DEVELOPING A SIMPLE TITRATION CELL-BASED SPLICING ASSAY

by

JAKE GACKE

A THESIS

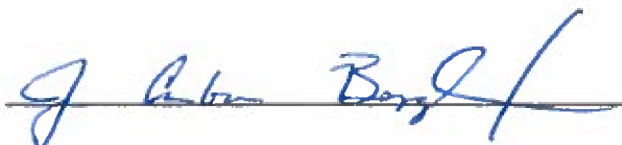
Presented to the Department of Chemistry and Biochemistry
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

June, 2015

An Abstract of the Thesis of

Jake Gacke for the degree of Bachelor of Science
in the Department of Chemistry and Biochemistry to be taken June, 2015

Title: Developing a Simple Titration Cell-Based Splicing Assay

Approved: 

J. Andrew Berglund

Muscleblind-like 1 (MBNL1) is a protein that regulates the expression of genes at the step known as pre-mRNA splicing or alternative splicing. Cells use alternative splicing to generate multiple mRNAs (and proteins) from a single gene. MBNL1-mediated splicing is dependent on MBNL1's ability to bind pre-mRNA targets, which is possible through two critical pairs of tandem zinc finger RNA binding motifs. However, these two distinct RNA binding domains (zinc fingers 1-2 and 3-4) have non-equivalent splicing function in MBNL1 dependent alternative splicing events. The zinc finger pair (1-2) has been shown to contribute significantly higher RNA binding affinity and splicing activity for several RNA substrates compared to the zinc finger pair (3-4). Therefore, a splicing factor with greater splicing activity might be constructed by increasing MBNL1's affinity through the creation of synthetic MBNL1 containing two copies of zinc finger pair (1-2).

To test this hypothesis, an MBNL1 variant replacing the zinc finger pair (3-4) domain with another copy of zinc finger pair (1-2) domain was created – MBNL1 (1-2, 1-2). A simple cell-based splicing assay was developed to compare the splicing

activities of WT-MBNL1 and MBNL1 (1-2, 1-2). Preliminary results showed that MBNL1 (1-2, 1-2) had similar splicing activity compared to the wild type MBNL1.

Acknowledgements

I would like to express my utmost gratitude to Professor Berglund for offering me the privilege to explore meaningful scientific research, as well as entrusting me to work independently. I would also like to thank Melissa Hale for her guidance and mentorship throughout my research because without her none of this would have been possible. Thank you Berglund lab for all of your help. Finally, I would like to thank my family for their unconditional love and support.

.

Table of Contents

INTRODUCTION 1

Alternative splicing 1

Myotonic dystrophy pathology 3

Muscleblind-like 1 (MBNL1) 4

Background experiments hint towards the possibility of an RNA-binding protein with increased affinity and activity 7

METHODS 12

Cell culture and transfection 13

Western blot analysis 14

Cell-based splicing assay 14

RESULTS 16

Western blot analysis confirms successful dosing 16

Both WT-MBNL1 and MBNL1 (1-2, 1-2) exhibit a splicing sensitivity response 19

DISCUSSION 23

Western blot analysis and splicing assay results confirm successful dosing simulation 23

Applications and advantages of this dosage simulation model (transfection titration) 23

Challenges faced in developing the transfection titration model 25

FUTURE DIRECTIONS AND REFINEMENT 28

REFERENCES 30

List of Figures

Figure 1: Splicing and assembly of the spliceosome 3

Figure 2: Cassette exon alternative splicing event and possible products 4

Figure 3: MBNL1 co-localizes with CUG repeat RNA and forms nuclear foci in neurons of DM1 patients 5

Figure 4: Schematic representation of MBNL1's structure including sequence alignment and crystal structure of the zinc finger domains 7

Figure 5: Graphic representation of the two constructed MBNL1 variants 9

Figure 6: Workflow illustration of a transfection titration cell-based splicing assay 13

Figure 7: MBNL1 (1-2, 1-2) is expressed at low levels in an overexpression system 17

Figure 8: Qualitative and quantitative confirmation of a successful dosing system 18

Figure 9: Splicing activity response to both constructs as a function of plasmid concentration 21

List of Tables

Table 1: MBNL1 (1-2, 1-2) shows similar splicing activity in comparison to WT-MBNL1 12

INTRODUCTION

Alternative splicing

Genes contain all genetic information necessary for cell survival in the form of DNA. Each gene encodes instructions to create and assemble proteins. These proteins are directly responsible for performing cellular functions vital for cell viability. In order to produce proteins from the corresponding DNA, a complex and highly regulated process must be carried out. First, DNA is converted into pre-mRNA through transcription. The pre-mRNA contains important coding regions of the protein called exons as well as untranslated regions referred to as introns. Before translation into protein, pre-mRNA is processed into mature RNA, commonly referred to as messenger RNA (mRNA). One of these processing steps is splicing. Splicing is a key process in which introns are removed from the pre-mRNA and the remaining exons are stitched together (Figure 1) [1].

