to compete away unspecific interactions. Only MblA incubated with Actn2 RNA withstood the

interaction signal (Fig. R.16B). A protein gel electrophoresis of a protein aliquot of MblA carried through the crosslinking protocol in parallel to a freshly thawed equivalent amount of protein showed that degradation was occurring during the crosslinking protocol and very little full length protein lasted at the end (Fig. R.16C). Western blot showed that fragments observed to bind RNA were amino terminal as they were detected by anti-GST antibody but not anti-His (Fig. R.16C',C").

To further demonstrate that Actn2 fragment, but not Actn1, contains the sequences where MblA binds *α-actinin* mRNA, we performed a competition assay. The reaction was set up in the same conditions using 200, 300 and 400 molar excess of Actn1 and Actn2 competitor RNAs. No lane showed a reduction in the binding signal thus indicating that the amount of competitor RNA was not enough. The concentration of competitor RNAs and reaction volume limitations did not permit to perform assays with higher amount of competitor RNA. Therefore, further optimization of protein extraction and in vitro transcription are necessary to properly show that the binding signal is specific.

In summary, Muscleblind proteins have been found to be insoluble and very unstable when expressed in bacteria. Despite of that, we managed to show that MblA might interact with a fragment of *α-actinin* mRNA that contains MBNL1 consensus binding sequences and not with the adjacent fragment, lacking those consensus sequences.

III. Analysis of Muscleblind molecular function.

Human MBNL proteins are RNA binding proteins that regulate alternative splicing of several muscular transcripts in cell culture [32]. MBNL1 splicing activity on *cTNT* pre-mRNA has been studied in depth [31] showing that a developmental switch in *cTNT* isoform ratio arises from the developmentally-regulated expression of splicing factors such as MBNL1 and CUG-BP1. Moreover, previous work in our laboratory showed functional conservation between *Drosophila* Muscleblind proteins and human MBNL1 [107] and we here describe specific alternative splicing defects in *Drosophila muscleblind* mutants (Fig. R.9 and 10). We therefore decided to confirm the hypothesis that *Drosophila* Muscleblind proteins regulate alternative splicing of specific pre-mRNAs by performing different *in vitro* minigene splicing assays.

HEK293T and COS cells give very good results in experiments using minigenes, since the transfection efficiency and also the levels of transgene expression are very high. Mammalian cells have been used to analyse the function of Drosophila proteins in a number of studies (for examples see [123, 124]). Given the strong functional conservation showed between Drosophila Muscleblind proteins and vertebrate MBNL proteins, we decided to undergo the analysis in mammalian cells. This approach has the limitation of a not being a physiological situation but it gives the opportunity to test the functional conservation in the other direction. Human MBNL1 protein was shown to be functional in the fly; with these experiments we tested if the fly proteins can modify alternative splicing in a vertebrate cell environment.

III.1. *Drosophila α-actinin* minigene splicing assay in human cells.

Upon finding that normal alternative splicing of *α-actinin* pre-mRNAs requires *muscleblind* and that Muscleblind proteins show differential binding activity to *α-actinin* pre-mRNAs (Fig. R.9 and R.16), we next tested the ability of Muscleblind isoforms to modify *α-actinin* alternative exon usage in cell culture. In order to perform these experiments we built a *Drosophila α-actinin* minigene in a mammalian expression vector, which was transfected along with epitope-tagged Muscleblind proteins. The *α-actinin* minigene included cassette exons 6, 7 and 8, and their intervening sequences, flanked by constitutive exons 5 and 9 (Fig. R.17A).

The *α-actinin* minigene and GFP-tagged *Drosophila* Muscleblind isoforms, or human MBNL1 protein, were co-transfected into human HEK293T cells. 24 h after co-transfection, RNA was extracted and minigene products analyzed by semi quantitative RT-PCR. The minigene basal splicing pattern was characterized by co-transfecting control vector expressing GFP alone. This analysis revealed splice isoforms *α-actnA*, *B*, *C* and a new isoform that hasn't been described to appear in flies, possibly skipping exons 6, 7 and 8. We will refer to as the *short isoform*. *Sac*I digestions of PCR products revealed that *α-actinin* isoform *B* (typical of adult muscle, see Fig.R.9) was predominantly generated in this cellular context (Fig. R.17C). Both MblC and human MBNL1 were able to alter *α-actinin* minigene splicing so that the short isoform and isoform C, typical of larval musculature, were reduced (Fig. R.17B). Activity of GFP-Bruno proteins on *α-actinin* minigene was also tested but no reproducible effects were observed.

When increasing the amount of protein-coding plasmids transfected, we detected that GFP protein itself could be interfering with splicing. For this reason, we tagged Muscleblind isoforms with a myc epitope and repeated the experiments. Constructs expressing myc-tagged Muscleblind proteins were co-transfected with the *α-actinin* minigene and the ratio of *α-actinin* minigene splice products quantified (Fig. R.17D, F). myc-tagged Muscleblind proteins had a similar effect onto *α-actinin* minigene isoform ratio that GFP-tagged proteins.

MblC showed the highest activity in these experiments as the *α-actinin* C and short isoforms were reduced compared with the splicing pattern obtained when the empty vector was transfected. MblA also reduced the amount of *α-actinin* short isoform. Other Muscleblind isoforms did not alter minigene splicing. These results were reproduced in COS cells by Mike Poulos (Dept. of Molecular Genetics and Microbiology, University of Florida, USA). When analysing protein levels, MblC appeared to be more stable than other isoforms both GFP and myc-tagged, which could influence in the higher activity detected (Fig. R.17C). MblD was not detected in any western blot but one made by Mike Poulos with protein extracts from COS cells.

From these experiments we conclude that MblC regulates *α-actinin* splicing in human HEK293T cells repressing isoform *C* and short isoform during *α-actinin* pre-mRNA splicing.

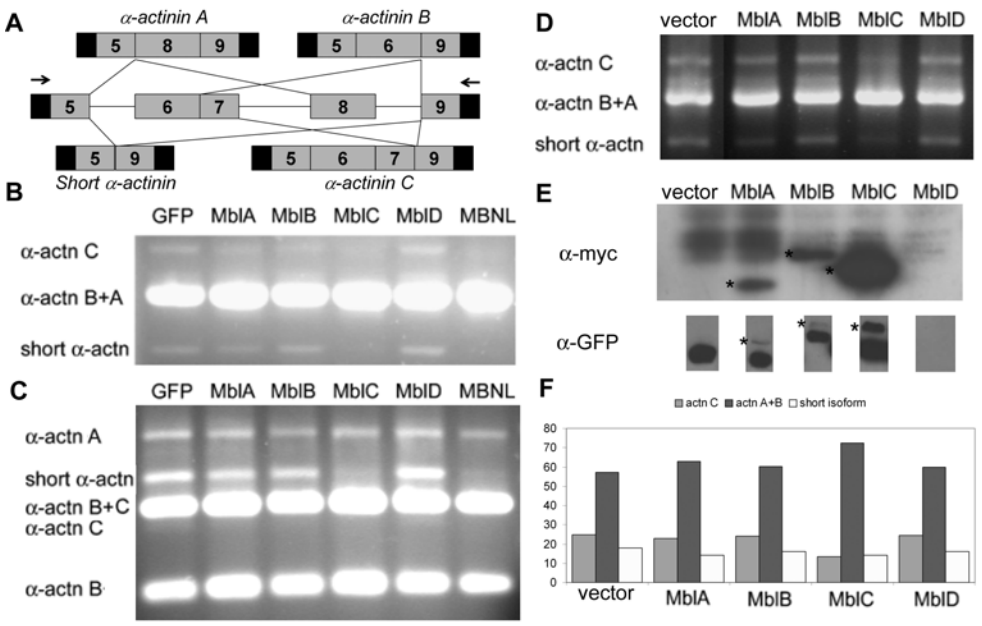


Figure R.17. Expression of Muscleblind proteins affects *Drosophila α-actinin* minigene splicing regulation in cell culture. A) Schematic representation of the structure of the *α-actinin* minigene (middle) and the three alternative splicing products. Arrows denote primers used in the RT-PCR. **B)** Electrophoretic separation of PCR products (2% agarose gel) after co-transfection of the *α-actinin* minigene and GFP-tagged *Drosophila* Muscleblind protein isoforms, human MBNL1 or vector alone in HEK293T cells, as indicated. MblC consistently repressed a *short α-actinin* transcript isoform (possibly skipping all alternative exons between exons 5 and 9) and *α-actinin C*, whereas MblA showed a weaker effect. **C)** *Sac*I digestion to differentiate *α-actininA* and *B* (see figure R.9) showed a marked predominance of *α-actininB*. **D)** 2% agarose gel of RT-PCR products obtained when expressing myc-tagged Mbl isoforms show the same effects observed with GFP-tagged proteins. **E)** Western blotting (anti-GFP and anti-Myc) of protein extracts from transfected HEK cells showed that all Muscleblind protein isoforms were expressed in this cell type, except MblD that was not detected, and that MblC was more stable or efficiently translated than other isoforms. Asterisks denote the migration of full length protein. **F)** Quantification of band intensity in the gel shown in D with 1DAdvanced Phoretrix software gave the following values of *α-actininA+B* (AB), *α-actininC* (C), and *short α-actininC* (s): vector: C= 24.7,

AB=57.2, s=18.1; MblA: C= 23, AB=62.9, s=14.1, MblB: C= 23.9, AB=60.3, s=15.8, MblC: C= 13.4, AB=72.5, s=14.1, MblD: C= 24.2, AB=59.8, s=16.

III.2. Expression of CUG repeat RNA alters *Drosophila α-actinin* minigene splicing in human cells.

Long non-coding CUG repeats can be pathogenic to humans and model organisms [52, 69, 84]. Expanded CUG repeat RNA forms double-stranded hairpins that interfere with the function of Muscleblind proteins [51]. Accordingly, expression of CUG repeat RNA in cell culture produces splicing defects in MBNL1 targets similar to depletion of MBNL1 proteins [32]. To test whether CUG repeats interfere with a *Drosophila* Muscleblind target, we transiently co-expressed the *α-actinin* minigene and human *DMPK* 3'UTR containing 160 to 200 CUG repeats (referred to as $(CUG)_{197}$) in cell culture. Minigene titration in HEK293T and COS cells showed that HEK cells were more convenient to study the effect of CUG expression (Fig. R.17A, B).

Expression of $(CUG)_{197}$ RNA interfered with regulated exon choice in *Drosophila α-actinin* pre-mRNA increasing the expression of *α-actinin* isoform *C* from a 9 to 50% of total transcript and reducing *α-actinin A + B* from 91 to 50% of the total product (representative gel is shown in Fig. R.18C and quantification of two experiments in Fig. R.18D). These results are similar to the splicing defects found in *α-actn* pre-mRNAs when analysing CUG expressing flies (note that the expression of CUG expanded repeat containin RNA in flies generated the reduction of *α-actinin* isoform *A*). In order to analyse the ability of Muscleblind proteins to rescue the effect of CUG expression, we also co-transfected the *α-actinin* minigene

and CUG repeats with Muscleblind isoforms (Fig. R.18C lanes 3-7). No conclusion could be drawn as the expression of GFP alone already changed the splicing pattern.

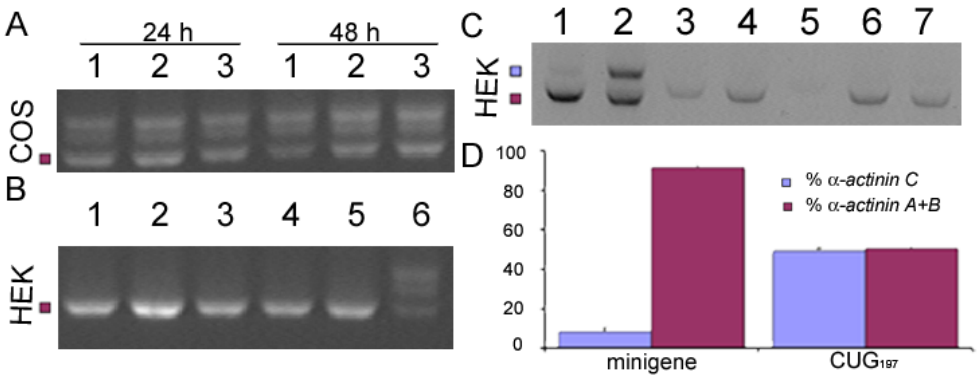


Figure R.18. Expression of CUG repeat containing RNA modifies *Drosophila α-actinin* minigene mRNA splicing in cell culture. Blue box marks *α-actininC* transcripts, Maroon box marks *α-actininA+B* RNAs. **A)** RT-PCR amplification of minigene products upon transfection of COS cells with 0.25 (1), 0.5 (2) or 1 μg (3) of *α-actinin* minigene performing the RNA extraction 24 or 48 h after transfection. No difference on RT-PCR sensitivity was detected when increasing either the amount of vector or the growing time after transfection. An unspecific band was detected. The upper band should correspond to the *α-actininC* product because of the size, but it was not confirmed by sequencing. **B)** RT-PCR amplification of minigene products upon co-transfection of HEK cells with 1 (1-3) or 2 μg (4-6) of minigene plasmid and 0.5 μg of pEGFP-C1 (GFP control) (1), pMV2-MblC (2) and pMV2-MblD (3) or 1 μg of pGEM42 (4), pMV-MblC (5) and pSP72-DMPK$(CUG)_{197}$ (6). **C)** Amplification of minigene splicing products after co-transfection of HEK cells with 0.5 μg *α-actinin* minigene and 0.5 μg pSP72-DMPK$(CUG)_{197}$ (2-7) together with 0.5 μg of pGEM42 (2), pEGFP-C1 (3), and GFP tagged Mbl proteins A (4), B (5), C (6) and D (7). Control lane 1 shows RT-PCR products when transfecting 0.5 μg *α-actinin* minigene and 1 μg of empty pGEM42. **D)** Quantification of the effect of CUG repeat RNA in α-*actinin* minigene isoform ratio showed a significant increase in α-actnC isoform and a decrease in α-actnA+B.

These results demonstrate that CUG repeat RNA misregulates *Drosophila α-actinin* pre-mRNA.

III.3. *Drosophila* Muscleblind isoforms regulate mouse *Tnnt3* minigene splicing in human cells.

troponin T genes are established MBNL1 targets. Human MBNL1, CUG-BP1 and ETR3 regulate human *cardiac Troponin T* transcript splicing [31]. *Mbnl1* knockout mice show splicing defects in both cardiac (*Tnnt2)* and skeletal muscle (*Tnnt3*) *Troponin T* [28]. *fast skeletal Troponin T* (*TNNT3*) splicing is altered in DM as patient myocytes show inclusion of an exon between exons 8 and 9 only present in developing normal muscles [28]. Inclusion of this exon reduces the sensitivity to Ca^{2+} of the resulting fast skeletal Troponin T [125]. We have shown that *muscleblind* mutant pupae have splicing defects in *tnT* mRNA. With these evidences, we decided to test the functional conservation between human and fly proteins expressing *Drosophila* proteins in a human cell context. We tested the ability of Muscleblind proteins to regulate the splicing of a vertebrate RNA, a *Tnnt3* minigene that has been described to be regulated by murine Muscleblind proteins [29]. In order to check if any of the *Drosophila* homologues of CUG-BP1 had activity on *Tnnt3* splicing, we also co-transfected the *Tnnt3* minigene with GFP-tagged Bruno proteins.

Semi quantitative RT-PCR after co-transfection of HEK293T cells with Muscleblind-GFP fusion proteins and murine *Tnnt3* minigene revealed strong activity of Mbl proteins in this assay. MblA, B and C modified the *Tnnt3* splicing pattern whereas no effect was detected when expressing MblD (Fig. R19A,C). Again, MblC was the isoform with the strongest activity

as one of the bands from *Tnnt3* completely disappeared. Western blot showed that MblC was again more stable or efficiently translated, which could influence the activity detected (Fig. R19B). No activity was detected for any of the Bruno proteins, although western blotting showed they were very unstable and the detection was very low, as it was for MblD isoform. The presence of the foetal exon in the medium band and the absence in the lower one was confirmed by sequencing. The upper band could not be sequenced.

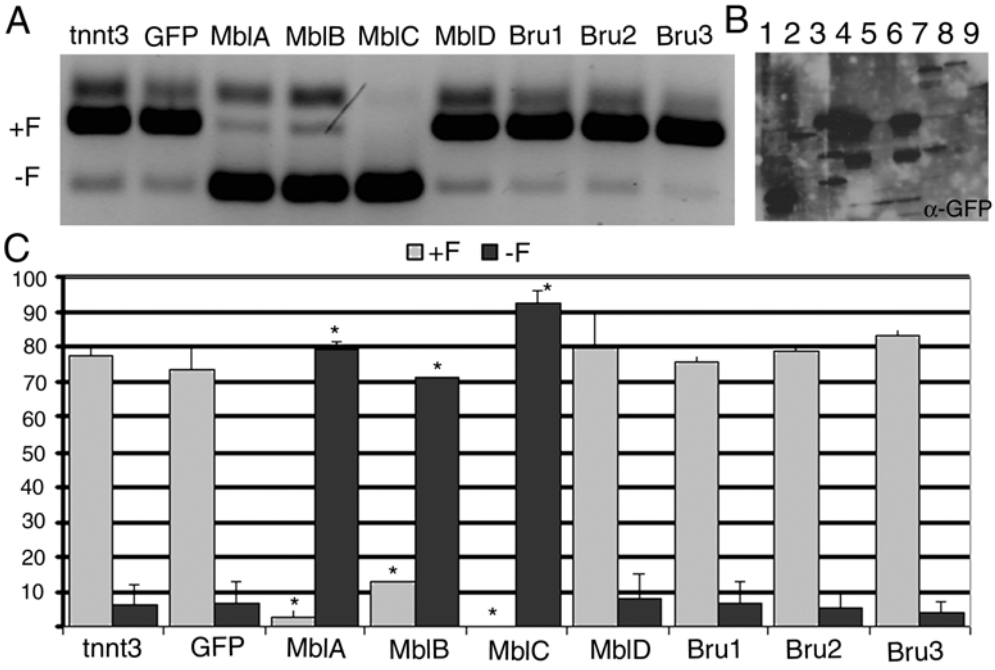


Figure R.19. Muscleblind proteins modify murine *TroponinT3* minigene splicing in a human cell culture assay. A) 2% agarose gel of RT-PCR products from HEK293T RNA extractions. +F indicates presence of foetal exon confirmed by sequencing. -F indicates absence. **B)** Western blot with α-GFP antibody of HEK293T protein extracts. 1: GFP vector; 2: MblA; 3: MblB; 4: MblC; 5:MblD; 6: MblC$^{\Delta SUMO}$; 7: Bruno1; 8: Bruno2; 9: Bruno3. **C)** Representation of average percentage of RT-PCR products including (+F) and excluding (-F) the foetal exon in three replicas except for MblB, which could be amplified only once.