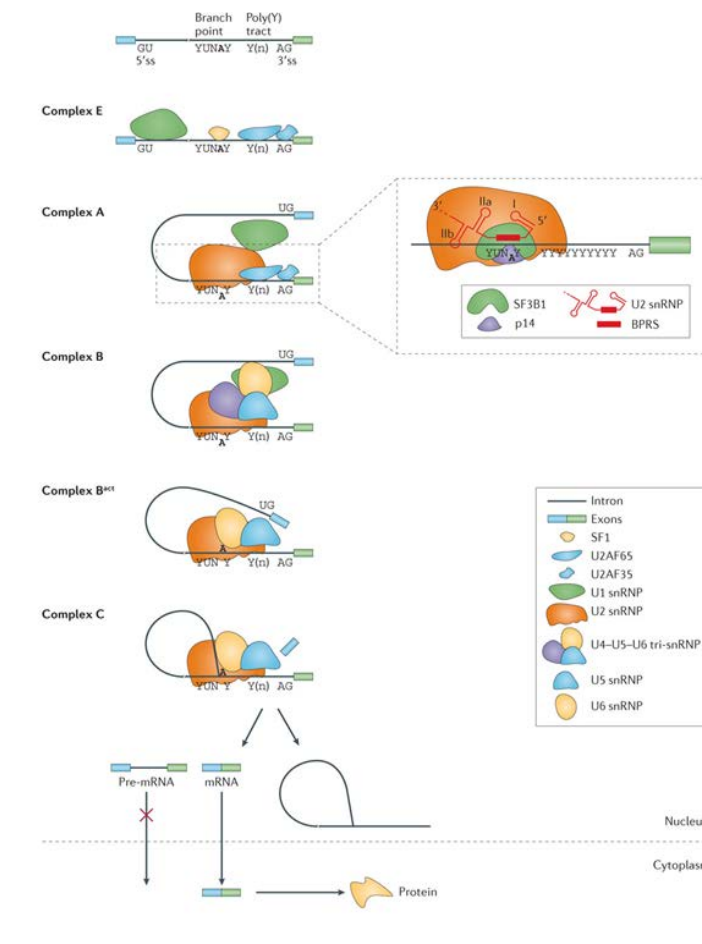


Figure 1. Splicing and the assembly of the spliceosome. Pre-mRNA with intronic sequences (thin line) and exonic sequences (coloured boxes) is represented above with the consensus sequences at the intron 5′ and 3′ ends (Y represents pyrimidines, N represents any nucleotide). The branch point adenosine is represented in bold. Many cis-elements and factors assist in the assembly of the spliceosome. The spliceosome is assembled sequentially forming the complexes illustrated. Once assembled, the splicing reaction is catalyzed, generating a mature RNA and releasing the intron in a lariat conformation. Box 1. BPRS, branch point recognition sequence; SF1, splicing factor 1; ss, splice site; U2AF, auxiliary factor, U2 snRNP (small nuclear ribonucleoprotein) (*nature reviews* 2012) [1].

One of many specialized splicing events is alternative splicing. Alternative splicing allows the cell to select which exons are ultimately included or excluded in the final processed mRNA, which in turn dictates which protein isoform is produced (Figure 2)[2]. Why is alternative splicing so important? Alternative splicing enables cells to express multiple protein isoforms from the same gene, thus generating increased protein diversity. A diverse proteome (entire set of proteins expressed by a genome) is critical in regulating gene expression both temporally and spatially[3]. Therefore, alternative splicing offers a method for regulating what protein isoforms are produced across various tissues and throughout development. It is estimated that over 90 percent

of human genes are alternatively spliced[4]. As such, the mis-regulation of alternative splicing, which can produce a dramatic shift in the potential protein isoforms expressed, can prove disastrous for the cell. Disease-like symptoms may develop as a consequence of aberrant splicing. It is estimated that over 50 percent of all diseases contain a mis-regulated splicing element[5]. Mis-regulation of alternative splicing plays a pivotal role in myotonic dystrophy Type 1 (DM1) pathology.

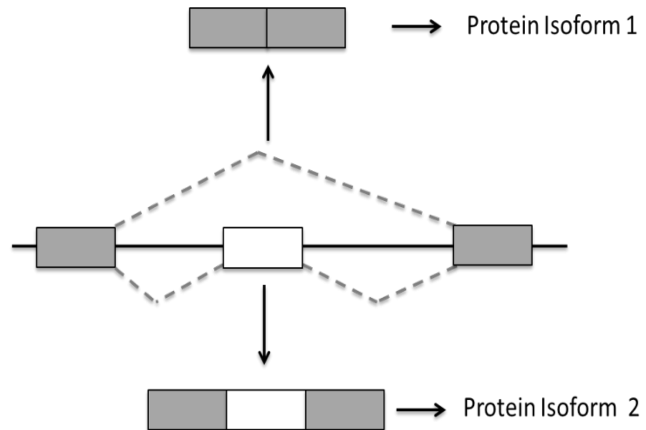


Figure 2. Cassette exon alternative splicing event. In a cassette exon event, the alternatively spliced exon is either included or "skipped." The pre-mRNA transcript is shown in the middle. Constitutively spliced exons are shown as grey boxes and the alternatively spliced exon is shown as a white box. The line represents introns. Two different mRNAs may be generated, one in which the exon is included and one in which the exon is excluded. As a result, two different protein isoforms can be produced[2].