Quantification was with 1DAdvanced Phoretrix software . Asterisks denote statistically significant differences compared to the effect of the expression GFP alone ($p<0.05$).

In summary, Muscleblind proteins influence exon choice of murine *Tnnt3* minigene in HEK293T cells. Expression of fly proteins in these cells led to the exclusion of *Tnnt3* foetal exon whereas no effect was detected for *Drosophila* Bruno proteins.

III.4. **Analysis of Mbl protein implication in cell death.**

Genetic data generated in our laboratory implicated *mbl* gene function with the apoptotic process. *reaper* and *Diap1* loss-of-function alleles respectively suppressed and enhanced a *mblC* over-expression phenotype in the *Drosophila* eye [126]. Furthermore, the over-expression of *mblC* in the posterior compartment of the wing imaginal disc (*engrailed-Gal4* driver) led to a lack of laminar tissue due to the activation of Caspase3. The genetic interactions with key apoptotic genes and the data from the analysis of the wing phenotype suggested *mbl* could be involved in the apoptotic process.

Over-expression of pro-apoptotic genes in *Drosophila* S2 cells activates programmed cell death, which can be measured in a cell viability assay [127]. We cloned myc-tagged Muscleblind isoforms into pIEI, a expression vector for *Drosophila* S2 cells. Monitoring cell viability by LacZ staining 48 h after co-transfection of myc-tagged Mbl isoforms and LacZ in pIEI vector showed an increase in cell death when expressing MblA, B and C. Increment in cell death was especially marked for MblB over-expression (Fig. R.20A), whereas MblD showed a weak effect in the opposite direction that, although not statistically significant, was reproducible in three

independent assays (Fig. R.20B). An experiment carried out fixing cells 24 h after transfection showed that the effect was quick.

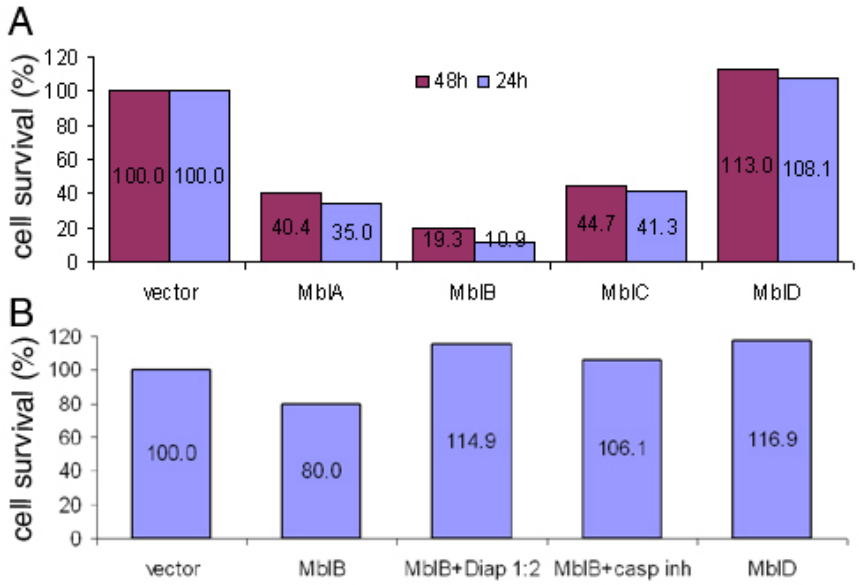


Figure R.20. Over-expression of Muscleblind proteins in *Drosophila* S2 cells increases cell death probably through apoptosis. The bar graph represents the percentage of cell survival observed when transfecting *Drosophila* S2 cells with Muscleblind isoforms tagged with myc in pIEI vector. **A)** Counting of surviving cells after 24 h and 48 h showed a statistically significant decrease in cell survival when expressing MblA, B and C (p-value<0.05). **B)** MuscleblindB induction of cell death was suppressed by co-transfection with anti-apoptotic Diap1 or addition of caspase inhibitor.

To confirm that the observed increase of cell death was due to apoptosis, we over-expressed anti-apoptotic protein Diap1 together with MblB or added a chemical caspase-inhibitor (Fig. R.20B). In order to detect suppression by Diap1, we reduced the amount of MblB-pIEI4 transfected (see M&M). Diap1 over-expression effectively suppressed MblB cell death activity, although the conditions should be optimized in order to ascertain the interaction (Fig. R20B and M&M section II.3.4.). Addition of a caspase

inhibitor also abrogated MuscleblindB effect on cell survival, thus suggesting that the activation of cell death by MblB is through the apoptotic pathway.

These results show that Muscleblind proteins activate cell death in an isoform specific manner. This cell death is probably due to activation of the apoptotic pathway although more experiments are necessary to fully support this proposal.

IV. **Molecular basis of isoform specific behaviour.**

Cell culture assays showed the differential activity of Muscleblind isoforms in regulating *Drosophila α-actinin* and murine *tnnt3* minigenes, and cell viability. Binding abilities of Muscleblind isoforms were also different in *in vivo* and *in vitro* studies. Muscleblind proteins share an important part of their N-terminal region but structural differences in their isoform-specific C-terminal ends may underlie biologically relevant functional differences as has been shown for other proteins. Just 16 amino acids of difference between *Drosophila alternative splicing factor* (dASF) and its human homolog (hASF) results in a different sub-cellular localisation, abolition of a phosphorylation site and a different splicing activity [128]. Specific carboxy-terminal regions can also result in specific protein-protein interactions that may influence RNA binding activity and regulate splicing activity [129]. As sub-cellular localisation and pot-translational modifications could be influencing the differential activities of Muscleblind isoforms, we performed cell culture assays in order to observe Mbl isoform localisation and the influence of the elimination of a putative sumoylation site.

IV.1. Analysis of sub-cellular localisation of Muscleblind proteins.

While the ultimate reason for functional differences between Muscleblind isoforms lies within the protein sequence, it is also possible that distinct cellular localisation influences protein activity. To test whether *Drosophila* Muscleblind protein isoforms were preferentially localised in the nucleus or cytoplasm, we transiently transfected COS cells with GFP-tagged Muscleblind proteins (Fig. 21A). In these experiments MblA appeared predominantly in the cytoplasm, whereas both MblB and MblC were enriched in the nucleus. These results were in agreement with the splicing activity of MblC in cell culture. Of note is the diffuse signal of GFP-MblD, probably arising from protein degradation, which precluded drawing any conclusion. Both MblA and MblC accumulated into cytoplasmic foci whereas MblB only occasionally did. These aggregates have been suggested to be a consequence of translational inhibition due to stress (Dr. J. Tazi, personal communication). To test if these aggregates were due to the amount of protein expressed by transfected cells or they were cell-type specific, we transfected HEK293T cells with a lower amount of plasmid DNA (Fig. 21B). MblA, B and C continued appearing in foci. The protein aggregates were much smaller and mainly nuclear for MblA and MblC. Therefore, isoforms showed the same sub-cellular localisation as in COS cells, suggesting that these were their physiological localisations *in vivo*.

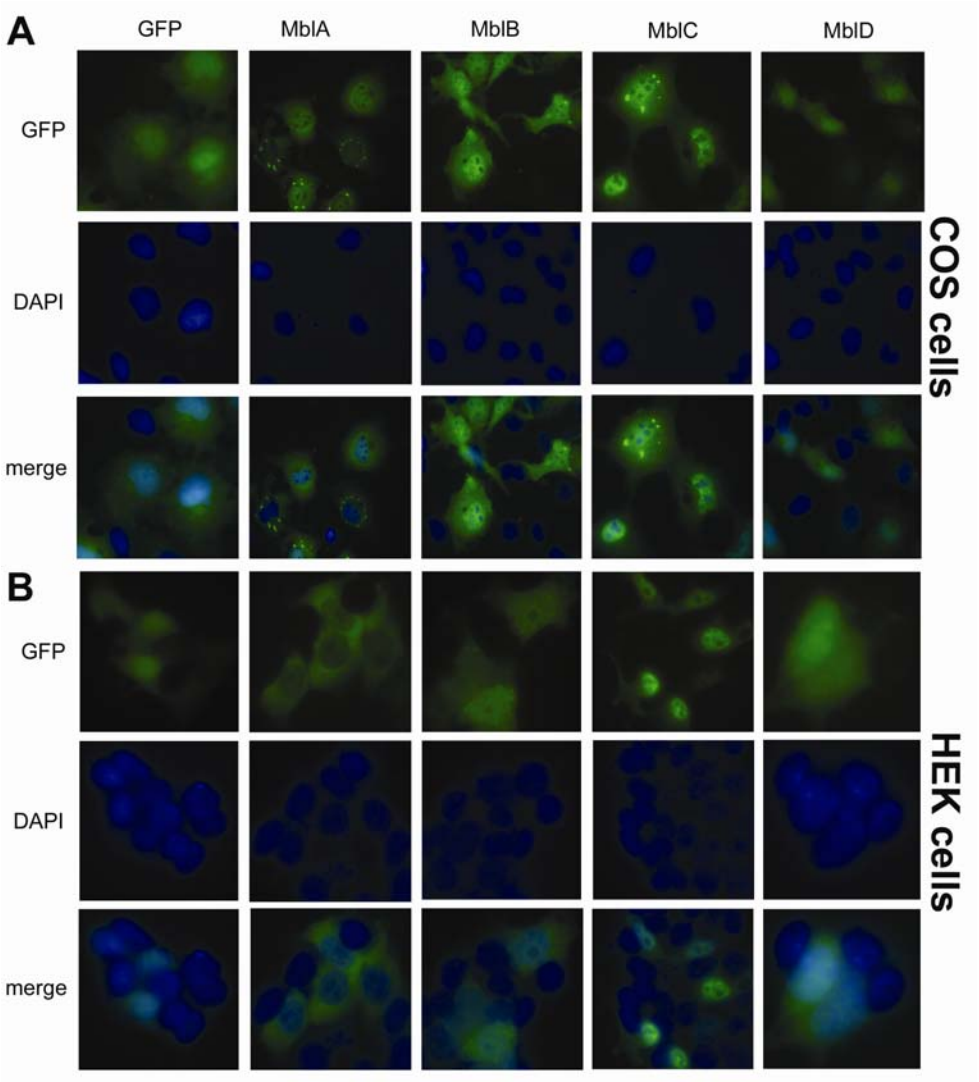


Figure R.21. Muscleblind proteins show different sub-cellular localisation in mammalian cell culture. A) 1 μg of GFP-fusion protein expression vector was transfected into COS cells. GFP fluorescence was detected from the GFP-tagged Muscleblind proteins and GFP control (green) and nuclei were counterstained with DAPI (shown in blue). MblA was found to be mainly cytoplasmic whereas MblC fluorescence was almost restricted to nuclei. MblB showed an intermediate behaviour, enriched in the nucleus but also intense in the cytoplasm, whereas MblD signal was diffuse and similar to that found in the GFP control, which suggested protein degradation . MblA, B and C formed big perinuclear aggregates. **B)**

300 ng of DNA were transfected into HEK293T cells. Aggregate size and number were reduced but sub-cellular localisation was similar to that found in COS cells.

These data coupled with the observation that Muscleblind distribution in adult *Drosophila* muscles, when only *mblC* and *mblD* transcripts are detected [45], is mainly in nuclear foci [83], suggest that MblC could normally be located in nuclear aggregates. The results also show that a high proportion of MblC is localised in the nucleus, whereas MblA is mostly cytoplasmic. This could be influencing the ability to regulate splicing in the cell culture assays previously shown (Figs. R.17 and R.19) and further supports functional diversification between Muscleblind isoforms.

IV.2. Analysis of the functional relevance of a putative MblC sumoylation site.

All four MBNL1 zinc fingers, and the linker in between them, were shown to be necessary for binding to CUG repeat-containing RNA in a yeast three hybrid assay [43]. CCCH zinc finger motifs are the only domains of Muscleblind proteins associated with a function. Nevertheless, we have found *Drosophila* Muscleblind proteins sharing the first amino terminal 179 amino acids, including both zinc fingers, show differences in sub-cellular localisation and splicing activity. We next sought the molecular basis for such differences.

in silico studies developed in our laboratory on MblC specific sequence identified a putative sumoylation consensus sequence (FKRP) conserved in *C. elegans*. (Fig. R.22A). Sumoylation is an ubiquitin-related posttranslational modification that regulates protein activity. Sumoylation does not typically lead to protein degradation. Instead, it usually regulates

intracellular protein localisation [130, 131]. In order to test whether the conserved FKRP motif was involved in MblC function, site-directed mutagenesis was performed to disrupt the site. The lysine residue located at position 202 was mutated to isoleucine (K202I; Fig R.22A). We tested the mutated protein (hereafter referred to as MblC$^{\Delta sumo}$) in the functional assays we performed before with wild type MblC.

COS and HEK293T cells were transfected with pEGFP-N3-MblC$^{\Delta sumo}$ to assess the sub-cellular localisation of mutated protein. Expression of GFP-tagged MblC$^{\Delta sumo}$ protein in COS cells showed a frequency of protein foci higher than when wild type MblC was expressed (Fig. R.22B). To check whether MblC aggregation ability was enhanced by the elimination of the putative sumoylation site, we also transfected HEK293T cells diminishing the amount of transfected vector expressing GFP-tagged wild type and mutated MblC. Foci formation when transfecting wild type MblC-GFP was severely reduced but MblC$^{\Delta sumo}$-GFP still formed prominent aggregates, mostly perinuclear (Fig. R.22B').

COS cells were also used to test the ability of MblC$^{\Delta sumo}$ to co-localise with expanded CUG repeat RNA in ribonuclear foci. Co-transfection of pEGFP-N3-MblC$^{\Delta sumo}$ and *DMPK* 3'UTR (CUG_{197}) in COS cells showed no disruption of protein co-localisation with expanded CUG ribonuclear foci (Fig. R.22C). The splicing factor activity of the mutated protein was also tested. Mutation of FKRP site did not affect Muscleblind activity on *Drosophila α-actinin* minigene (not shown). When co-transfecting the protein expression vector with mouse *Tnnt3* minigene, the splicing activity of the protein was also maintained (Fig. R.22D).

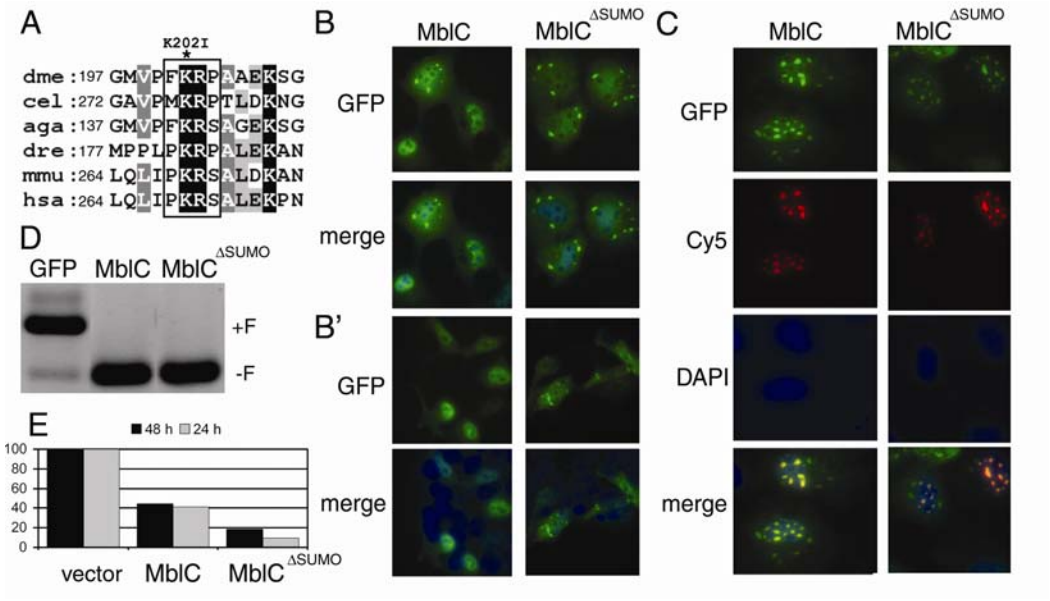


Fig.R.22. Elimination of a conserved putative sumoylation site increases protein aggregation and cell death-inducing activity of MblC. (A) *Drosophila* MblC sumoylation site (dmel) is conserved in *C. elegans* (cel). Sequence is partially conserved in *Anopheles gambiae* (aga), *Danio rerio* (dre), *mus musculus* (mmu), and humans (hsa). Sequence change resulting from site-directed mutagenesis is denoted by an asterisk. Figure adapted from alignment by J.M.Fernandez (Depto. Genética Facultad de Biología U.V.) **B)** COS cells transfected with 1 μg of GFP-tagged proteins showed a higher frequency of protein aggregates when the site was mutated. **B')** HEK293T cells transfected with 300 ng of GFP-tagged proteins showed small nuclear aggregates in wild type MblC transfections whereas MblC$^{\Delta SUMO}$ still aggregated in big foci also found in cytoplasm. FKRP site disruption did not interfere with co-localisation with CUG ribonuclear foci **(C)** or *Tnnt3* splicing regulation in cell culture **(D)**. Statistical significance (p-value<0.01) was found for both +F and -F bands when co-expressing MblC$^{\Delta SUMO}$ and the *tnnT3* minigene compared to the empty vector, showing no change in the splicing activity of the protein in this system when the site was mutated. **E)** MblC$^{\Delta SUMO}$-GFP induced twice more cell death than wild type MblC tagged with GFP in *Drosophila* S2 cells. The increment was observed both 24 and 48 h after transfection.

We finally checked the influence of the K202I change in the activation of cell death. Interestingly, pro-apoptotic activity of Muscleblind was enhanced by elimination of the putative sumoylation site. When expressing myc-tagged MblC$^{\Delta sumo}$ in *Drosophila* S2 cells, survival of cells was reduced to 20% after 24 h compared to controls transfected with the empty vector. In contrast, samples transfected with wild type protein showed a survival rate of 40% (Fig. R.22E). Experiments carried out in our laboratory later on confirmed the increase of cell death activation when removing the FKRP site.

In summary, our results suggest a role for the FKRP site, possibly sumoylated, in Muscleblind C activity regulation. The mutation of the site does not alter RNA binding to CUG repeats or tnnT3 splicing but enhances formation of protein aggregates and the activation of cell death in cell culture.

Discussion

Yo no puedo cambiar la dirección del
viento, pero si ajustar las velas
para que me lleven a mi destino
-James Dean-

Since *muscleblind* gene was first described in *Drosophila*, a function in RNA metabolism was proposed for the proteins coded because of the zinc finger motifs [44]. Studies carried out later on showed that human MBNL proteins have the ability not only to bind to RNA molecules but also to modify their splicing [32] and subcellular localisation [36]. Muscleblind proteins are being implicated in an increasing number of diseases. Expression levels of human MBNL2 are altered in several tumours [36] and MBNL1 is upregulated in schizophrenia [54] and sporadic idiopathic pulmonary arterial hypertension [55]. MBNL1 function is also impaired in Myotonic Dystrophy [28, 29] and a similar mechanism has been proposed to participate in Huntington Disease Like 2 pathogenesis [53]; also a fly model of SCA8 genetically interacted with *Drosophila muscleblind* loss-of-function alleles [52]. Mouse and fly *muscleblind* mutants present defects in muscle and photoreceptor differentiation [28, 41, 44]. *Drosophila muscleblind* mutant defects [46] and expression patterns of fly and vertebrate *muscleblind* genes [39-41] suggest that it might be also implicated in nervous system development.

Drosophila presents advantages to be used as model organism to study Muscleblind function as it presents a single gene that generates four transcript isoforms, whereas mouse and human present three MBNL genes and at least 15 confirmed transcript isoforms [38, 42, 43]. *Drosophila* Muscleblind proteins share their N-terminal region. MblA, B and C share 179 amino acids, including two zinc finger motives. This is the only putatively functional domain described in these proteins and the domain that defines Muscleblind proteins as part of the CCCH zinc finger domain family of proteins. MblD, the shortest isoform, shares just 63 amino acids with the others and it has only a zinc finger. All four proteins present

specific carboxy-terminal sequences that might modify the binding activity of the zinc fingers or provide other functionality to the protein.

A guiding theme throughout my PhD work was to better understand the molecular function of *Drosophila* Muscleblind proteins and their relevance to myotonic dystrophy. To this end, I initially looked for myotonic dystrophy-like defects in a hypomorphic *muscleblind* allelic combination and a CUG toxicity model in flies. I found that *muscleblind* mutant flies showed defects in alternative splicing of defined pre-mRNAs. These potential Muscleblind targets were also altered in flies expressing CUG repeats, thus suggesting conservation of the mechanism of pathogenesis in the fly. Then, I analyzed the RNA-binding properties of Muscleblind, both to physiological and pathogenic CUG repeat containing RNA. I showed the conservation of direct binding to targets through the zinc fingers *in vitro,* and the interaction with CUG containing RNA in cell culture. I confirmed Muscleblind activity as splicing factor with minigene splicing assays in cell culture and I showed the ability of MuscleblindA, B and C to activate cell death, probably through the apoptotic pathway. I also performed experiments in order to analyse the basis of isoform-specific activity detected for Muscleblind proteins. I described Muscleblind protein isoforms have different subcellular localization, and I discovered that mutation of a conserved putative sumoylation motif in MuscleblindC protein activates the pro-apoptotic activity of the protein in cell culture.

With these results, we conclude that the function of Muscleblind proteins and also the mechanism of DM pathogenesis are well conserved between vertebrates and flies (both aspects will be discussed deeply below). *Drosophila Muscleblind* function, as vertebrate *MBNL* function, is necessary

for the proper regulation of alternative splicing of specific genes. Furthermore, *Drosophila* Muscleblind proteins, as the vertebrate MBNL proteins, can modify the splicing pattern of transcripts in cell culture experiments. In the other hand, the expression of CUG repeat containing RNA in flies alters the splicing pattern of *muscleblind* targets. This opens the possibility to obtain important data from *Drosophila* with biomedical relevance. In summary, this work shows that Drosophila is a good model to study myotonic dystrophy. Both Muscleblind mutant flies and flies expressing CUG repeat containing RNA are useful tools not only in the research to clarify the role of Muscleblind proteins in DM but also in the analysis of other factors involved in DM and the characterisation of Muscleblind protein function.

1. *mbl^{E27}/mbl^{k7103}* mutants, a model to study *muscleblind* function in late development.

An accurate counting of hypomorph flies arriving to the adult stage gave a 70% of viable flies (Fig. R.3). Null *muscleblind* mutant embryos show defects in sarcomere organisation and tendon matrix reduction [41]. Somehow, these or other defects not yet described are too severe for the embryo to survive. They die as larvae inside the chorion. Contrarily, *mbl^{E27}/mbl^{k7103}* allelic combination might be giving enough *muscleblind* function for the eggs to undergo a proper embryonic development, but fails in later stages as the adult flies show defects that can be severe and even the flies that look externally fine, have reduced viability. Then, this hypomorphic allelic combination emerges as a useful tool to analyse *muscleblind* function in late development and in consequence, a model to study Muscleblind implication in myotonic dystrophy.

During *Drosophila* wing development, the dorsal and ventral wing epithelia fuse together by highly specific cell-cell adhesions, and defects in this process result in blistered wings. The blistered wing phenotype observed in the mbl^{E27}/mbl^{k7103} mutants (Fig. R.1) suggests that Muscleblind may participate in such a cell adhesion process. During *Drosophila* wing maturation, components of extra-cellular matrix (ECM) are expressed that bond the dorsal and ventral cuticular surfaces of the wing following migration of the cells [132]. The origin of ECM components is not well understood although it has been proposed that hemocytes are implicated in this process. Hemocytes are known to mediate ECM formation between dorsal and ventral epithelia during prior pupal appositions of the wing epithelia [133, 134]. Hemocytes are also known to express several tendon matrix molecules [135]. *muscleblind* null mutants where described to have a severe reduction of the extracellular tendon matrix (TM) at indirect muscle-epidermis attachments [41]. Muscleblind expression analysis by immunohistochemistry showed no expression in hemocytes [41]. Muscle tissue has not been described to secrete any component of TM and the study of *muscleblind* mutant clones in *Drosophila* photoreceptors showed *muscleblind* function being cell-autonomous. So, if *muscleblind* mechanism of action is the same in muscle and eye development, one wouldn't expect the secretion of TM components to be dependent on *muscleblind*. Thus, the reduction in TM might be caused by the lost of anchoring of the components to proteins present in muscle membrane which presence or functionality require *muscleblind* function.

Integrin is well known as a central molecule that mediates adhesion between wing layers [136, 137]. The maintenance of muscle attachment to

epidermal tendon cells of *Drosophila* also depends on the PS1 and PS2 integrins. Muscles and tendon cells express complementary pairs of heterodimeric integrins (PS1 and PS2), composed of a common β-subunit and different α-subunits such as α_{PS1} in tendon cells, and α_{PS2} in muscles (reviewed in [138, 139]). Whereas mutations in β_{PS} subunit cause severe detachment phenotypes, mutations in *multiple edematous wings* (*mew*) and *inflated* (*if*), which encode the α_{PS1} and α_{PS2} subunits, respectively, result in either weaker or no embryonic muscle detachment phenotype [140-142]. Moreover, hypomorphic alleles of integrin mutants result in wing blisters in surviving adult flies, a phenotype likely to be caused by the loss of adhesion between dorsal and ventral wing epidermal blades which express PS1 and PS2, respectively [138].

mbl^{E27}/mbl^{k7103} adult flies showed defects in wing lamina adhesion (Fig. R.1). 15% of the mutant flies that arrived to adult stage showed wing blisters and over a 20% of the escapers showed an acute phenotype have unexpanded wings in which the two surfaces are completely detached. The loss of *PS integrin* function in the wings, by loss of function of either the gene itself or the adaptor molecules that mediate attachment of extracellular matrix -integrin complexes to cytoskeleton, causes the formation of a fluid-filled blister [137, 143]. So, the absence of cell adhesion molecules or the presence of inadequate isoforms with different functionality may also generate lamina detachment. Interestingly, *viking*, a gene encoding the α2 chain of collagen IV, which together with α1 chain form the basal membranes that stabilise cell-matrix interactions in *Drosophila* [144] was detected to interact with the expression of expanded CUGs in *Drosophila* [85]. These flies reproduce the main features of Myotonic Dystrophy and genetically interact with *muscleblind* function.

Another molecule implicated in cell-cell adhesion, the kinase coded by *turtle*, interacted with the over-expression of *muscleblindC* in the adult eye [126]. Then, mbl^{E27}/mbl^{k7103} mutants are a good scenario to analyse the mechanism by which *muscleblind* function is required for wing development and determine *muscleblind* implication in extra-cellular matrix deposition or anchoring.

We also detected splicing defects in *Drosophila troponinT* mRNA specific of *muscleblind* mutant pupae (Fig. R.10). Human MBNL1, PTB, CUB-BP and ETR-3 (the last two homologs to *Drosophila bruno* genes) cooperate to regulate the splicing of *cardiac TroponinT* (*cTNT*, [32]). Both *cTNT* and *fast skeletal TroponinT* (*TnnT2*) are mis-spliced in *Muscleblind-like1* knockdown mice ($Mbnl1^{\Delta 3/\Delta 3}$). Interestingly, *Drosophila TroponinT* mutation was described to generate an upheld phenotype [145], typical of disrupted IFMs, also found when expressing expanded CUG repeat RNA in a muscular pattern [85]. *Drosophila TroponinT* has been described to undergo tissue-specific alternative splicing, also regulated during development. Interestingly, we detected an alteration in the ratio of inclusion of the exon 3, encoding a glutamic acid-rich domain homologous to the fetal exon of *cTNT* that is regulated by human MBNL1 [121]. This exon is only absent in the *troponinT* isoform expressed in TDT and IFM muscles and probably confers specific functional properties as the foetal exon does in humans [125]. This defect was already detectable in heterozygous flies carrying the strong allele mbl^{E27} pointing that half dose of *muscleblind* is not enough to perform all the functions required for a normal development. Despite of the presence of this molecular defect, heterozygous flies do not present any external phenotype and no viability problems have been observed. So, in laboratory growth conditions, these flies can cope with this defect and

maybe others not yet described. This gives us a situation in which *muscleblind* dose can be somehow monitored as we have a gradation of *muscleblind* loss of function. It would be interesting to correlate the defects found in the fly model of RNA toxicity (see below) with those found in mbl^{E27}/mbl^{k7103} hypomorphs to approximately measure the level of *muscleblind* function that remains in CUG repeat RNA expressing flies in vivo.

2. Conservation of Muscleblind molecular function.

Splicing machinery is one of the most complicated macromolecular complexes in the cell, with more than 300 proteins described to be implicated in splicing [146] and a highly dynamic composition [147] that requires numerous protein-protein and protein-RNA interactions to be accurate. MBNL1 is an alternative splicing factor that binds to physiological targets through a YGCU(U/G)Y sequence [31, 32]. We found splicing defects in *muscleblind* mutants (Fig. R9, 10) that pointed to the conservation of the splicing activity in flies. The demonstration arrived when over-expression of *Drosophila* Muscleblind proteins modified *α-actinin* minigene splicing in cell culture (Fig. R.16). Insect and vertebrate cellular environments are known to be very different. Interestingly, *Drosophila* Muscleblind proteins were active in vertebrate cell culture and they could regulate *α-actinin* minigene splicing. Even stronger reinforcement of the conservation is the fact that *Drosophila* Muscleblind proteins were active on the murine *TroponinT3* minigene in human cell culture (Fig.R.18). They were able to interact not only with vertebrate proteins required for the splicing but also with vertebrate sequence elements. These data reinforce previous *in vivo* conservation data generated by rescue of *muscleblind*

mutant background by MBNL1 [107]. Very little is known about the factors that are required for MBNL to recognise the RNA and to be active as splicing factor but they were similar enough in flies for MBNL1 to perform a good part of *muscleblind* functions.

α-actinin isoform *C*, typical of larval muscle, was severely reduced in *muscleblind* mutants. The correspondent increase mainly in the non-muscle *α-actinin* isoform *A,* and also in *α-actinin* isoform *B*, typical of adult muscle, was detected (Fig.R.9). MBNL1 is implicated in a developmental switch in alternative splicing of *cTNT* E5 in vertebrates in coordination with CUG-BP1, ETR-3 and PTB [31]. Whereas CUG-BP1 and ETR-3 favour E5 exclusion and their expression is down-regulated during heart development, MBNL1 and PTB repress E5 inclusion and their expression is maintained in the adult heart. As result of the regulation of splicing factors levels, E5 is excluded in the developing heart but it is included in the adult tissue. The splicing alterations in flies were only found when all samples were collected in parallel and strictly controlling the egg laying and development times, indicating that we might be working in a developmental stage in which there is a switch in splicing regulation. Also the synchronisation of the pupae was essential to get reproducible results. In any case, what we observed in the fly seems to be a more complicated equilibrium than that described in *cTNT*. The splicing alteration that we found in *α-actinin* splicing is not only a change from larval to adult isoforms but also from muscle to non-muscle isoform. The splicing alteration in *TroponinT* found in hypomorphic *muscleblind* mutant pupae was a change from hypodermic muscle transcript to the transcript specific of TDT and IFM muscles (Fig.R.10). Then, it is again an alteration in cell-specific alternative splicing regulation although cloning and sequencing analysis should be

performed to discard a change in inclusion of the adult E4 that would mean an additional adult to larval switch. Interestingly, splicing defects in mutants were observed coinciding with the two expression peaks of *muscleblindC* (Fig. I.5). It would be interesting to check whether *mblC* transcript shows a tissue-specific expression pattern in the adult fly to better understand the mechanism of control of this cell-type specific splicing events. It is also worth noting that the changes in exon usage are not from zero to 100% or vice versa, pointing to the implication of other factors in the regulation of these splicing events. Identification of those factors and a detailed description of their expression patterns will be necessary to understand the complete mechanism of *α-actinin* and *TroponinT* splicing in *Drosophila*.

The RNA-protein interaction mechanism of Muscleblind proteins seems to be conserved between human and flies too. Human MBNL proteins are RNA binding proteins that directly interact with RNA, both with their physiological targets and the expanded CUG containing RNA causing DM [32, 51]. A yeast three hybrid study showed that MBNL1 specifically binds to CHHG sequences (where H is A, U or C) [43]. It also showed that the four zinc finger domains and the linker between them are necessary to interact with expanded CUG repeats although the last zinc finger might be not needed to interact with the CCUG repeats that originate DM2. Analysis of splicing defects in *muscleblind* mutants revealed two potential targets of Muscleblind proteins: *α-actinin* and *TroponinT*. Bioinformatics analyses of genomic sequences showed a consensus sequence for MBNL1 binding close to *Drosophila TroponinT* exon E3 and a cluster of seven overlapping MBNL1 binding sequences close to *α-actinin* exon E7 3' end. The preliminary binding results, point to the binding between Muscleblind zinc fingers and a fragment of the *α-actinin* mRNA containing MBNL1 binding

sequences (Fig.R.13, 16). The degradation of the protein advises against any strong conclusion about the specificity of these studies. So, more experiments, either expressing fragments of the protein to obtain soluble and stable expressed fragments, or expressing the protein in a different system (ideally *Drosophila* S2 cells) should be performed to confirm the results obtained.

Seven overlapping MBNL1 consensus binding sequences were found in *α-actinin* I6 (see Fig. R.9 for nomenclature). N terminal fragments of MblA appeared to interact with a fragment of 161 bp containing these sequences and not the adjacent fragment with none of these sequences (Fig.R.13, R.16) in *in vitro* crosslinking assays. Other techniques as crosslinking immunoprecitation (CLIP) or systematic evolution of ligands by exponential enrichment (SELEX) should be considered to definitely prove that Muscleblind binding sequence is conserved between humans and flies. Further experiments are needed to probe that Muscleblind splicing activity on *α-actinin* requires the binding to mRNA through these sequences. Assays testing protein functionality when mutating these sequences and when introducing these sequences in an unrelated context would tell whether those nucleotides are required and/or sufficient for the protein to bind and to be active. It will also be interesting to analyse how many of these sequences are necessary for Muscleblind proteins to bind. As no big cluster has been found in *TroponinT* mRNA, the other target we identified, it is possible that just one sequence is enough for Muscleblind to bind and be active.

Also the interaction with CUG repeat containing RNA is conserved *in vivo* as shown by the localisation of Muscleblind proteins with RNA foci in

human cells (Fig.R.11,12). MBNL proteins have been described to co-localise *in vivo* with ribonuclear foci formed by *DMPK* 3'UTR RNA in many tissues from DM patients [38, 51, 73-75, 148, 149]. Human GFP-tagged MBNL proteins expressed in DM fibroblasts form big aggregates that co-localise with RNA foci containing CUG expanded repeats [38, 51, 148]. Similarly, when we co-expressed *Drosophila* Muscleblind proteins and *DMPK* 3'UTR RNA with around 180 CUG repeats, we found big aggregates that perfectly co-localised with the RNA foci in COS and HEK cells. MBNL proteins have also shown ability to interact with other trinucleotide repeat sequences. MBNL proteins showed co-localisation with RNA foci formed by CAG repeats [78] and yeast three hybrid analysis defined a binding sequence CHHG [43]. Interestingly, our three hybrid preliminary results show that Muscleblind proteins might be able to bind CAG repeats as well. In conclusion, sequestration of Muscleblind proteins by CUG repeats is conserved *in vivo* and studies on *Drosophila* Muscleblind RNA binding activity can provide information of biomedical relevance.