Myotonic dystrophy pathology

Myotonic dystrophy (type 1) is the most common form of adult-onset muscular dystrophy affecting one in every 8,000 individuals. Those afflicted with DM1 can experience a wide range of symptoms that occur within multiple systems of the body. Varying in severity, these symptoms include myotonia (hyperexcitability of skeletal muscle), muscle wasting, insulin resistance, cardiac defects, cataracts, and cognitive dysfunction[5]. Pathology of DM1 is highly studied and relatively well understood. DM1 is characterized by a CTG trinucleotide repeat expansion located within the 3' untranslated region of the DMPK (dystrophia myotonica protein kinase) gene[5]. Upon

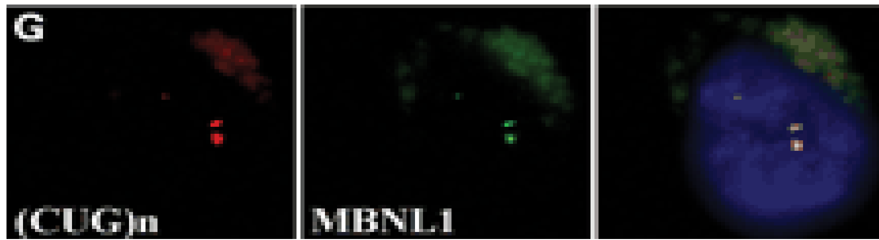


Figure 3. MBNL1 co-localizes with CUG repeat RNA in the nucleus of neurons in DM1 patients, forming foci. CUG repeats are shown in red (visualized via fluorescent *in situ* hybridization, FISH), MBNL1 protein is shown in green (visualized via immunofluorescence), and DAP1 (fluorescent DNA stain) in blue. No foci or co-localization is observed from non-DM1 individuals[8].

transcription, the CTG repeats in DNA are converted into CUG repeats in RNA. Consequentially, the CUG repeats act as toxic RNA. These CUG repeats sequester the alternative splicing factor MBNL1 (MBNL1 is discussed in detail on pages 5-7) [7]. MBNL1's high affinity for YGCY (UGCU) sequences in the toxic RNA is responsible for the formation of nuclear foci (a combination of toxic RNA and proteins in the nucleus) (Figure 3) [8].

The co-localization of MBNL1 and toxic CUG-repeat RNA effectively reduces the concentration of free MBNL1 within the cell. If MBNL1 is bound to the toxic RNA, it is unable to perform its primary regulatory function. Aberrant splicing of MBNL1's pre-mRNA targets is responsible for a DM1 phenotype.

Muscleblind-like 1 (MBNL1)

Muscleblind-like protein 1 (MBNL1) is involved in numerous cellular processes including mRNA localization, cell fate, and most importantly, alternative splicing. MBNL1 is a master splicing regulator that promotes both exon inclusion and exclusion events, acting as a splicing enhancer or repressor[9]. As an alternative splicing regulator, MBNL1 is involved in the alternative splicing of numerous pre-mRNA transcripts including its own. For instance, MBNL1 promotes insulin receptor exon 11 inclusion[10],

skeletal muscle chloride channel exon 7A exclusion[11], and exon 5 exclusion of its own pre-mRNA transcript (among many others) [12].

The mechanism through which MBNL1 regulates alternative splicing has yet to be deduced. One possible mechanistic explanation revolves around the idea that MBNL1 inhibits splicing by directly interfering with the splicing machinery upon binding its pre-mRNA transcripts. Recent work has shown that MBNL1 competes directly with the essential splicing factor U2AF65 (a splicing factor involved in the recruitment of a snRNP necessary for spliceosome assembly) (Figure 1) during the MBNL1-mediated splicing event of exon 5 in the cardiac troponin T pre-mRNA. When U2AF65 is prevented from binding the pre-mRNA, the U2 snRNP can no longer be recruited in the formation of the spliceosome and the following exon 5 is excluded[13]. Although this is only one example of an alternative splicing event, the general consensus follows that MBNL1's splicing activity is dependent on its ability to bind its pre-mRNA targets. Generally, an exon inclusion event is promoted if MBNL1 binds downstream from an alternatively spliced exon. If MBNL1 binds upstream, an exon exclusion event is promoted[7]. Conserved RNA binding domains called zinc fingers enable MBNL1 to bind to its pre-mRNA targets.