3. Functional interaction with Bruno proteins.

We have seen that the lack of *muscleblind* function does not result in the complete absence of a transcript isoform (Fig. R.9, R.10), thus suggesting that other factors might be implicated in the regulation of *α-actinin* and *TroponinT* splicing. It is also remarkable that expression of CUG repeat containing RNA in the fly altered splicing of *α-actinin* but the defects were different from those in *muscleblind* mutant flies. Thus, CUG expansion might be affecting other factors implicated in *α-actinin* splicing regulation. Human MBNL proteins, CUG-BP1, ETR-3 and PTB act on *cTNT* pre-mRNA to regulate a developmental switch in its isoform ratio [31]. CUG-BP1 and

ETR-3 belong to the CELF family of vertebrate proteins, homologues to *Drosophila* Bruno proteins. *bruno1* (*bru* or *aret*) function was described to modify splicing of a *Dscam* minigene in S2 cells [150] although main *aret* functions have been related to translation repression in *Drosophila* oocyte [151]. *bruno3* conserves the ability to bind to EDEN sequences bound by CUG-BP1 and it has been suggested to be its ortholog by sequence analysis [152]. A genetic screen for dominant modifiers of a *muscleblindC* over-expression phenotype performed in our laboratory identified *aret* as an interacting gene as the loss of function of *aret* enhanced *mblC* over-expression phenotype. Alleles of the other two *bruno* genes were tested but neither *bru2* nor *bru3* genetically interacted. I checked all three Bruno proteins for their ability to modify splicing of *Drosophila* Muscleblind targets under the same conditions where Muscleblind proteins showed activity. I could not detect any effect of *Drosophila* Bruno proteins on splicing of Muscleblind targets *α-actinin* and *TroponinT*. The low levels of protein found when over-expressing Bruno proteins in cell culture could prevent us from detecting any activity and also the tnnT3 minigene assay could not be sensitive to Bruno activity if it is antagonistic to Muscleblind. It would be interesting to perform a co-transfection assay in which Bruno over-expression was tested to antagonise the effect obtained when over-expressing Muscleblind isoforms.

Alternatively, the antagonism could be conserved in a different process. Muscleblind could be involved in *bruno1* dependent processes in the oocyte different to splicing or *bruno* could be regulating splicing of other Muscleblind targets. CUG-BP1 and MBNL1 act antagonistically on *cardiac TroponinT* (*TNNT2*) mRNA to regulate its splicing [31]. Both proteins also have the ability to modify the splicing of *Insulin receptor* mRNA in cell

culture [32, 33, 80] and transgenic mice expressing the active form of CUG-BP1, show the same splicing defects in *Tnnt2* and *Clcn1* Mbnl1 knockdown [28, 81]. DM patients show increased levels of CUG-BP1 and the sequestration of MBNL1 by *DMPK* 3'UTR is supposed to generate MBNL loss of function (reviewed in [50, 108, 153, 154]). Thus, CUG-BP1 and MBNL proteins seem to have related functions in vertebrate muscular tissue. *Drosophila bruno1* function is well understood in the oocyte [151, 155] but its expression pattern makes it difficult to expect an implication in muscle development as its RNA is only intensely detected in embryos from 0-2 h AEL and weakly detected in pupae and adult flies [151]. Little is known about *bruno3*. It has been shown to conserve binding activity to EDEN element in RNA and it has been proposed as the CUG-BP1 orthologue because of that [152]. The expression data available in the BDGP gene expression database show a late embryonic expression centred in the sensory complexes, the ventral nerve cord and visceral musculature. Then, a coordinated action with *muscleblind* in the regulation of skeletal muscle-specific transcripts is also improbable. Finally, no studies have been published about *bruno2*. In summary, although other Bruno isoforms should be tested to discard any activity on *Drosophila α-actinin* and *TroponinT* mRNA alternative splicing regulation, it is unlikely that the antagonism of vertebrate MBNL and CUG-BP1 is conserved in flies in the regulation of skeletal muscle development by alternative splicing.

4. Conservation of Muscleblind sequestration by expanded CUG containing RNA in flies.

Human MBNL proteins are the only factors described to bind CUG repeat expansions originating myotonic dystrophy in a length-dependent

manner, observation that provided a molecular explanation for the observed correlation between expansion length and severity of the disease [51]. MBNL proteins co-localise with ribonuclear foci containing expanded CUG repeat tracts in fibroblasts, muscle cells and neurons of Myotonic Dystrophy1 patients [51, 73-75]. They have also been shown to co-localise with expanded CCUG repeat-containing RNA in DM2 [76]. We found that all *Drosophila* Muscleblind isoforms that contain two zinc fingers, MuscleblindA, B and C, co-localise with CUG repeat-containing ribonuclear foci, although preliminary yeast three hybrid results showed differences in affinity. These results support the conservation in *Drosophila* of protein-RNA interaction of Muscleblind proteins with pathogenic expanded RNAs. Then, although no expanded RNA has been described to occur naturally in *Drosophila*, the expression of a transgene carrying long CTG repeats in the fly probably interferes with Muscleblind function as it does in Myotonic Dystrophy.

α-actinin mRNA splicing is altered when expressing CUG repeat RNA, both in cell culture and in the fly (Fig. R.9, R.17). We have shown that Muscleblind proteins regulate *α-actinin* mRNA alternative splicing and *muscleblind* genetically interacts with the rough eye phenotype generated by expression of expanded CUG containing RNA [85]. These results clearly indicate that there is a functional relation between *Drosophila muscleblind* and toxic CUG repeat containing RNA. The reproduction of main myotonic dystrophy features in flies expressing 480 interrupted CUG repeat RNA (muscle histological defects, RNA foci formation; Muscleblind co-localisation with RNA foci; splicing defects) demonstrates that toxic RNA pathogenic pathway is conserved in flies. The different defects found in *α-actinin* mRNA splicing in *muscleblind* mutant flies and flies expressing the

CUG repeat containing RNA indicate that this model might be sensitive to other factors which will be important to characterise to dissect myotonic dystrophy.

5. Novel *muscleblind* targets affected by expression of expanded CUG repeat RNA.

We found defects in the alternative splicing of two muscle transcripts in *muscleblind* mutant flies by following a candidate gene approach. We detected defects in alternative splicing of *Drosophila α-actinin* mRNA in *muscleblind* mbl^{E27}/mbl^{E16} mutant embryos 16-18 h AEL. α-Actinin is the major component of vertebrate and insect Z-bands. Transcripts from the unique *Drosophila α-actinin* gene undergo alternative splicing to generate both non-muscle (*α-actininA*) and muscle-specific (*α-actininB* in adult muscle and *α-actininC* in larval muscle) isoforms [156]. These isoforms are cell type-specific and presumably bind Actin differently as alternative exons encode a peptide located at the junction of the Actin binding domain and the first central repeat. I found that *Drosophila muscleblind* mutant embryos present a reduction of the larval muscle isoform *α-actininC*. Interestingly, *Drosophila α-actinin* mutants present disruption of Z discs and muscle insertions, and also muscle paralysis [120, 156], similarly to *muscleblind* mutant phenotype. Then, *α-actinin* splicing impairment might be contributing to the muscle phenotype of *muscleblind* mutant embryos. Our results are consistent with a recent publication in which defects in splicing of *α-actinin* and the homolog to the vertebrate *ZASP*/*Cypher*, *CG30084,* have been shown [157]. *ZASP* encodes a protein component of Z-bands in humans [158]. Then, the impairment of muscle development observed in

muscleblind mutant embryos is probably the consequence of the addition of several isoform ratio alterations.

Drosophila α-actinin splicing regulation was also altered in the fly model of expanded CUG repeat RNA toxicity. We found the isoform *A* to be reduced using two different Gal4 strains and two different CTG transgenes. Thus, the defect is clearly specific, but it is different from what we found in *muscleblind* mutant embryos. Two main reasons could explain this difference: first, the amount of *muscleblind* function that is compromised might be different in CUG repeat containing RNA expressing flies than in *muscleblind* mutants; second, other factors might be involved. The discrepancy could be reflecting the different sensitiveness of the fly to *muscleblind* function. The reduction of *muscleblind* function by CUG expanded transcripts might not be the same than in the mutants tested. Furthermore, *muscleblind* function is altered during the whole embryonic development and all tissues in *muscleblind* mutants whereas CUG repeat RNA is only expressed in the *Mhc* or *da* pattern. Flies expressing CUG repeats with the drivers used are viable whereas strong alleles of *muscleblind* are embryonic lethal, thus suggesting that most of *muscleblind* function is being performed properly in these flies. It would be interesting to check whether the different alleles or transheterozygous allelic combinations show the same defect. On the other hand, expanded CUG repeat RNA might be also affecting other factors implicated in *Drosophila α-actinin* splicing. In this concern, Bruno proteins were good candidates to be acting on splicing regulation of *muscleblind* targets. However, no effect of Bruno proteins on *Drosophila α-actinin* and mouse *Tnnt3* minigenes was detected (see section 3 for discussion).

α-actinin mRNA splicing was sensitive to expanded CUG containing RNA expression in cell culture and also in flies. Regarding the biomedical relevance of *α-actinin* mis-splicing, it is important to keep in mind that, even though α-Actinin is the major component of Z discs, a single *α-actinin* gene provides both muscle and non-muscle specific isoforms in *Drosophila*. Thus, its contribution to tissues other than muscle might be also relevant. Moreover, *muscleblind* expression is detected in tissues other than muscle during embryonic development. Rescue assays showed that the expression of Muscleblind isoforms in a general pattern rescues more than the expression in muscular tissue alone [45], thus indicating that *muscleblind* might be controlling isoform ratio of other mRNAs in the different tissues where it is expressed. Muscleblind dependent isoform ratio of neurogenic proteins could explain the increase in rescuing ability when expressing Muscleblind proteins in a general embryonic pattern and also the defects found in development of peripheral nervous system and neural photoreceptor structures in *muscleblind* mutants [44, 109]. It has been suggested that *α-actinin* defects in *muscleblind* mutant embryos were not conserved in DM1 patients because human *α-actinin2* splicing was normal in DM1 samples [157]. However, PTB, CUB-BP and ETR-3 regulate inclusion of the smooth muscle and non muscle exons in rat *α-actinin* [119]. Recent data suggest that MBNL1 regulates this splicing event ([159] Dr. Kino, personal communication), and we showed that human MBNL1 was able to modify *Drosophila α-actinin* minigene splicing. Therefore, we suggest that a more exhaustive study is required to determine which human *α-actinin*, if any, is regulated by MBNL proteins and hence might be altered in DM patients.

Splicing of vertebrate *troponinT2* and *3* mRNAs is regulated by MBNL1 and is mis-regulated in Myotonic Dystrophy patients [28, 31, 32]. We found that *muscleblind* function is required for the proper splicing regulation of *Drosophila TroponinT* mRNA (Fig. R.10). Splicing of this transcript was also mis-regulated in pupae expressing expanded CUG containing RNA. A dominant effect of the Gal4 strain was detected when analyzing *troponinT* mRNA splicing in flies carrying *Mhc-Gal4* insertion, which indicated an unspecific alteration of this splicing event. Analysis carried out afterwards in our laboratory showed that flies expressing 60 CUG repeat containing RNA under the control of the same *Mhc-Gal4* driver have a weaker splicing alteration meaning that the defect is probably specific (Monferrer L. and García-López, A., unpublished).

6. Muscleblind isoforms are functionally distinct.

One of the aims of my study was to identify protein motifs implicated in Muscleblind function. The four *Drosophila* Muscleblind isoforms give a scenario with enough variability as to identify parts of the protein implicated in different activities, much like in a deletion analysis. MblA, B and C present two complete zinc fingers and, although sharing an important part of the protein sequence, they present specific carboxy-terminal regions. MblD only presents a zinc finger as the second one is truncated before the conserved Histidine. These structural differences might be reflecting biologically relevant functional differences as it has been shown to occur in other cases [128, 129].

muscleblind function is required for alternative splicing regulation as shown by the splicing defects in *α-actinin* and *TroponinT* mRNAs in *muscleblind*

mutants. We demonstrated with cell culture assays that MuscleblindA, B and C are alternative splicing factors when acting on *Drosophila α-actinin* and mouse *TroponinT3* minigenes (Fig. R.17, R.19). All human MBNL proteins (MBNL1-3) have been described to modify alternative splicing of *cardiac TroponinT* (*cTNT* or *TnnT2*) in cell culture [32]. However, *Mbnl1* knockdown mice show defects in *tnnT2* and several other transcripts [28, 30] but no splicing defect was found in *Mbnl2* knockout mouse. In fact, MBNL2 was implicated in the localization of Integrin α3 protein through the localization of its mRNA [36]. No physiological target for MBNL3 has been described. Actually, its molecular function might be different from the other MBNL proteins as it is an anti-myogenic factor [37]. Thus, although MBNL2 and MBNL3 show the ability to modify splicing in cell culture, they could not be acting like splicing factors or they could be acting onto different targets. The same situation could occur in *Drosophila*.

Indeed, we found MblA, B and C to show different activities as splicing factors (Fig. R.17, R.19). Whereas MblC strongly modified both *Drosophila α-actinin* and mouse *TroponinT3* minigenes, MblA marginally regulated *Drosophila α-actinin* minigene splicing in cell culture and MblB did not have any effect on it. The expression pattern of *mblB* transcripts almost discards its implication in *α-actinin* splicing control during embryo development as it is restricted to late larva and early pupa, but it could be active onto *troponinT* pre-mRNA during those stages. *mblA* is expressed during the embryonic development and in the late larva-early pupa stages so it could be implicated in the defects of splicing in both RNAs. *mblC* is broadly expressed during the entire fly life cycle and it is the only transcript detected in adult flies. Then, it could be implicated in both splicing events. Some muscular defects appear when over-expressing CUG repeat

containing RNA in the adult muscle [85] but the molecular basis has not been described and it is not known whether they are caused by *muscleblind* loss of function or not. Then, more studies are needed to clarify to which extent MuscleblindC is acting as splicing factor in adult tissues. Over-expression of Muscleblind isoforms in *Drosophila* S2 cells showed that they have different ability to activate cell-death, being MblB the isoform with strongest effect on cell viability (Fig. R.20). Also the binding activity was different between the isoforms (Fig. R.11-R.14, R.16 and Table R.1). All together these results show that Muscleblind isoforms are not functionally equivalent and they might not only have different targets but also act in a different manner on them. Works in our lab are now trying to elucidate whether the splicing defects found in *muscleblind* mutants are due to lack of function of any particular isoform and which of them are required for the proper splicing of *α-actinin* and *troponinT* mRNAs. Isoform-specific RNA interference transgenic flies are being generated in which each isoform will be specifically silenced. We are also checking if MblA isoform has a function in subcellular localisation of mRNAs and genome-wide strategies will be used to identify the targets of each Muscleblind isoform.

We found MuscleblindC to be the most active isoform as splicing factor when acting on *α-actinin* and *TroponinT* (Fig. R.17, R.18). Evolutionary studies showed that MuscleblindC is the only isoform conserved in protostomes, thus suggesting it has been under higher selective pressure than the other protein isoforms [126]. Indeed, it has been recently demonstrated that MuscleblindA, B and C have different ability to rescue a *muscleblind* mutant background, being MuscleblindC the one showing the highest rescuing potential [45]. Taken together these data suggested that

MuscleblindC was performing most of the *muscleblind* functions required for embryonic development, in agreement with its wide expression in embryogenesis. Although marginal, MblA had some effect on *α-actinin* minigene splicing and both MblA and B modified *Tnnt3* minigene transcript isoform ratio (Fig. R.17, R.18). Sequences similar to MblA and MblB isoforms were only found within the melanogaster group [126]. The lack of conservation and their restricted expression patterns [45] suggest that they are probably carrying out specialized functions in the species in which they have been detected. MblD isoform showed the lowest conservation as similar sequences were only found in *Drosophila simulans*. This, together with the high instability that we have observed in all the cell types and organisms that we have used, and the absence of significant effect in any of the different assays we have performed might suggest that MblD is a recently acquired variant, which has lost the function of the original protein and it has not acquired a new function yet. The appearance of this non-functional isoform could be a new mechanism for *muscleblind* function regulation by alternative splicing.

7. Basis of isoform specific behaviour.

Alternative splicing can generate transcripts encoding proteins with subtle or opposing functional differences. Aberrant relative levels of alternative spliced isoforms are expected to affect cellular functions and an increasing number of human diseases are being related to aberrant isoform ratios. This way, regulation of splicing regulators is a clue in the proper development of the organisms as it is the regulation of transcription factors. This regulation can arise by controlling the levels of splicing factors, their activation (phosphorylation) and their efficiency (subcellular localization)

[160]. Well studied splicing factors as PTB have different splicing isoforms themselves with different splicing activity [161]. Then, regulation of alternative splicing is a mechanism to regulate the activity of splicing factors.

We have shown Muscleblind isoforms have different affinity to bind RNA and that they have different capacity to regulate alternative splicing. Rescue assays where the different isoforms were tested for their ability to rescue the lethality of *muscleblind* mutant embryos showed they have different potential [45]. *in vitro* assays showed that Muscleblind RNA binding activity probably resides in the zinc fingers (Fig. R.16), as it does in the human MBNL proteins. The analyses of the C-t sequences of Muscleblind proteins led to the identification of several putative functional domains, such as the alanine and phenylalanine regions only present in MblB, which can give a different functionality to each isoform [42] but none of them have been experimentally shown to modify protein activity. Muscleblind proteins share the N-terminal region in which the binding activity resides. Thus, the differences in binding affinity might reside in C-terminal sequences. Isoform specific regions might confer particular tertiary structure, diverse protein-protein interaction abilities or trigger specific post-translational modifications that modify RNA affinity. The specific RNA sequence to which the proteins bind might be defined by these C-terminal amino acids. These sequences might be also introducing the other activities that we detected, the splicing regulation and apoptosis activation, and others that we have not identified yet.

We found two interesting data that could contribute to explain the functional diversification of Muscleblind isoforms. When expressed in cell culture,

Muscleblind proteins show distinct sub-cellular localisations (Fig. R.21). MuscleblindC, the isoform with highest alternative splicing activity, shows a coherent nuclear distribution while MuscleblindA, that shows a higher RNA affinity *in vitro* but a lower activity as splicing factor, is mainly cytoplasmic. MuscleblindB shows an intermediate behaviour being both nuclear and cytoplasmic signals strong. Its activity as splicing factor is similar to MuscleblindA although it is more enriched in the nucleus, where the splicing occurs. The difference in RNA affinity could help to explain these results as MblA bound *α-actinin* RNA with more affinity than MblC and MblB did not show interaction with the same fragment in the three hybrid experiments in which MblC showed strong binding activity.

The second data I presented about the basis of functional diversity was the analysis of the relevance of the MuscleblindC putative sumoylation site (SUMO). We identified a putative sumoylation site with the consensus sequence FKRP that is conserved in *C.elegans* (Fig. R.22). Sumoylation consists in the covalent attachment of a small ubiquitin-related peptide, SUMO. This post-traductional modification regulates protein activity in a substrate-specific manner and it usually results in intracellular localization regulation [131]. The molecular mechanism by which sumoylation targets intracellular protein localization is not well understood although it probably generates conformational changes that alter protein-protein interactions. This would explain the different ability of MblC to aggregate when mutating sumoylation site compared with the wild type protein. Often those aggregates appear in response to stress when the protein has the ability to form dimers [162-164]. Thus, the elimination of the FKRP site could be changing the ability of MuscleblindC to for dimmers. The localisation experiments were carried out in human cells where the cellular environment

might be quite different from the physiological situation of MuscleblindC. Interestingly, I also observed an alteration of MuscleblindC function in a *Drosophila* cell environment that is much closer to a physiological situation. The increase of pro-apoptotic activity when eliminating SUMO site could be reflecting an over-activation of MblC on RNA processing. Muscleblind proteins might need to interact with other factors that define its precise functionality as many other RNA binding proteins do [17]. Sumoylation could have the role of modifying these interactions to regulate Muscleblind function. Whether the FKRP sequence can be sumoylated is our next priority. Alternatively, we could be interfering in protein folding or eliminating a protein-protein interaction site, thus modifying MuscleblindC functionality.

It would be interesting to address whether the subcellular localisation is coded into the protein sequence (any signal peptide not identified yet) or it depends on post-translational modifications. In fact, sumoylation of MuscleblindC could be related to its major nuclear localisation. Experiments in *Drosophila* S2 cells transfecting wild type MblC and MblC$^{\Delta SUMO}$ are being addressed in our laboratory to determine if MuscleblindC is sumoylated *in vivo* and to analyse the functional consequences of elimination of sumoylation site in a *Drosophila* cell environment.

8. Is the binding to physiological targets separable from the binding to pathogenic CUG repeat containing RNA?

The cell culture assay showed that MuscleblindA, B and C co-localise with the RNA foci containing expanded CUG repeats (Fig. R.11, R.12). Regarding to MuscleblindD, the signal in the co-localisation assay is blurry both in COS and HEK cells. Despite of that, the signal is different to that

observed when transfecting the GFP vector alone and experiments made afterwards by J.M. Fernández in *Drosophila* S2 cells indicate that the signal detected in MblD over-expression experiments is specific. Furthermore, the only western blot in which full length MblD has been observed was in a similar experiment carried out by Mike Poulos in COS cells. Thus, we suggest MuscleblindD does not co-localise with expanded CUG containing RNA in human cells and it is so unstable that is being degraded during the protein extraction prior to western blot detection. This leaves a situation in which the two zinc fingers seem to be required for Muscleblind proteins to be sequestered, although it is yet to be known whether they are sufficient.

The minigene splicing assay showed that only isoforms with the two zinc fingers can modify the splicing of the transcripts analysed, but they also showed that the C-terminus sequence can modify the protein activity. Thus, Muscleblind activity as splicing factor is not completely dependent on the binding activity of the zinc fingers. Interestingly, the three hybrid results point to the requirement of just one zinc finger to bind *α-actinin* RNA as MblD interacted with the intron fragment of α-actinin in two independent experiments. It was shown by three hybrids that the last zinc finger in MBNL is required to interact with CUG repeats but not with CCUG expansions [43]. Then, it would be interesting to test whether the four MBNL1 zinc finger motifs are required for the binding to its physiological targets or not. If the requirements for the binding to physiological targets and to pathological expansions are different enough, an important perspective of therapy would be open.

9. Implication in new processes: apoptosis.

We found that over-expression of Muscleblind isoforms in *Drosophila* S2 cells considerably reduced cell viability in an isoform-specific manner (Fig. R.20A). The reduction of the cell-death inducing activity when co-expressing the apoptotic inhibitor Diap1 or using chemical caspase inhibitors indicate that they probably activate cell death through the apoptotic pathway (Fig. R.20B). According to this, several key apoptotic genes showed genetic interaction with *mblC* over-expression in *Drosophila* adult eye. Immunostaining of wing imaginal discs over-expressing *mblC* showed increased caspase activity in specific regions inside the area where *mblC* expression was driven [126]. An accurate analysis of the interaction of Muscleblind induced cell death and the apoptotic pathway would clarify if this is the mechanism by which Muscleblind is killing the cells. Supression of the cell death by co-expression of *Diap1* or administration of apoptosis inhibitors would confirm the genetic interactions and the preliminary cell culture experiment presented in this thesis. The work performed in our laboratory is the first study that relates muscleblind function with the apoptotic pathway. The relevance of this insight is yet to be revealed but it is worth to note that myoblast cells transiently expressing expanded *DMPK* alleles showed increased susceptibility to oxidative stress [165]. These cells underwent cell death with characteristics of apoptosis.

Further analyses are necessary to elucidate whether MblC activation of apoptosis is due to a direct regulation of any apoptotic gene at RNA level or not. Apoptotic genes produce pro-apoptotic or anti-apoptotic isoforms depending on the regulation of their splicing [166] and Muscleblind proteins could be affecting the isoform ratio of a key apoptotic gene. However, other

explanations are also possible. Muscleblind proteins could be regulating the isoform ratio of a cell adhesion molecule causing apoptosis by inefficient attachment to substrate. Furthermore, Muscleblind proteins could also perform other functions in RNA metabolism. Human MBNL proteins are implicated not only in splicing but also in RNA localization [32, 36] and other RNA binding proteins in *Drosophila* as Bruno1 have been described to be splicing factors and translation repressors [150, 151]. In fact, elimination of sumoylation site in MuscleblindC increased its pro-apoptotic activity without affecting its activity as splicing factor on *TroponinT3* minigene (Fig. R.22). The possibility of Muscleblind implication in apoptosis being independent of its splicing regulation activity would be in agreement with the fact that MuscleblindB, an isoform that cannot alter *Drosophila α-actinin* splicing and has a lower activity on mouse *TroponinT3* minigene isoform ratio, is the isoform with a higher pro-apoptotic activity. In order to test this possibility in depth, another splicing assay should be used. The *TroponinT3* minigene assay cannot detect hyperactivation of MuscleblindC activity on splicing by lack of sumoylation because the wild type protein already excludes completely the inclusion of the foetal exon into the mature mRNA. In any case, it is also possible that MuscleblindB targets substantially differ from those of MuscleblindC. It is worth noting that *muscleblindB* expression is restricted to late larva and early pupa stages [45] and, at that moment, the metamorphic process is at maximum speed and a large amount of cells is undergoing apoptosis. MuscleblindB could be in the pathway that activates this generalised apoptosis.

Conclusions

No es posible descender dos veces al mismo río, tocar dos veces una sustancia mortal en el mismo estado, ya que a causa del ímpetu y la velocidad de los cambios, se dispersa, vuelve a reunirse, y aflora y desaparece.

-Heráclito de Éfeso-

Taken together our results support the following main conclusions:

I-*Drosophila* Muscleblind proteins are alternative splicing regulators. We found *α-actinin* and *TroponinT* transcripts to be among their physiological targets.

II- Muscleblind proteins likely recognize the same binding sequences as human MBNL1 since MuscleblindA binds an *α-actinin* pre-mRNA fragment that contains a cluster of seven overlapping MBNL1 consensus binding sequences through the zinc finger-containing N-terminal region.

III-Over-expression of Muscleblind proteins activates cell death, probably through the apoptotic pathway. MuscleblindB reduced the most cell viability. MuscleblindA and C also increased cell death whereas MblD seemed to enhance cell viability.

IV- A conserved putative sumoylation site (FKRP) is implicated in the cell-death induction by MuscleblindC and its ability to aggregate. Site-directed mutagenesis of the FKRP site enhanced the ability of MuscleblindC to activate cell death upon over-expression, with no discernible effect on splicing.

V- Muscleblind isoforms are functionally distinct and show distinct preferences in their sub-cellular localisation in human cells. In particular, MuscleblindA, B, C and D show different RNA binding affinities in a yeast three-hybrid assay and differ in their ability to modify the alternative splicing of *Drosophila α-actinin* and murine *TroponinT3* minigenes.

VI- Non-coding CUG repeat RNAs are pathogenic to *Drosophila* cells. MuscleblindA, B, and C co-localise with expanded CUG repeat RNA foci in human cells and expression of 480 CUG repeat containing RNA in *Drosophila* misregulates the alternative splicing of *α-actinin* pre-mRNA.

Resumen en castellano

Lo difícil se consigue, lo imposible, se intenta.
-Napoleón Bonaparte-

INTRODUCCIÓN Y OBJETIVOS

El gen *muscleblind* (*mbl*) fue descrito inicialmente en *Drosophila* como un gen cuya función era necesaria para el adecuado desarrollo del sistema nervioso periférico embrionario [46, 109]. La caracterización detallada de los mutantes *muscleblind* mostró su implicación en la diferenciación terminal de los fotorreceptores [44] y los músculos [41]. La generación de clones mitóticos mutantes para *muscleblind* en las células precursoras de los fotorreceptores no originó fenotipo externo, pero cortes tangenciales mostraron que los rabdómeros de los fotorreceptores presentaban un tamaño reducido y no se extendían correctamente hasta las regiones basales de la retina [44]. Mutaciones nulas para *muscleblind* son letales en homocigosis. Las moscas mueren en embriogénesis tardía, como larvas que no pueden salir del corion [41]. Los músculos de estos embriones mutantes aparecen hipercontraidos, con ausencia de bandas I y discos Z. Además, la matriz tendinosa extracelular de las uniones indirectas del músculo a la epidermis está fuertemente reducida.

De acuerdo con este fenotipo mutante, el análisis de la expresión de Muscleblind reveló un patrón principalmente mesodérmico, con detección de señal en la musculatura esquelética y visceral [41]. En cuanto al ectodermo, Muscleblind se detectó en el sistema nervioso central y en los órganos de Bolwig, que son los precursores de los fotorreceptores larvarios. Hay pocos datos acerca de la regulación de la expresión de *muscleblind*. La expresión de *muscleblind* en el músculo desaparece en mutantes *Dmef2* [41] y recientemente se detectó la unión del factor de transcripción pro-miogénico Dmef2 a secuencias reguladoras de *mbl* [97].

A partir del único gen *muscleblind* de *Drosophila* se generan cuatro transcritos por procesado alternativo que codifican para cuatro proteínas caracterizadas por la presencia de dedos de zinc del tipo CCCH. Las isoformas MuscleblindA, B y C poseen dos dedos de zinc de este tipo mientras MuscleblindD sólo tiene uno [44]. Las tres primeras comparten 179 aminoácidos entre ellas y 63 aminoácidos con MblD. Todas ellas poseen regiones carboxilo terminales específicas susceptibles de conferirles especializaciones funcionales, aunque no se han realizado estudios al respecto.

Los homólogos de *muscleblind* en vertebrados son los genes *Muscleblind-like1*, *2* y *3* (*MBNL1-3*). Los ratones de falta de función de *Mbnl1* sufren miotonía (incapacidad para relajar los músculos), tienen defectos histológicos en el músculo, cataratas y, a nivel molecular, defectos en la regulación del procesado alternativo de varios transcritos [28, 30]. Las proteínas humanas MBNL1, 2 y 3 tienen la capacidad de modificar el procesado alternativo de los transcritos de la *troponinaT cardiaca* (*cTNT*) y el *receptor de insulina* (*IR*) en cultivo celular [32]. MBNL1 une el transcrito de la *cTNT* en una secuencia YGCU(U/G)Y (Y es pirimidina) intrónica y reprime la inclusión del exon E5 en el transcrito maduro. Del balance entre los niveles de MBNL1, PTB, CUG-BP1 y ETR-3 resulta la inclusión o exclusión final del E5 en el transcrito maduro [31]. En el caso del transcrito del *IR*, las proteínas CUG-BP1 y hnRNP H forman un complejo represor de la inclusión del exon E11, mientras que MBNL1 reprime la acción de este complejo al unirse a hnRNP H [33].

Se han detectado niveles alterados de expresión de los genes *MBNL* en diversas patologías [36, 54, 55]. Las proteínas MBNL tienen un papel

importante en la patogénesis de la distrofia miotónica (DM). La DM es una enfermedad autonómica dominante generada por la expansión del trinucleótido CTG en una región no codificante del gen *DMPK*. Esta secuencia repetitiva expandida está presente en el RNA de *DMPK* y genera una estructura secundaria en horquilla que provoca el secuestro de factores nucleares, fundamentalmente las proteínas MBNL, impidiendo su funcionamiento normal [51, 70, 71, 167]. Dos observaciones fundamentales apoyan la hipótesis de que gran parte de los síntomas de la enfermedad se deben a la falta de función de las proteínas MBNL. Primero, el ratón de falta de función de *Mbnl1* reproduce muchos de los síntomas de la enfermedad mencionados anteriormente [28]. En segundo lugar, la expresión de la proteína MBNL1 en ratones modelo de DM que expresan repeticiones largas como las que originan la enfermedad, restaura el fenotipo normal [29].

Desde 1997 nuestro laboratorio ha mantenido un interés continuo en la caracterización de la función molecular de las proteínas Muscleblind de *Drosophila* y, más recientemente, en su relevancia en el mecanismo de patogénesis de la distrofia miotónica. En este contexto general fue donde iniciamos el presente proyecto de tesis doctoral. Hipotetizamos que *Drosophila* podía servir como modelo biomédico y, en particular, que las proteínas Muscleblind humanas y de *Drosophila* desarrollaban funciones semejantes *in vivo*. Más allá de esto, supusimos que la unión de las proteínas Muscleblind a sus dianas fisiológicas y a RNAs que contienen repeticiones CUG, eran actividades separables. De este modo, el objetivo general de este proyecto de tesis fue la demostración de la conservación funcional entre las proteínas Muscleblind de *Drosophila* y humanas, así como encontrar mutaciones en Muscleblind que inhibieran la unión a

repeticiones CUG sin comprometer su función fisiológica. Este objetivo general requirió el desarrollo de diversos reactivos y técnicas, y el abordaje de los siguientes objetivos específicos:

1. Análisis de la conservación de la ruta de patogénesis de la distrofia miotónica en *Drosophila*.

2. Estudio de la conservación en las proteínas Muscleblind de *Drosophila* de la función como regulador del procesado alternativo descrita para las proteínas MBNL humanas.

3. Evaluación de la diversificación funcional de las isoformas de Muscleblind como variantes proteicas naturales.

4. Análisis de la unión de Muscleblind a dianas de RNA fisiológicas y a RNA con repeticiones CUG.

5. Caracterización de dominios funcionales en las proteínas Muscleblind *de Drosophila*.

RESULTADOS Y CONCLUSIONES

1. *muscleblind* es necesario para el correcto procesado alternativo de los transcritos de la *α-actinina* y la *TroponinaT* de *Drosophila*.

Los ratones de falta de función de *Mbnl1* presentan una alteración de la regulación del procesado alternativo que resulta en un mantenimiento en el adulto de un conjunto de transcritos típicos de la etapa fetal. [29, 39]. Por ello decidimos analizar los mutantes *muscleblind* de *Drosophila* en busca de defectos en el procesado alternativo.

Los embriones mutantes *muscleblind* muestran ausencia de bandas Z [41]. La α-Actinina (α-Actn) es la proteína más abundante en esta estructura conectiva. Moscas mutantes para la *α-actinina* presentan defectos en las uniones del músculo al epitelio y ausencia de bandas Z [120], de modo similar a los mutantes *muscleblind*. Además, en vertebrados, la *α-actinina* presenta isoformas específicas de tejido como consecuencia del procesado alternativo de los transcritos inmaduros y proteínas relacionadas con la función de MBNL1 como regulador del splicing alternativo, como PTB, CUG-BP, ETR-3 y CELF4, están implicadas en la regulación de este suceso de procesado alternativo [119]. Los transcritos del único gen de la *α-actinina* de *Drosophila* generan por procesado alternativo una isoforma específica de tejido no muscular (*α-actinina A*) y dos de tejido muscular (*α-actinina B* en músculo adulto y *α-actinina C* en músculo larvario) (Fig. C.1A). Todo ello nos llevó a considerar que los transcritos de la *α-actinina* podrían ser una diana directa de Muscleblind en *Drosophila*.

Por otro lado, el RNA mensajero de la *TroponinaT* es una diana conocida de la proteína MBNL1 humana. En concreto, dos genes parálogos de la TroponinaT, la *TroponinaT cardiaca* (*tnnT2*) y la *TroponinaT de fibras esqueléticas rápidas* (*tnnT3*), presentan defectos de procesado en ratones de falta de función de *Mbnl1* [28]. Además, se ha propuesto un modelo de equilibrio dinámico entre MBNL1, PTB, CUGBP y ETR-3 para la regulación del procesado alternativo del mRNA de *tnnT2* en cardiomiocitos de ratón y pollo [31]. El único gen de la *TroponinaT* en *Drosophila* genera cuatro isoformas por procesado alternativo de los exones E3, E4 y E5 [121]. Estos transcritos son específicos de tipo muscular y su expresión está regulada a lo largo del desarrollo (Fig. C.1B).