Zinc finger (ZnF) binding motifs are highly conserved throughout many nucleic acid binding proteins[14]. There are two pairs of zinc finger RNA binding motifs within muscleblind-like 1. These four zinc fingers are split into two pairs, ZnF1-2 and ZnF3-4. Each ZnF pair (1-2 and 3-4) form two distinct RNA binding domains, which are separated by a 110 amino acid linker. Each of the four tandem zinc fingers are CCCH type (meaning three cysteine residues and one histidine residue) and assist in the

chelation of a zinc ion and act like fingers grasping the Watson-Crick face of single-stranded pre-mRNA transcripts[15]. In vitro RNA binding experiments reveal that MBNL1 selectively recognizes and binds to YGCY (where Ys are pyrimidines, either C or U) sequences in its RNA substrates, explaining its high affinity for toxic CUG repeat RNA[7]. It is important to note, zinc fingers 1 and 3 show high sequence similarity as do zinc fingers 2 and 4 (Figure 4) [15].

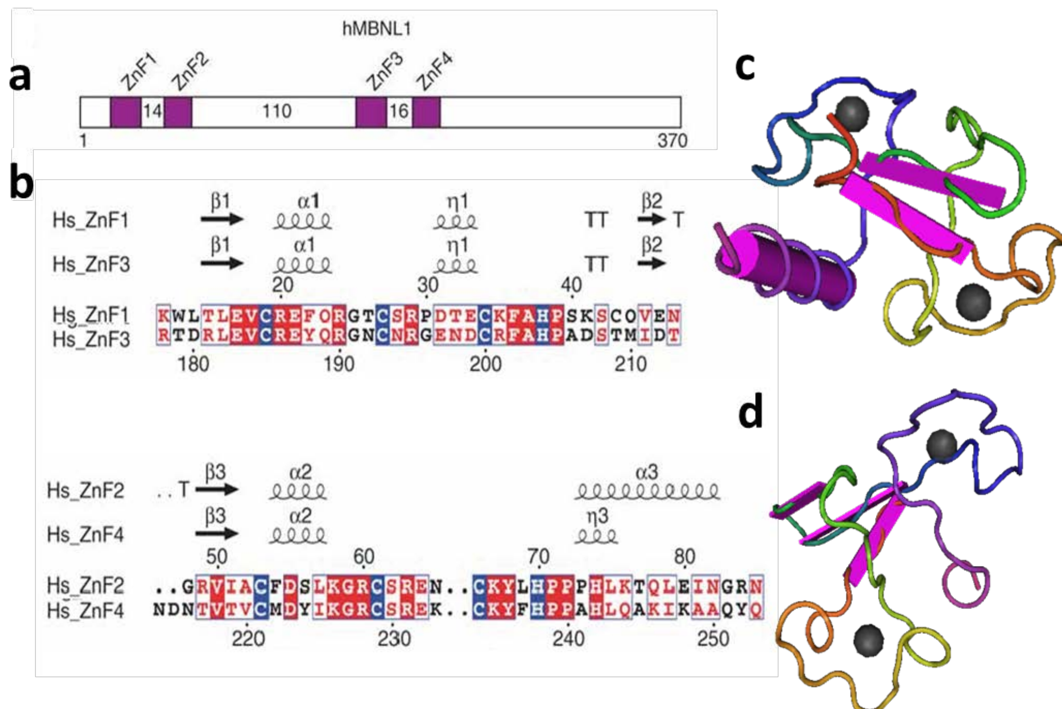


Figure 4. (a) Schematic representation of human MBNL1's zinc finger domains. Boxed numbers indicate the length of linker segments. (b) Sequence alignment of tandem ZnF1-2 and ZnF3-4 tandem zinc finger domains from human MBNL1. The residues involved in the chelation of zinc ions are highlighted in blue. Secondary-structure alignment of human MBNL1 ZnF (1-2 and 3-4) is shown above the sequences. Numbering above the sequences corresponds to human MBNL1 ZnF1-2 and numbering below corresponds to ZnF3-4. (c) Crystal structure of the ZnF (1-2) domain, PDB ID: 3D2N. (d) Crystal structure of the ZnF3-4 domain, PDB ID: 3D2Q. Grey spheres denote zinc ions[15].

The ZnFs also show high secondary structure similarity apart from an extended alpha helix towards the C-terminus of ZnF2 as compared to ZnF4[15]. A complete crystal structure of MBNL1 has yet to be determined, but crystal structures of both ZnF1-2 and ZnF3-4 exist in isolation (Figure 4).

Background experiments hint towards the possibility of an RNA-binding protein with increased affinity and activity

Our lab has recently shown that ZnF1-2 and ZnF3-4 RNA-binding domains do not possess equivalent splicing activity for MBNL1-regulated splicing events nor do they possess equivalent RNA-binding affinities. In the experiment, key residues, implicated in the interaction between protein and RNA, were mutated to alanine residues, which eliminated the ability to bind RNA targets. These mutations thereby resulted in an inactive ZnF domain at the site of mutagenesis enabling us to test the activity and affinity of each domain (1-2 and 3-4) individually. A mutation of ZnF3-4, producing a mutant MBNL1 containing only an intact, functional ZnF1-2, retained about 80 percent splicing activity, as well as 80 percent RNA-binding affinity normalized for several tested targets. Conversely, a mutation of ZnF1-2, producing a solely functional ZnF3-4 mutant, only retained about 50 percent splicing activity and 40 percent RNA-binding affinity in comparison to wild type MBNL1[16].