Para testar la hipótesis de que el gen *muscleblind* de *Drosophila* regula el procesado alternativo de los genes *α-actinina* y *TroponinaT*, analizamos por RT-PCR los transcritos presentes en moscas mutantes y control en estadío embrionario, larva, pupa y adulto. Los embriones mutantes nulos para *muscleblind* mostraron defectos en el procesado del RNA mensajero de la *α-actinina* (Fig. C.1C), al presentar reducido el transcrito de la *α-actininaC* e incrementados los de la *α-actininaA* y *B*. Por su parte, pupas trans-heterocigotas mbl^{E27}/mbl^{K7103}, combinación alélica que reduce la función *muscleblind* pero permite el desarrollo hasta adulto de algunas moscas, presentan defectos en el procesado alternativo del mensajero de la *TroponinaT* (Fig. C.1D), viéndose un incremento en el transcrito típico de los músculos indirectos del vuelo y del músculo depresor tergal del trocánter. El defecto en el procesado del mensajero de la TroponinaT es ya detectable en heterocigotos portadores del alelo nulo mbl^{E27}, pero no en los heterocigotos para el alelo que provoca una pérdida de función más leve mbl^{K7103}, indicando un posible efecto de la dosis de *muscleblind*.

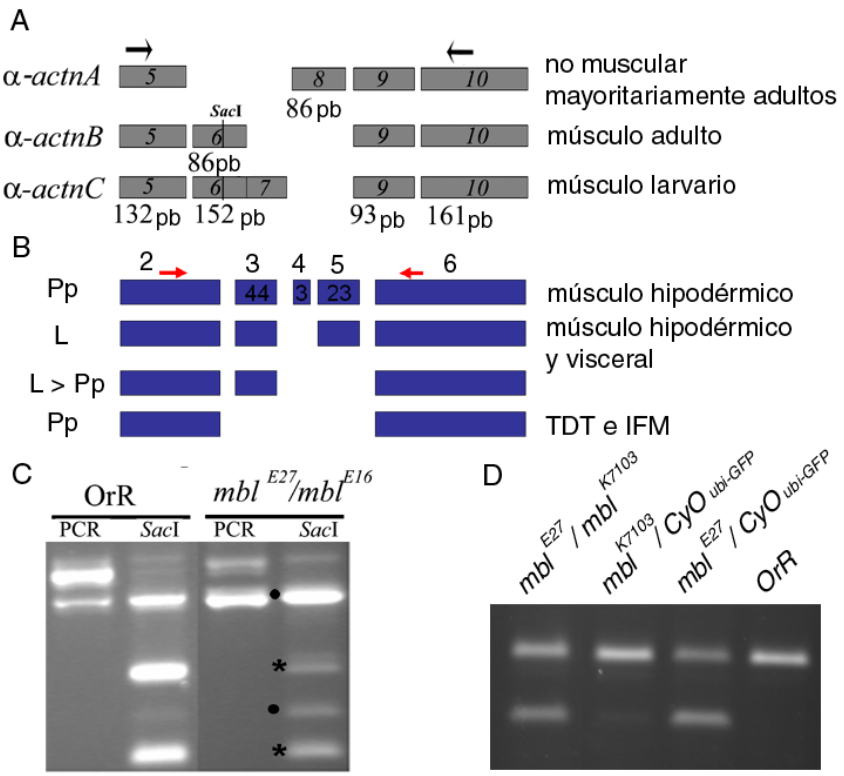


Figura C.1. Moscas mutantes *muscleblind* presentan defectos de procesado alternativo en los transcritos de la *α-actinina* y la *TroponinaT*. A) Esquema de los transcritos de la *α-actinina* de *Drosophila* mostrando la nomenclatura utilizada. La regulación durante el desarrollo se muestra a la derecha. El tamaño de los exones en pares de bases (pb) se indica debajo. Las isoformas A y B tienen el mismo tamaño y se diferencian mediante digestión con *Sac*I. Los cebadores utilizados en la RT-PCR se representan con flechas. **B)** Esquema de los transcritos generados a partir del RNA mensajero de la *TroponinaT* de *Drosophila* mostrando la nomenclatura utilizada. El tamaño de los exones en pb se representa en su interior. La regulación específica de tejido se muestra a la derecha (TDT = músculo depresor tergal del trocánter; IFM = músculo indirecto del vuelo). La regulación durante el desarrollo se muestra a la izquierda (Pp = pupa; L = larva). Las flechas representan los cebadores empleados en la RT-PCR. **C)** Electroforesis en agarosa al 2% de los productos de la PCR de la *α-actinina* en embriones de 16 a 18 h, y su correspondiente digestión con *Sac*I. Los embriones mbl^{E27}/mbl^{E16} presentan reducida la isoforma *α-actininaC* (*) e incrementada la *α-actininaA* y *B* (puntos negros). **D)** Separación

electroforética en agarosa al 2% de los productos de la RT-PCR sobre los transcritos de la *TroponinaT* en pupas tempranas (menos de 48 h). Las pupas mbl^{E27}/mbl^{K7103} presentan incrementada la isoforma típica de TDT e IFM. Este defecto ya se aprecia en heterocigotos portadores del alelo nulo mbl^{E27}.

En conjunto, los resultados muestran que la función de *muscleblind* es necesaria para la correcta regulación de eventos específicos de corte y empalme en momentos determinados del desarrollo.

2. **El efecto tóxico de la expresión de un RNA no codificante portador de repeticiones CUG expandidas está conservado en *Drosophila*.**

Las proteínas MBNL humanas son secuestradas por los transcritos portadores de repeticiones CUG expandidas que generan la distrofia miotónica [51]. Este secuestro se comprueba por la co-localización con el RNA portador de estas expansiones en células de pacientes y se hipotetiza que genera la falta de función de la proteína, desencadenando gran parte de los síntomas de la enfermedad. Los pacientes de DM muestran numerosos síntomas que son reproducidos por la falta de *Mbnl1* en ratones, entre ellos, la miotonía, las cataratas y la alteración del procesado alternativo de muchos transcritos, que muestran retención de exones fetales en tejido adulto [28, 30].

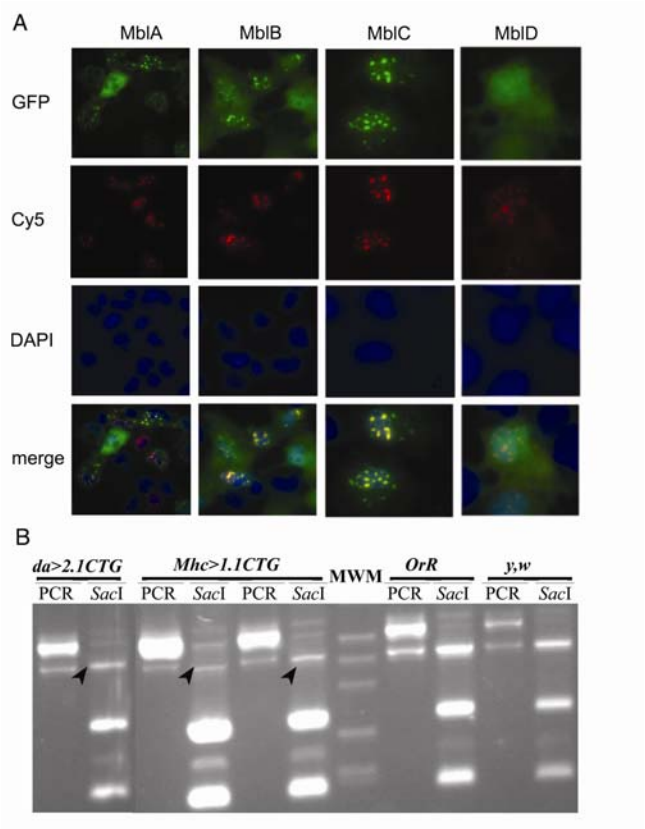


Figura C.2. El mecanismo de toxicidad de los transcritos portadores de repeticiones CUG expandidas está conservado en *Drosophila*. A) Las proteínas Muscleblind se expresaron en células COS junto con un transcrito no codificante portador de las repeticiones CUG expandidas. Se utilizó 1 μg de cada plásmido. Se muestra la detección de proteínas fusionadas a GFP (GFP; verde); sonda CAG hibridando con el RNA portador de las repeticiones CUG (Cy5; rojo); y tinción DAPI de los núcleos (DAPI; azul); así como la superposición de las imágenes (merge). Todas las proteínas Muscleblind excepto MblD co-localizan con las inclusiones ribonucleares. **B)** Separación electroforética (agarosa 2%) de los productos de RT-PCR de la *α-actinina* y sus correspondientes digestiones con *Sac*I de embriones: control (*OrR* y *y,w*), expresando repeticiones CUG en un patrón general (*da*) o muscular (*Mhc*). 2.1 y 1.1 indican la cepa transgénica utilizada. MWM = marcador de peso molecular; bandas: 517, 453, 394, 298, 234 y 154 pb. Las cabezas de flecha señalan la reducción del transcrito *α-actininaA* en moscas que expresan RNAs portadores de repeticiones CUG.

Para analizar si la ruta de patogénesis de los transcritos portadores de repeticiones CUG se conserva en *Drosophila*, nos propusimos inicialmente comprobar que las proteínas Muscleblind de *Drosophila* co-localizan con dichos transcritos en células de vertebrado. Posteriormente, estudiamos si la expresión de RNAs portadores de 480 repeticiones CUG en moscas generaba alteraciones en el procesado de RNAs, del mismo modo que lo hace en pacientes.

La co-expresión de las isoformas de Muscleblind fusionadas con GFP y RNA con expansiones largas de repeticiones CUG en células COS mostró que MblA, MblB y MblC co-localizan con los foci ribonucleares generados por los RNAs portadores de repeticiones CUG (Fig. C.2A).

El análisis del procesado alternativo de la *α-actinina* en moscas que expresan RNAs portadores de 480 repeticiones CUG mostró defectos en la regulación del procesado, aunque fueron diferentes a los hallados anteriormente en mutantes *muscleblind* (Fig. C.2B). En moscas que expresan RNAs portadores de repeticiones CUG, el transcrito de la *α-actininaA* aparece fuertemente reducido.

De estos resultados se concluye que las proteínas Muscleblind de *Drosophila* son secuestradas por los RNAs portadores de repeticiones de un modo similar a como lo hacen las MBNL humanas. Así mismo, la expresión de RNAs portadores de repeticiones en la mosca genera alteraciones del procesado alternativo en dianas fisiológicas de *mbl*.

3. Las proteínas Muscleblind modifican el procesado alternativo de minigenes en cultivo celular y unen por su región N-terminal un fragmento de RNA con secuencias consenso para la unión de MBNL1.

Las proteínas MBNL humanas se unen directamente al RNA a través de una secuencia consenso YGCUU/GY (Y = pirimidina) para modificar su procesado alternativo [32]. Ensayos en levadura mostraron que para su interacción con RNAs portadores de repeticiones, los dedos de zinc de MBNL1 y los aminoácidos que los separan son necesarios [43].

Con el fin de establecer si la unión al RNA de las proteínas Muscleblind de *Drosophila* es directa, realizamos ensayos *in vitro* incubando las proteínas con fragmentos del RNA inmaduro de la *α-actinina* de *Drosophila*. El ensayo de entrecruzamiento con luz ultravioleta mostró que fragmentos N-terminales de MblA, conteniendo los dedos de zinc, interaccionan con un fragmento de la *α-actinina* de *Drosophila* que contiene siete secuencias consenso para la unión de MBNL1 solapadas (Fig. C3A-A"). Además, para demostrar que, al igual que las proteínas humanas, las proteínas Muscleblind de *Drosophila* son factores reguladores del procesado alternativo, realizamos ensayos en cultivo celular de la actividad de estas proteínas sobre el procesado de minigenes de la *α-actinina* de *Drosophila* y de la *troponinT3* de ratón (Fig. C3B,C). MblC mostró gran actividad sobre ambos minigenes. MblA modificó levemente el procesado de *α-actinina*, y MblA y MblB alteraron el patrón de procesado de *tnnT3*, siendo ligeramente menos activos que MblC.

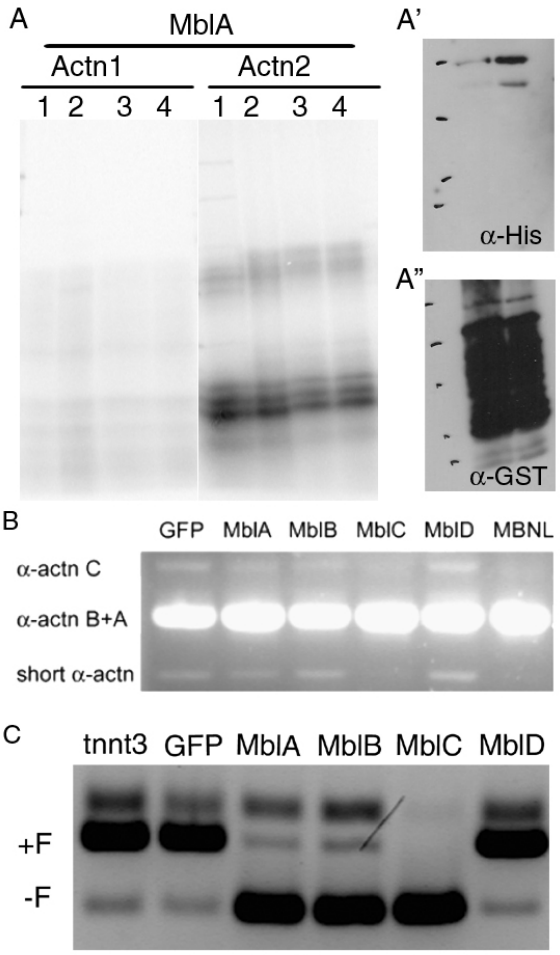


Figura C.3. Las proteínas Muscleblind son factores de procesado alternativo que unen directamente el RNA. A) Ensayo de entrecruzamiento por luz ultravioleta en el que se incubó MblA con un fragmento intrónico de la *α-actinina* que contiene secuencias consenso para la unión de MBNL (Actn2) y un fragmento contiguo que no contiene ninguna de estas secuencias (Actn1). El ensayo se realizó incubando el RNA y la proteína en presencia de 1: H_2O; 2: 0.2 μg/μl rRNA; 3: 5.5 μg/μl heparina; 4: 0.2 μg/μl rRNA y 5.5 μg/μl heparina **A')** El anticuerpo α-His detecta la presencia de proteína completa (His_6 se sitúa en el extremo C-t de MblA) **A")** El anticuerpo α-GST detecta la presencia de numerosos fragmentos de degradación conteniendo el fragmento N-t con los dedos de zinc (GST se sitúa en el extremo N-t de MblA) **B-C)** Electroforesis (agarosa 2%) de los productos de RT-PCR de extractos de células humanas en las que se co-expresaron las proteínas Mbl y MBNL1 fusionadas a GFP junto con el minigen de la *α-actinina* de *Drosophila* **(B)** y la tnnT3 de ratón **(C). B)** La expresión de MblC y MBNL1 elimina la formación de transcritos *α-actnC* y *short α-actn*. MblA disminuye ligeramente *short α-actn*. **C)** +/-F indica la presencia/ausencia del exón fetal. La expresión de MblA, B y C disminuye la inclusión del exón fetal de *tnnT3*.

En resumen, las proteínas Muscleblind son factores de procesado alternativo que regulan el corte y empalme de los transcritos de la *α-actinina* y la *TroponinaT*. Además, MblA une directamente un fragmento del mensajero de la *α-actinina* que contiene siete secuencias consenso de unión de MBNL1. Esta unión se produce a través de una región N-terminal que contiene los dedos de zinc.

4. La sobreexpresión de las proteínas Muscleblind activa la muerte celular en células de *Drosophila*.

Trabajos realizados en nuestro laboratorio mostraron que la reducción de la dosis de genes reguladores de la ruta apoptótica, como *Diap1* y *reaper,* interaccionan con un fenotipo de sobre-expresión de *mblC* en el ojo de *Drosophila*. La sobre-expresión de proteínas pro-apoptóticas en células S2 provoca un incremento la muerte celular que puede ser monitorizado [127]. Para comprobar si las proteínas Muscleblind de *Drosophila* tenían la capacidad de activar la muerte celular, sobre-expresamos las distintas isoformas marcadas con el epítopo myc. Todas las isoformas salvo MblD incrementaron la muerte celular, siendo MblB la isoforma más activa en este ensayo (Fig. C.4).

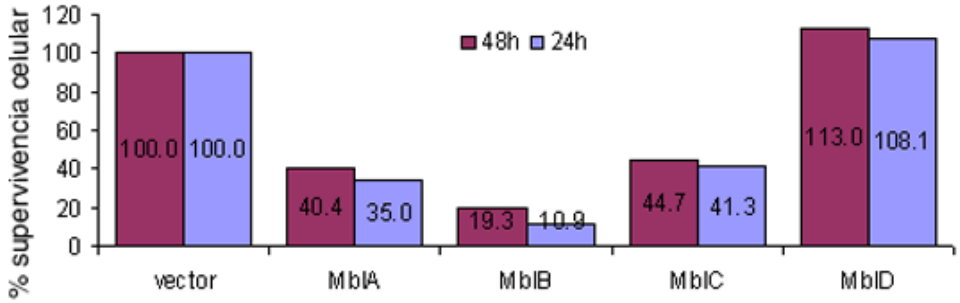


Figura C.4. Las proteínas Muscleblind activan la muerte celular en células S2. Se representa el porcentaje medio (dos réplicas) de células vivas respecto a las contabilizadas en las muestras control. La sobre-expresión de MblA, MblB y MblC fusionadas a myc redujo la viabilidad en recuentos realizados 48 y 24 h después de la transfección.

Estos resultados muestran la capacidad de las proteínas Muscleblind para la activar la muerte celular. Las interacciones genéticas observadas en nuestro laboratorio apuntan a que esta muerte celular es debida a apoptosis.

5. La localización subcelular de las distintas isoformas Muscleblind y un sitio potencial de sumolización presente únicamente en MuscleblindC contribuyen a su diversificación funcional.