These results suggest ZnF1-2 has greater affinity and splicing activity than ZnF3-4, which poses the question: Can we engineer an MBNL1 protein with increased RNA-binding affinity and splicing activity in comparison to wild type? The idea stems from the differences in affinity and splicing activity between the ZnF pairs. Because ZnF1-2 exhibits a greater affinity and splicing activity than ZnF3-4, it should be

possible to engineer an enhanced MBNL1 with two ZnF1-2 domains. A successfully engineered MBNL1 with greater RNA-binding affinity and splicing activity may serve as a possible therapeutic protein by either displacing sequestered wild type MBNL1 or by more effectively regulating alternative splicing in DM1 patients. The displaced wild type MBNL1 could theoretically resume function and alleviate DM1 symptoms by rescuing mis-splicing events. Creating these MBNL1 variants also provides a tool to study the importance of the ZnF domains to MBNL1 function.

In order to answer the aforementioned question, two protein constructs were engineered. The first construct (WT-MBNL1) represents wild type. WT-MBNL1 includes both ZnF1-2 and ZnF3-4. The other construct replaces ZnF3-4 with ZnF1-2 to generate a double ZnF1-2 MBNL1 or MBNL1 (1-2, 1-2) protein (Hale, M., unpublished). Each construct contains a human influenza hemagglutinin (HA) at the N-terminus for western blot detection (Figure 5). The lab predicted the overall splicing activity of the constructs would increase with ZnF1-2 content to produce the following trend: MBNL1 (1-2, 1-2) > wild type (WT) MBNL1. To test the splicing activity of these proteins, a cell-based splicing assay was utilized.

In this assay, HeLa cells (immortal cell line derived from a human cervical adenocarcinoma) are co-transfected with a reporter minigene and plasmid DNA

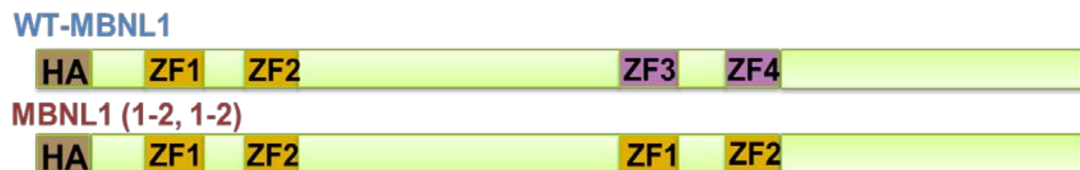


Figure 5. Graphic representation of the two engineered MBNL1 constructs. MBNL1 variants each contain an HA tag (brown box). Zinc fingers 1 and 2 and represented by yellow boxes and zinc fingers 3 and 4 are represented by purple boxes.

containing recombinant WT or synthetic MBNL1 protein. The reporter minigenes are, as their name implies, miniature versions of genes that are involved in an alternatively splicing event. Each minigene contains an MBNL1-regulated alternatively spliced exon flanked by two intervening introns and two neighboring exons. Those tested include the insulin receptor (*INSR*), cardiac Troponin T type 2 (*TNNT2*), and *MBNL1*. Following co-transfection of the minigene and protein plasmid, the HeLa cells resumed normal cellular activity (transcription and translation) and expressed the protein construct from the plasmid containing the MBNL1 protein. The cells will also express the MBNL1-mediated reporter minigene mRNA isoforms. The mRNA isoform product ratio is regulated by the overexpressed protein. Due to the high efficiency of the constitutive promoter, MBNL1 will flood the system causing an overabundance of the splicing factor.

Once the cells have incubated for 18 hours, the RNA was harvested and purified, removing any DNA. The harvested RNA contains the alternatively spliced products as well as all other endogenous RNA. Due to RNA's instability, it must be converted to its stable counterpart, DNA. Reverse transcriptase circumvents this dilemma by converting the desired RNA, through reporter-specific primers, back into complementary DNA (cDNA). After the RNA is converted into cDNA, the cDNA is subject to polymerase chain reaction amplification or PCR. Through reporter-specific primers, PCR allows for selective amplification of the cDNA of interest. In other words, PCR is used to substantially increase the amount of DNA to an amount that is more favorable for measurement and management. The PCR samples of the alternatively

spliced products are then run on a native gel, using gel electrophoresis, and analyzed with imaging.

Gel electrophoresis distinguishes and separates the alternatively spliced products as a function of their size. Larger products, such as those that contain the alternatively spliced exon (inclusion products), take longer to migrate through the matrix of the acrylamide gel due to their large size. On the other hand, smaller products (exclusion products) are able to migrate more quickly through the gel. A difference in migration speed causes the larger products to appear higher up on the gel because the exon inclusion product contains more nucleotides. The opposite can be said for the exclusion product, which appears lower on the gel. Finally, splicing activity can be determined by measuring the intensity of the resulting bands. A comparison of percent exon inclusion, which is calculated by dividing the intensity of the inclusion band by the sum of the intensities of both the inclusion and exclusion bands, provides a means to measure splicing activity. Percent exon inclusion is used to measure the splicing activity by comparing the percent inclusion of the overexpressed protein samples to the control cells' splicing activity without expression of MBNL1.