Hemos visto que las distintas proteínas Muscleblind tienen distinta actividad en los diferentes ensayos que hemos realizado. Aunque la razón última para esta diversidad funcional debe residir en la secuencia, es posible que la localización subcelular esté influyendo en la actividad de la proteína. Con el fin de comprobar si las isoformas Muscleblind se localizan de manera preferente en los distintos compartimentos celulares, transfectamos células de mamífero con las proteínas fusionadas a GFP. Al transfectar con gran cantidad de plásmido, observamos que MblA, B y C formaban agregados proteicos y se distribuían de manera diferente en la célula (Fig. C5A). MblA apareció fundamentalmente citoplasmática, MblC enriquecida en el núcleo, y MblB con un comportamiento intermedio, presente en núcleo y citoplasma. En cuanto a MblD, se observó una señal fundamentalmente nuclear bastante difusa, posiblemente relacionada con la inestabilidad de la proteína.

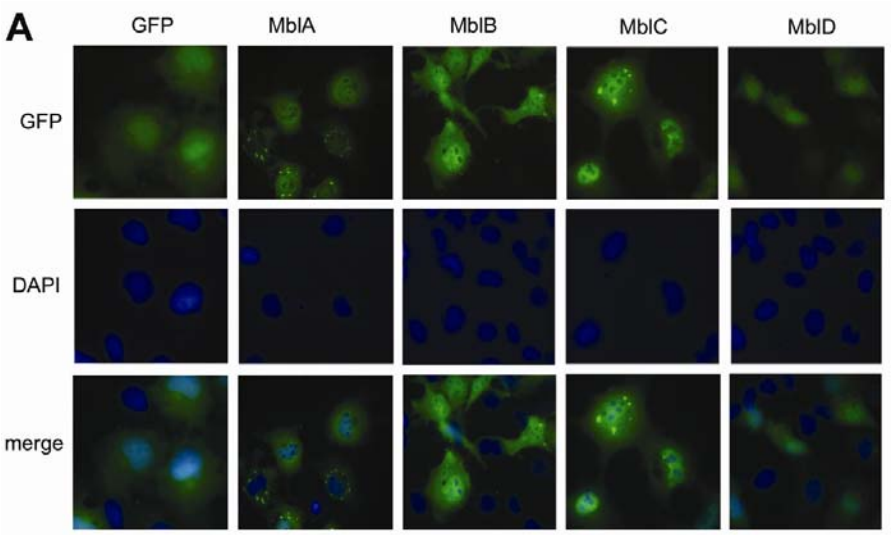


Figura C.5. Las proteínas Muscleblind presentan distinta localización subcelular. Las proteínas Muscleblind fusionadas a GFP se transfectaron en células COS. Se muestra la detección de las proteínas Mbl-GFP (GFP; verde) y la tinción DAPI de los núcleos (DAPI; azul), así como la superposición de las imágenes (merge).

Mientras que la isoforma MblD difiere sustancialmente de las demás al tener un solo dedo de zinc, MblA, MblB y MblC comparten una gran parte de su secuencia (179 aa). Las diferencias funcionales entre ellas, por tanto, deben residir en las secuencias específicas C-terminales. El análisis *in silico* de estas secuencias detectó la existencia de regiones de baja complejidad ricas en alanina o fenilalanina en MblB [42]. Posteriormente, trabajos en nuestro laboratorio llevaron a la detección de un sitio potencial de sumolización que aparecía conservado en *C.elegans* (Figura C.6A). La alteración de dicha secuencia por mutagénesis dirigida (cambio K202I representado en Fig. C.6A) no modificó la actividad reguladora del procesado alternativo del minigen de la *tnnT3* de ratón ni la co-localización con los transcritos portadores de repeticiones CUGs, pero incrementó la frecuencia de agregados proteicos en células COS (Fig. C6B) y aumentó la

capacidad de la proteína para inducir muerte celular en cultivo de células S2 de *Drosophila* (Fig. C6C).

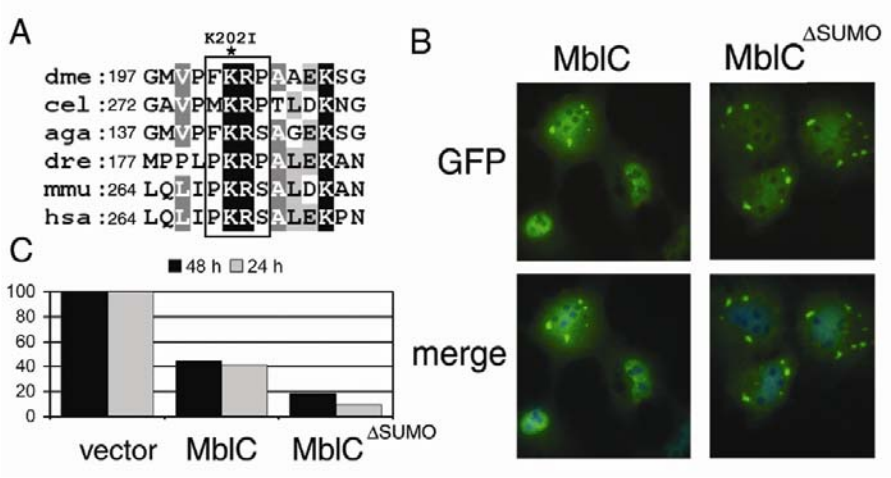


Figura C.6. La mutación de un sitio potencial de sumolización incrementa la formación de agregados y la activación de la muerte celular por la proteína MblC. A) El alineamiento muestra la conservación de la secuencia FKRP, un sitio potencial de sumolización. Se indica el cambio de aminoácido (K202I). dme: *Drosophila melanogaster*; cel: *Caenorhabditis elegans*; aga: *Anopheles gambiae*; dre: *Dario rerio*; mmu: *Mus musculus*; hsa: *Homo sapiens*. **B)** Transfección de células COSM6 con la proteína MblC y la forma mutada MblC$^{\Delta SUMO}$. Se muestra la detección de proteínas Mbl-GFP (GFP; verde), y la superposición con la tinción DAPI (azul) de núcleos (merge). **C)** Se representa el porcentaje medio de supervivencia celular al expresar las proteínas MblC silvestre y mutada respecto a la encontrada al transfectar el vector vacío.

Con estos resultados concluimos que la diferente localización subcelular de las isoformas proteicas de Muscleblind podría contribuir a la especialización funcional que detectamos en distintos ensayos. Asimismo, identificamos un sitio potencial de sumolización como motivo funcional de la proteína MblC, puesto que un fenotipo de sobre-expresión se ve alterado al mutar esta secuencia.

Tomados en conjunto, nuestros resultados apoyan las siguientes conclusiones:

I-Las proteínas Muscleblind de *Drosophila* son reguladores del procesado alternativo. Entre sus dianas fisiológicas encontramos los transcritos de la *α-actinina* y la *TroponinaT*.

II-Las proteínas Muscleblind probablemente reconocen la misma secuencia de unión que la proteína humana MBNL1 ya que MuscleblindA une un fragmento del pre-mRNA de la *α-actinina* que contiene siete secuencias de unión para MBNL1 solapadas a través de una región N-terminal que contiene los dedos de zinc.

III-La sobre-expresión de las proteínas Muscleblind activa la muerte celular, probablemente a través de la ruta apoptótica. MuscleblindB fue la isoforma que más redujo la viabilidad celular. MuscleblindA y MuscleblindC también incrementaron la muerte celular mientras MuscleblindD pareció incrementar la viabilidad celular.

IV-Un sitio conservado, posiblemente de sumolización, (FKRP) está implicado en la regulación de la inducción de la muerte celular por MuscleblindC así como de su capacidad para agregarse. La mutación del sitio FKRP incrementó la capacidad de MuscleblinC para activar muerte celular sin tener un efecto detectable en su actividad como factor de procesado alternativo.

V-Las isoformas Muscleblind son funcionalmente distintas y muestran localizaciones sub-celulares preferentes en células de mamífero. En

concreto, MuscleblindA, B, C y D tienen diferente afinidad por el RNA en un ensayo de tres híbridos en levaduras y distinta capacidad para modificar el procesado alternativo de minigenes de *α-actinina* de *Drosophila* y *TroponinaT3* de ratón.

VI-RNAs no codificantes portadores de repeticiones CUG son patogénicos para células de *Drosophila*. MuscleblindA, B, y C co-localizan con foci ribonucleares formados por RNA portadores de repeticiones CUG en células humanas. La expresión de RNAs portadores de 480 repeticiones CUG en *Drosophila* altera el procesado alternativo del transcrito de la *α-actinina*.

Bibliografía

Una palabra mal colocada estropea el
más bello pensamiento
-autor desconocido-

1. Schmucker, D., et al., *Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity.* Cell, 2000. **101**(6): p. 671-84.
2. Modrek, B. and C. Lee, *A genomic view of alternative splicing.* Nat Genet, 2002. **30**(1): p. 13-9.
3. Matlin, A.J., F. Clark, and C.W. Smith, *Understanding alternative splicing: towards a cellular code.* Nat Rev Mol Cell Biol, 2005. **6**(5): p. 386-98.
4. Faustino, N.A. and T.A. Cooper, *Pre-mRNA splicing and human disease.* Genes Dev, 2003. **17**(4): p. 419-37.
5. Le Hir, H., et al., *The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay.* Embo J, 2001. **20**(17): p. 4987-97.
6. Nilsen, T.W., *The spliceosome: the most complex macromolecular machine in the cell?* Bioessays, 2003. **25**(12): p. 1147-9.
7. Proudfoot, N.J., A. Furger, and M.J. Dye, *Integrating mRNA processing with transcription.* Cell, 2002. **108**(4): p. 501-12.
8. Bernstein, S.I., P.T. O'Donnell, and R.M. Cripps, *Molecular genetic analysis of muscle development, structure, and function in Drosophila.* Int Rev Cytol, 1993. **143**: p. 63-152.
9. Black, D.L. and P.J. Grabowski, *Alternative pre-mRNA splicing and neuronal function.* Prog Mol Subcell Biol, 2003. **31**: p. 187-216.
10. Grabowski, P.J. and D.L. Black, *Alternative RNA splicing in the nervous system.* Prog Neurobiol, 2001. **65**(3): p. 289-308.
11. Schiaffino, S. and C. Reggiani, *Molecular diversity of myofibrillar proteins: gene regulation and functional significance.* Physiol Rev, 1996. **76**(2): p. 371-423.
12. Stamm, S., et al., *Function of alternative splicing.* Gene, 2005. **344**: p. 1-20.
13. Arner, A., M. Lofgren, and I. Morano, *Smooth, slow and smart muscle motors.* J Muscle Res Cell Motil, 2003. **24**(2-3): p. 165-73.
14. Daoud, R., et al., *Activity-dependent regulation of alternative splicing patterns in the rat brain.* Eur J Neurosci, 1999. **11**(3): p. 788-802.
15. Pelisch, F., et al., *Cross-talk between signaling pathways regulates alternative splicing: a novel role for JNK.* J Biol Chem, 2005. **280**(27): p. 25461-9.
16. Shin, C. and J.L. Manley, *Cell signalling and the control of pre-mRNA splicing.* Nat Rev Mol Cell Biol, 2004. **5**(9): p. 727-38.

17. Black, D.L., *Mechanisms of alternative pre-messenger RNA splicing.* Annu Rev Biochem, 2003. **72**: p. 291-336.
18. Lopez, A.J., *Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation.* Annu Rev Genet, 1998. **32**: p. 279-305.
19. Mendes Soares, L.M. and J. Valcarcel, *The expanding transcriptome: the genome as the 'Book of Sand'.* Embo J, 2006. **25**(5): p. 923-31.
20. Garcia-Blanco, M.A., A.P. Baraniak, and E.L. Lasda, *Alternative splicing in disease and therapy.* Nat Biotechnol, 2004. **22**(5): p. 535-46.
21. Nissim-Rafinia, M. and B. Kerem, *The splicing machinery is a genetic modifier of disease severity.* Trends Genet, 2005. **21**(9): p. 480-3.
22. Baralle, D. and M. Baralle, *Splicing in action: assessing disease causing sequence changes.* J Med Genet, 2005. **42**(10): p. 737-48.
23. Laity, J.H., B.M. Lee, and P.E. Wright, *Zinc finger proteins: new insights into structural and functional diversity.* Curr Opin Struct Biol, 2001. **11**(1): p. 39-46.
24. Tenenhaus, C., et al., *PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of Caenorhabditis elegans.* Genes Dev, 2001. **15**(8): p. 1031-40.
25. Blackshear, P.J., *Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover.* Biochem Soc Trans, 2002. **30**(Pt 6): p. 945-52.
26. Lai, W.S., et al., *Interactions of CCCH zinc finger proteins with mRNA. Binding of tristetraprolin-related zinc finger proteins to Au-rich elements and destabilization of mRNA.* J Biol Chem, 2000. **275**(23): p. 17827-37.
27. Hudson, B.P., et al., *Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d.* Nat Struct Mol Biol, 2004. **11**(3): p. 257-64.
28. Kanadia, R.N., et al., *A muscleblind knockout model for myotonic dystrophy.* Science, 2003. **302**(5652): p. 1978-80.
29. 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, 2006.

30. Lin, X., et al., *Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy.* Hum Mol Genet, 2006. **15**(13): p. 2087-97.
31. Ladd, A.N., et al., *Dynamic balance between activation and repression regulates pre-mRNA alternative splicing during heart development.* Dev Dyn, 2005. **233**(3): p. 783-93.
32. Ho, T.H., et al., *Muscleblind proteins regulate alternative splicing.* Embo J, 2004. **23**(15): p. 3103-12.
33. Paul, S., et al., *Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing.* Embo J, 2006. **25**(18): p. 4271-83.
34. Charlet, B.N., et al., *Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing.* Mol Cell, 2002. **9**(3): p. 649-58.
35. Philips, A.V., L.T. Timchenko, and T.A. Cooper, *Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy.* Science, 1998. **280**(5364): p. 737-41.
36. Adereth, Y., et al., *RNA-dependent integrin alpha(3) protein localization regulated by the Muscleblind-like protein MLP1.* Nat Cell Biol, 2005. **7**(12): p. 1140-7.
37. Squillace, R.M., D.M. Chenault, and E.H. Wang, *Inhibition of muscle differentiation by the novel muscleblind-related protein CHCR.* Dev Biol, 2002. **250**(1): p. 218-30.
38. 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, 2002. **11**(7): p. 805-14.
39. Kanadia, R.N., et al., *Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3.* Gene Expr Patterns, 2003. **3**(4): p. 459-62.
40. Fernandes, J.M., J.R. Kinghorn, and I.A. Johnston, *Characterization of two paralogous muscleblind-like genes from the tiger pufferfish (Takifugu rubripes).* Comp Biochem Physiol B Biochem Mol Biol, 2007. **146**(2): p. 180-6.
41. Artero, R., et al., *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**(2): p. 131-43.

42. Pascual, M., et al., *The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing.* Differentiation, 2006. **74**(2-3): p. 65-80.
43. Kino, Y., et al., *Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats.* Hum Mol Genet, 2004. **13**(5): p. 495-507.
44. Begemann, G., et al., *muscleblind, a gene required for photoreceptor differentiation in Drosophila, encodes novel nuclear Cys3His-type zinc-finger-containing proteins.* Development, 1997. **124**(21): p. 4321-31.
45. Vicente, M., et al., *Muscleblind isoforms are functionally distinct and regulate alpha-actinin splicing.* Differentiation, 2007.
46. Kania, A., et al., *P-element mutations affecting embryonic peripheral nervous system development in Drosophila melanogaster.* Genetics, 1995. **139**(4): p. 1663-78.
47. Shang, C.A., et al., *Genes induced by growth hormone in a model of adipogenic differentiation.* Mol Cell Endocrinol, 2002. **189**(1-2): p. 213-9.
48. Phillips, R.L., et al., *The genetic program of hematopoietic stem cells.* Science, 2000. **288**(5471): p. 1635-40.
49. Romero, D.G., et al., *Angiotensin II early regulated genes in H295R human adrenocortical cells.* Physiol Genomics, 2004. **19**(1): p. 106-16.
50. Day, J.W. and L.P. Ranum, *RNA pathogenesis of the myotonic dystrophies.* Neuromuscul Disord, 2005. **15**(1): p. 5-16.
51. Miller, J.W., et al., *Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy.* Embo J, 2000. **19**(17): p. 4439-48.
52. Mutsuddi, M., et al., *The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in Drosophila.* Curr Biol, 2004. **14**(4): p. 302-8.
53. Rudnicki, D.D., et al., *Huntington's disease--like 2 is associated with CUG repeat-containing RNA foci.* Ann Neurol, 2007. **61**(3): p. 272-82.
54. Dean, B., et al., *Gene expression profiling in Brodmann's area 46 from subjects with schizophrenia.* Aust N Z J Psychiatry, 2007. **41**(4): p. 308-20.
55. Edgar, A.J., et al., *Upregulated genes in sporadic, idiopathic pulmonary arterial hypertension.* Respir Res, 2006. **7**: p. 1.

56. Ranum, L.P. and J.W. Day, *Dominantly inherited, non-coding microsatellite expansion disorders.* Curr Opin Genet Dev, 2002. **12**(3): p. 266-71.
57. Aslanidis, C., et al., *Cloning of the essential myotonic dystrophy region and mapping of the putative defect.* Nature, 1992. **355**(6360): p. 548-51.
58. 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, 1992. **69**(2): p. 385.
59. Buxton, J., et al., *Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy.* Nature, 1992. **355**(6360): p. 547-8.
60. Fu, Y.H., et al., *An unstable triplet repeat in a gene related to myotonic muscular dystrophy.* Science, 1992. **255**(5049): p. 1256-8.
61. Harley, H.G., et al., *Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy.* Nature, 1992. **355**(6360): p. 545-6.
62. Mahadevan, M., et al., *Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene.* Science, 1992. **255**(5049): p. 1253-5.
63. Liquori, C.L., et al., *Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9.* Science, 2001. **293**(5531): p. 864-7.
64. Jansen, G., et al., *Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice.* Nat Genet, 1996. **13**(3): p. 316-24.
65. Reddy, S., et al., *Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy.* Nat Genet, 1996. **13**(3): p. 325-35.
66. Frisch, R., et al., *Effect of triplet repeat expansion on chromatin structure and expression of DMPK and neighboring genes, SIX5 and DMWD, in myotonic dystrophy.* Mol Genet Metab, 2001. **74**(1-2): p. 281-91.
67. Taneja, K.L., et al., *Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues.* J Cell Biol, 1995. **128**(6): p. 995-1002.
68. Wang, J., et al., *Myotonic dystrophy: evidence for a possible dominant-negative RNA mutation.* Hum Mol Genet, 1995. **4**(4): p. 599-606.