The following minigenes were tested in the above-mentioned manner: *MBNL1*, *INSR*, *hTNNT2*, *TNNT2*, *VLDLR*, *Nfix*, and *ATP2A1*. The results of the overexpression assay reveal that MBNL1 (1-2, 1-2) splicing activity is similar to that of wild type, achieving an average splicing activity of 114.8% across all tested minigenes normalized to WT-MBNL1 (Hale, M., unpublished) (Table 1).

These results suggest that MBNL1 (1-2, 1-2) is about the same or slightly better than wild type at regulating splicing activity. However, this overexpression system

provides limited information because only a single data point can be collected. With only a single data point of comparison, it is difficult to definitively answer our original question: Can we engineer an MBNL1 protein with increased RNA binding affinity and splicing activity in comparison to wild type. It is possible that MBNL1 (1-2, 1-2) can regulate alternative splicing at lower concentrations of protein in comparison to wild type. The overexpression system does not let us confirm that possibility.

In order to more definitively answer our original question, it is necessary to compare splicing activities as a function of protein concentration. Unfortunately, few models are available that test splicing activity response in regards to the concentration of the splicing factor. **This research aims to develop a simple dosing system in which the splicing activity (% exon inclusion) can be determined over a wide range of protein (MBNL1) concentrations**. Such a model will theoretically measure the splicing activities as a function of the construct concentration and provide a more accurate comparison in splicing activities not possible with an overexpression assay.

Table 1. Percent splicing activity across all seven tested minigenes normalized to WT-MBNL1

Minigene	MBNL1 (1-2, 1-2)
MBNL1	103.8%
INSR	92.8%
hTNNT2	155.0%
TNNT2	144.0%
VLDLR	97.6%
Nfix	110.2%
ATP2A1	100.0%
Average	114.8%

METHODS

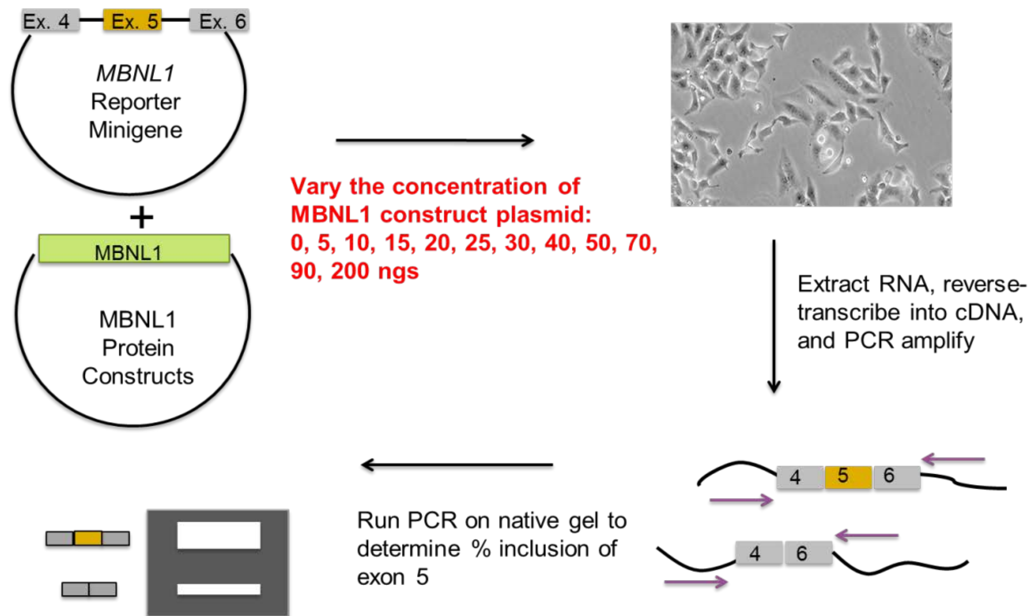


Figure 6. Workflow illustration of a transfection titration cell-based splicing assay. Circles represent plasmid vectors. Colored boxes (grey and orange) represent exons 4,5, and 6 of the *MBNL1* minigene. Purple arrows represent exon-specific PCR primers.

A modification of the cell-based splicing assay as described earlier will be utilized (workflow shown in Figure 6 and described in Purcell, *et al*). However, this model will address some of the pitfalls associated with an overexpression assay. Instead of transfecting the full concentration of protein plasmid (200 nanograms), varying amounts of protein plasmid (ranging from 0 to 200 nanograms) will be transfected, creating a titration-like effect. A titration of protein plasmid ranging from a low concentration to the full concentration will enable us to determine the splicing activity at varying concentrations of protein. The sensitivity of a splicing event to each of the protein constructs' concentration will help answer our original question and offer a simple alternative to other dosing methods. In this study, we investigate the constructs'

splicing efficacy in the auto-regulation of exon 5 exclusion of its own pre-mRNA transcript.

Cell culture and transfection

HeLa cells are routinely cultured as a monolayer in Dulbecco's Modified Eagle Medium (DMEM high glucose with GlutaMAX™, Life Technologies) supplemented with 10% fetal bovine serum (FBS) at 37°C under 5% CO_2 to approximately 85-90% confluence. Prior to transfection, the cells were plated in twelve-well plates at a density of $8x10^4$ cells/well. Plated cells are grown for approximately 36 hrs in DMEM, after which they are washed with 1X PBS (phosphate buffered saline) and the DMEM is replaced with Opti-MEM® reduced serum media (Life Technologies) and transfected with 400 ng total of plasmid DNA using Lipofectamine 2000® (Life Technologies) as described by the manufacturer's protocol. In each of the samples, cells are transfected with 200 ng of *MBNL* reporter minigene plasmid in addition to 200 ngs of a combination of construct protein plasmid and empty pCI vector plasmid. Control samples (mock) are transfected with 200 ngs of empty pCI vector. The remaining samples are transfected with 5, 10, 15, 20, 25, 30, 40, 50, 70, 90, and 200 ngs of construct protein plasmid to create a titration range. Empty pCI vector is transfected in these samples to keep the total concentration of DNA transfected constant at 400 ng (5 ng of protein plasmid and 195 ng of empty plasmid for example). Cells were harvested with TrypLE (Invitrogen) 24 h after transfection and pelleted via centrifugation.

Western blot analysis

HeLa cell pellets were lysed in radioimmunoprecipitation assay (RIPA) buffer (50mMTris, pH 7.4, 150mMNaCl,1% NP-40, 0.25% sodium deoxycholate, 0.1% SDS, 1 mM phenylmethylsulfonyl fluoride [PMSF]) supplemented with 1X protease inhibitor cocktail (Sigma*FAST*; Sigma) prior to resolution on 10% SDS-PAGE gels. MBNL1 construct proteins were probed with 1:1,000 primary antibody (anti mouse), followed by 1:1,000 goat anti-rabbit secondary. The GAPDH loading control was probed using GAPDH primary (1:1,000) antibody (anti rabbit) followed by 1:1,000 goat anti-mouse secondary. Fluorescence was visualized and quantified using a LI-COR imaging system (LI-COR Biosciences).

Cell-based splicing assay

RNA was isolated from HeLa cell pellets using an RNeasy kit (Qiagen). Isolated RNA (500 ng) was incubated with 1 unit of RQI DNase (Promega) in a 10 µl reaction mixture for 1 h at 37 °C. After DNase treatment, RNA (100 ng) was reverse transcribed (RT) in a 10 µl reaction volume with Superscript II (Invitrogen), according to the manufacturer's protocols, using reporter-specific primers for 50 min at 42 °C. The resulting cDNAs (2-µl volumes) are then amplified *via* polymerase chain reaction (PCR) in a 20-µl reaction mixture using flanking exon-specific primers[16]. The number of amplification cycles was determined to be within the linear range for the primers used. The resulting PCR products were resolved by gel electrophoresis on 6% native polyacrylamide (19:1) gels run at 300 V for 90 min. Alternatively spliced products were

visualized and quantified using SYBR green I nucleic acid stain (Invitrogen) in combination with an AlphaImager HP system (Alpha Innotech). Bands representing alternatively spliced products are quantified using Alpha Imager HP software (Alpha Innotech), and percent exon inclusion is calculated as the ratio of the intensity of the inclusion band to the sum of the intensities of inclusion and exclusion bands.

RESULTS

Western blot analysis confirms successful dosing

Western blot analysis was performed to confirm the effectiveness of the dosing experiment. In this experiment, HeLa cells were co-transfected with a titration of the construct protein plasmid at 0, 5, 10, 15, 20, 25, 30, 40, 50, 70, 90 and 200 nanograms