69. Mankodi, A., et al., *Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat.* Science, 2000. **289**(5485): p. 1769-73.
70. Michalowski, S., et al., *Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein.* Nucleic Acids Res, 1999. **27**(17): p. 3534-42.
71. Ebralidze, A., et al., *RNA leaching of transcription factors disrupts transcription in myotonic dystrophy.* Science, 2004. **303**(5656): p. 383-7.
72. Kim, H. and J. Yin, *Robust growth of human immunodeficiency virus type 1 (HIV-1).* Biophys J, 2005. **89**(4): p. 2210-21.
73. Mankodi, A., et al., *Nuclear RNA foci in the heart in myotonic dystrophy.* Circ Res, 2005. **97**(11): p. 1152-5.
74. Jiang, H., et al., *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, 2004. **13**(24): p. 3079-88.
75. Mankodi, A., et al., *Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2.* Ann Neurol, 2003. **54**(6): p. 760-8.
76. Mankodi, A., et al., *Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2.* Hum Mol Genet, 2001. **10**(19): p. 2165-70.
77. Harbison, S.T., et al., *Quantitative genomics of starvation stress resistance in Drosophila.* Genome Biol, 2005. **6**(4): p. R36.
78. Ho, T.H., et al., *Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy.* J Cell Sci, 2005. **118**(Pt 13): p. 2923-33.
79. Charlet, B.N., et al., *Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing.* Mol Cell, 2002. **10**(1): p. 45-53.
80. Savkur, R.S., A.V. Philips, and T.A. Cooper, *Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy.* Nat Genet, 2001. **29**(1): p. 40-7.
81. Ho, T.H., et al., *Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy.* Hum Mol Genet, 2005. **14**(11): p. 1539-47.

82. Krol, J., et al., *Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets.* Mol Cell, 2007. **25**(4): p. 575-86.
83. Houseley, J.M., et al., *Myotonic dystrophy associated expanded CUG repeat muscleblind positive ribonuclear foci are not toxic to Drosophila.* Hum Mol Genet, 2005. **14**(6): p. 873-83.
84. Haro, M.D., et al., *MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of Myotonic Dystrophy Type 1.* Hum Mol Genet, 2006.
85. Monferrer-Sales, L., *Un modelo en Drosophila del mecanismo de patogénesis de las expansiones CTG en la distrofia miotónica*, in *Departamento de Genética*. 2007, Universidad de Valencia: Valencia.
86. Pette, D. and R.S. Staron, *Mammalian skeletal muscle fiber type transitions.* Int Rev Cytol, 1997. **170**: p. 143-223.
87. Reininghaus, J., et al., *The myotonic mouse mutant ADR: physiological and histochemical properties of muscle.* Muscle Nerve, 1988. **11**(5): p. 433-9.
88. Jockusch, H., et al., *Reduction of myosin-light-chain phosphorylation and of parvalbumin content in myotonic mouse muscle and its reversal by tocainide.* Eur J Biochem, 1988. **171**(1-2): p. 101-5.
89. Schleef, M., et al., *Subtractive cDNA cloning as a tool to analyse secondary effects of a muscle disease. Characterization of affected genes in the myotonic ADR mouse.* Neuromuscul Disord, 1994. **4**(3): p. 205-17.
90. Schimmelpfeng, J., H. Jockusch, and P. Heimann, *Increased density of satellite cells in the absence of fibre degeneration in muscle of myotonic mice.* Cell Tissue Res, 1987. **249**(2): p. 351-7.
91. Farkas-Bargeton, E., et al., *Immaturity of muscle fibers in the congenital form of myotonic dystrophy: its consequences and its origin.* J Neurol Sci, 1988. **83**(2-3): p. 145-59.
92. Begemann, G., et al., *The Drosophila orphan nuclear receptor seven-up requires the Ras pathway for its function in photoreceptor determination.* Development, 1995. **121**(1): p. 225-235.
93. Kramer, S., S.R. West, and Y. Hiromi, *Cell fate control in the Drosophila retina by the orphan receptor seven-up: its role in the decisions mediated by the ras signaling pathway.* Development, 1995. **121**(5): p. 1361-1372.

94. Brown, N.H., *Integrins hold Drosophila together.* Bioessays, 1993. **15**(6): p. 383-90.
95. Kiehart, D.P. and R. Feghali, *Cytoplasmic myosin from Drosophila melanogaster.* J. Cell Biol., 1986. **103**(4): p. 1517-1525.
96. Prokop, A., et al., *Absence of PS integrins or laminin A affects extracellular adhesion, but not intracellular assembly, of hemiadherens and neuromuscular junctions in Drosophila embryos.* Dev Biol, 1998. **196**(1): p. 58-76.
97. Sandmann, T., et al., *A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development.* Dev Cell, 2006. **10**(6): p. 797-807.
98. Marek, K.W., et al., *A genetic analysis of synaptic development: pre- and postsynaptic dCBP control transmitter release at the Drosophila NMJ.* Neuron, 2000. **25**(3): p. 537-47.
99. SenGupta, D.J., et al., *A three-hybrid system to detect RNA-protein interactions in vivo.* Proc Natl Acad Sci U S A, 1996. **93**(16): p. 8496-501.
100. Bernstein, D.S., et al., *Analyzing mRNA-protein complexes using a yeast three-hybrid system.* Methods, 2002. **26**(2): p. 123-41.
101. O'Neill, J.W. and E. Bier, *Double-label in situ hybridization using biotin and digoxigenin-tagged RNA probes.* Biotechniques, 1994. **17**(5): p. 870, 874-5.
102. Ule, J., et al., *CLIP identifies Nova-regulated RNA networks in the brain.* Science, 2003. **302**(5648): p. 1212-5.
103. Sambrook, J.M., P.; Russell, D., *Molecular cloning: A Laboratory Manual*. Third ed. 2001: CSHL Press. 2344.
104. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucl. Acids Res., 2003. **31**(13): p. 3406-3415.
105. Ito, H., et al., *Transformation of intact yeast cells treated with alkali cations.* J Bacteriol, 1983. **153**(1): p. 163-8.
106. various, *Current Protocols in Molecular Biology*. 2006: John Wiley & Sons, Inc. 4800.
107. Monferrer, L. and R. Artero, *An Interspecific Functional Complementation Test in Drosophila for Introductory Genetics Laboratory Courses.* J Hered, 2006.
108. Harper, P., *Myotonic dystrophy*. 2001, London: Saunders.

109. Prokopenko, S.N., et al., *Mutations affecting the development of the peripheral nervous system in Drosophila: a molecular screen for novel proteins.* Genetics, 2000. **156**(4): p. 1691-715.
110. Kenyon, C., *The plasticity of aging: insights from long-lived mutants.* Cell, 2005. **120**(4): p. 449-60.
111. Liang, H., et al., *Genetic mouse models of extended lifespan.* Exp Gerontol, 2003. **38**(11-12): p. 1353-64.
112. Seino, S. and G.I. Bell, *Alternative splicing of human insulin receptor messenger RNA.* Biochem Biophys Res Commun, 1989. **159**(1): p. 312-6.
113. Fernandez, R., et al., *The Drosophila insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential.* Embo J, 1995. **14**(14): p. 3373-84.
114. Mankodi, A., et al., *Expanded CUG Repeats Trigger Aberrant Splicing of ClC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy.* Molecular Cell, 2002. **10**(1): p. 35.
115. De Luca, A., et al., *Evaluation of the antimyotonic activity of mexiletine and some new analogs on sodium currents of single muscle fibers and on the abnormal excitability of the myotonic ADR mouse.* J Pharmacol Exp Ther, 1997. **282**(1): p. 93-100.
116. Orso, G., et al., *Disease-related phenotypes in a Drosophila model of hereditary spastic paraplegia are ameliorated by treatment with vinblastine.* J Clin Invest, 2005. **115**(11): p. 3026-34.
117. Wicher, D., C. Walther, and C. Wicher, *Non-synaptic ion channels in insects -- basic properties of currents and their modulation in neurons and skeletal muscles.* Progress in Neurobiology, 2001. **64**(5): p. 431.
118. Ludatscher, R.M., et al., *Myotonia dystrophica with heart involvement: an electron microscopic study of skeletal, cardiac, and smooth muscle.* J Clin Pathol, 1978. **31**(11): p. 1057-64.
119. Gromak, N., et al., *Antagonistic regulation of alpha-actinin alternative splicing by CELF proteins and polypyrimidine tract binding protein.* Rna, 2003. **9**(4): p. 443-56.
120. Fyrberg, E., et al., *Molecular genetics of Drosophila alpha-actinin: mutant alleles disrupt Z disc integrity and muscle insertions.* J Cell Biol, 1990. **110**(6): p. 1999-2011.

121. Benoist, P., et al., *Differential Muscle-type Expression of the Drosophila Troponin T Gene. A 3-BASE PAIR MICROEXON IS INVOLVED IN VISCERAL AND ADULT HYPODERMIC MUSCLE SPECIFICATION.* J. Biol. Chem., 1998. **273**(13): p. 7538-7546.
122. Pascual-Lucas, M., *Muscleblind, relevancia clínica y análisis de la función molecular en Drosophila melanogaster*, in *Departamento de Genética*. 2005, Universidad de Valencia: Valencia.
123. Robard, C., A. Daviau, and M. Di Fruscio, *Phosphorylation status of the Kep1 protein alters its affinity for its protein binding partner alternative splicing factor ASF/SF2.* Biochem J, 2006. **400**(1): p. 91-7.
124. Parra-Palau, J.L., et al., *The Drosophila protein kinase LK6 is regulated by ERK and phosphorylates the eukaryotic initiation factor eIF4E in vivo.* Biochem J, 2005. **385**(Pt 3): p. 695-702.
125. Chaudhuri, T., et al., *Role of the Fetal and [alpha]/[beta] Exons in the Function of Fast Skeletal Troponin T Isoforms: Correlation with Altered Ca2+ Regulation Associated with Development.* Journal of Molecular Biology, 2005. **352**(1): p. 58.
126. Vicente-Crespo, M., Pascual M., Fernandez-Costa, J.M., Miranda, M.E., Zhou, L., Smith C. and Artero, R.D., *Genetic and cell culture data implicate Muscleblind in troponin T alternative splicing and programmed cell death in Drosophila.* in preparation.
127. Zhou, L., et al., *Michelob_x is the missing inhibitor of apoptosis protein antagonist in mosquito genomes.* EMBO Rep, 2005. **6**(8): p. 769-74.
128. Allemand, E., et al., *Distinctive features of Drosophila alternative splicing factor RS domain: implication for specific phosphorylation, shuttling, and splicing activation.* Mol Cell Biol, 2001. **21**(4): p. 1345-59.
129. Singh, R. and J. Valcarcel, *Building specificity with nonspecific RNA-binding proteins.* Nat Struct Mol Biol, 2005. **12**(8): p. 645-53.
130. Bossis, G., et al., *Down-regulation of c-Fos/c-Jun AP-1 dimer activity by sumoylation.* Mol Cell Biol, 2005. **25**(16): p. 6964-79.
131. Wilson, V.G. and D. Rangasamy, *Intracellular Targeting of Proteins by Sumoylation.* Experimental Cell Research, 2001. **271**(1): p. 57.
132. Kiger, J.A., Jr., et al., *Tissue remodeling during maturation of the Drosophila wing.* Dev Biol, 2007. **301**(1): p. 178-91.

133. Brabant, M.C., et al., *Distinct spatial and temporal functions for PS integrins during Drosophila wing morphogenesis.* Development, 1996. **122**(10): p. 3307-17.
134. Murray, M.A., L.I. Fessler, and J. Palka, *Changing distributions of extracellular matrix components during early wing morphogenesis in Drosophila.* Dev Biol, 1995. **168**(1): p. 150-65.
135. Fessler, J.H. and L.I. Fessler, *Drosophila extracellular matrix.* Annu Rev Cell Biol, 1989. **5**: p. 309-39.
136. Brabant, M.C. and D.L. Brower, *PS2 integrin requirements in Drosophila embryo and wing morphogenesis.* Dev Biol, 1993. **157**(1): p. 49-59.
137. Wilcox, M., A. DiAntonio, and M. Leptin, *The function of PS integrins in Drosophila wing morphogenesis.* Development, 1989. **107**(4): p. 891-7.
138. Bokel, C. and N.H. Brown, *Integrins in development: moving on, responding to, and sticking to the extracellular matrix.* Dev Cell, 2002. **3**(3): p. 311-21.
139. Brown, N.H., *Cell-cell adhesion via the ECM: integrin genetics in fly and worm.* Matrix Biol, 2000. **19**(3): p. 191-201.
140. Bogaert, T., N. Brown, and M. Wilcox, *The Drosophila PS2 antigen is an invertebrate integrin that, like the fibronectin receptor, becomes localized to muscle attachments.* Cell, 1987. **51**(6): p. 929-40.
141. Brower, D.L., et al., *Nonequivalent requirements for PS1 and PS2 integrin at cell attachments in Drosophila: genetic analysis of the alpha PS1 integrin subunit.* Development, 1995. **121**(5): p. 1311-20.
142. Leptin, M., et al., *The function of PS integrins during Drosophila embryogenesis.* Cell, 1989. **56**(3): p. 401-8.
143. Lee, S.B., et al., *blistery encodes Drosophila tensin protein and interacts with integrin and the JNK signaling pathway during wing development.* Development, 2003. **130**(17): p. 4001-10.
144. Borchiellini, C., J. Coulon, and Y. Le Parco, *The function of type IV collagen during Drosophila muscle development.* Mech Dev, 1996. **58**(1-2): p. 179-91.
145. Fyrberg, E., et al., *Drosophila melanogaster troponin-T mutations engender three distinct syndromes of myofibrillar abnormalities.* J Mol Biol, 1990. **216**(3): p. 657-75.
146. Jurica, M.S. and M.J. Moore, *Pre-mRNA splicing: awash in a sea of proteins.* Mol Cell, 2003. **12**(1): p. 5-14.

147. Makarov, E.M., et al., *Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome.* Science, 2002. **298**(5601): p. 2205-8.
148. Fardaei, M., et al., *In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts.* Nucleic Acids Res, 2001. **29**(13): p. 2766-71.
149. Wheeler, T.M., M.C. Krym, and C.A. Thornton, *Ribonuclear foci at the neuromuscular junction in myotonic dystrophy type 1.* Neuromuscul Disord, 2007. **17**(3): p. 242-7.
150. Park, J.W., et al., *Identification of alternative splicing regulators by RNA interference in Drosophila.* Proc Natl Acad Sci U S A, 2004. **101**(45): p. 15974-9.
151. Webster, P.J., et al., *Translational repressor bruno plays multiple roles in development and is widely conserved.* Genes Dev, 1997. **11**(19): p. 2510-21.
152. Delaunay, J., et al., *The Drosophila Bruno paralogue Bru-3 specifically binds the EDEN translational repression element.* Nucleic Acids Res, 2004. **32**(10): p. 3070-82.
153. Ranum, L.P. and T.A. Cooper, *RNA-Mediated Neuromuscular Disorders.* Annu Rev Neurosci, 2006.
154. Ranum, L.P. and J.W. Day, *Myotonic dystrophy: RNA pathogenesis comes into focus.* Am J Hum Genet, 2004. **74**(5): p. 793-804.
155. Snee, M.J., et al., *A late phase of Oskar accumulation is crucial for posterior patterning of the Drosophila embryo, and is blocked by ectopic expression of Bruno.* Differentiation, 2007. **75**(3): p. 246-55.
156. Roulier, E.M., C. Fyrberg, and E. Fyrberg, *Perturbations of Drosophila alpha-actinin cause muscle paralysis, weakness, and atrophy but do not confer obvious nonmuscle phenotypes.* J Cell Biol, 1992. **116**(4): p. 911-22.
157. Machuca-Tzili, L., et al., *Flies deficient in Muscleblind protein model features of myotonic dystrophy with altered splice forms of Z-band associated transcripts.* Hum Genet, 2006. **120**(4): p. 487-499.
158. Faulkner, G., et al., *ZASP: a new Z-band alternatively spliced PDZ-motif protein.* J Cell Biol, 1999. **146**(2): p. 465-75.
159. Ishiura, S., et al., *Regulation of splicing by MBNL and CELF family of RNA-binding protein.* Acta Myol, 2005. **24**(2): p. 74-7.
160. Nissim-Rafinia, M. and B. Kerem, *Splicing regulation as a potential genetic modifier.* Trends Genet, 2002. **18**(3): p. 123-7.

161. Wollerton, M.C., et al., *Differential alternative splicing activity of isoforms of polypyrimidine tract binding protein (PTB).* Rna, 2001. **7**(6): p. 819-32.
162. Kirsh, O., et al., *The SUMO E3 ligase RanBP2 promotes modification of the HDAC4 deacetylase.* Embo J, 2002. **21**(11): p. 2682-91.
163. Stege, G.J., et al., *Importance of the ATP-binding domain and nucleolar localization domain of HSP72 in the protection of nuclear proteins against heat-induced aggregation.* Exp Cell Res, 1994. **214**(1): p. 279-84.
164. Yano, M., et al., *A novel function of 14-3-3 protein: 14-3-3zeta is a heat-shock-related molecular chaperone that dissolves thermal-aggregated proteins.* Mol Biol Cell, 2006. **17**(11): p. 4769-79.
165. Usuki, F. and S. Ishiura, *Expanded CTG repeats in myotonin protein kinase increase susceptibility to oxidative stress.* Neuroreport, 1998. **9**(10): p. 2291-6.
166. Schwerk, C. and K. Schulze-Osthoff, *Regulation of apoptosis by alternative pre-mRNA splicing.* Mol Cell, 2005. **19**(1): p. 1-13.
167. Kim, D.H., et al., *HnRNP H inhibits nuclear export of mRNA containing expanded CUG repeats and a distal branch point sequence.* Nucleic Acids Res, 2005. **33**(12): p. 3866-74.