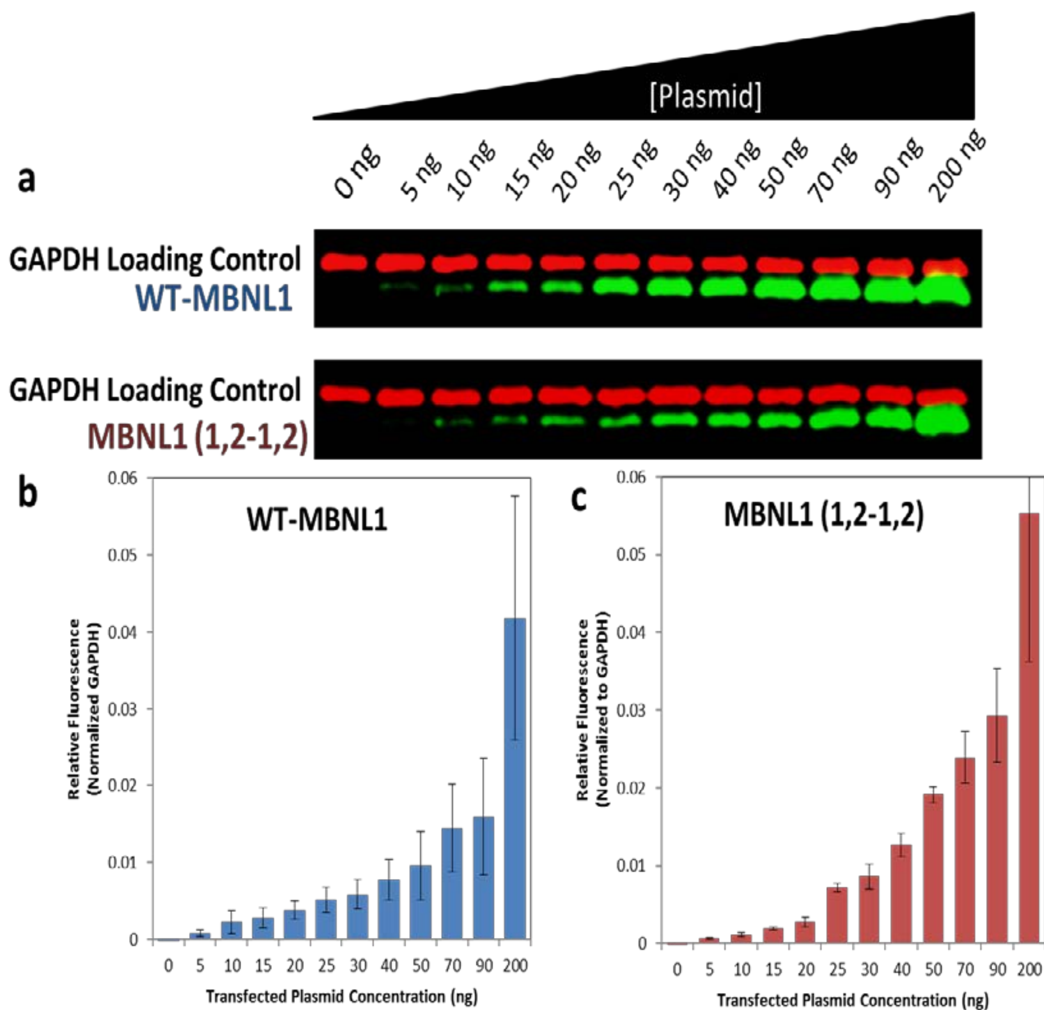


Figure 7. Western blot analysis. (a) Qualitative confirmation of dosing system for both WT-MBNL1 and MBNL1 (1-2, 1-2). (b) Quantification of WT-MBNL1 protein concentration calculated as the relative fluorescence normalized to GAPDH. (c) Quantification of MBNL1 (1-2, 1-2) protein concentration. Values are reported as an average of three separate western blots.

and a constant concentration of *MBNL1* reporter minigene (200 ng). A qualitative assessment of the western blot reveals a direct correlation between the concentration of the construct protein plasmid transfected and the concentration of the construct protein expressed within the cells (Figure 7). The concentration of protein expressed *in vivo* increases as the amount of the plasmid transfected increases, thus, confirming a successful dosing system.

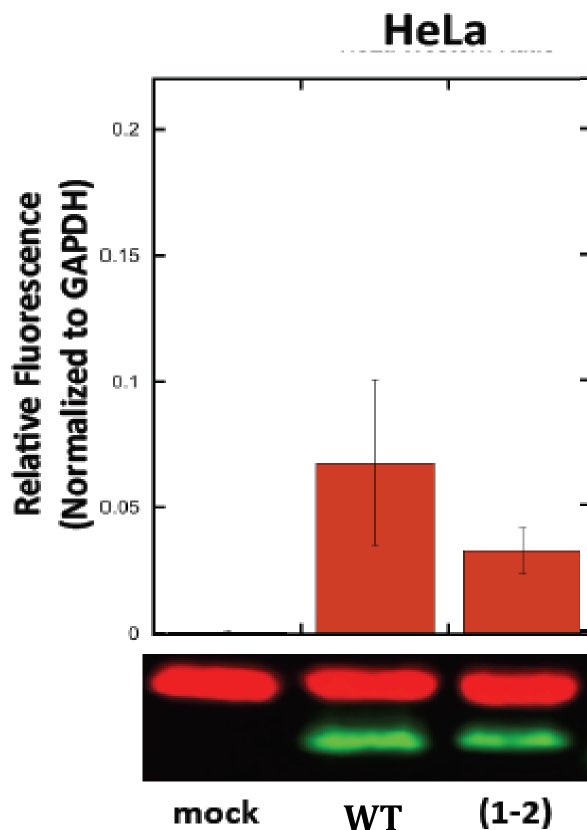


Figure 8. Western blot quantification of an overexpression of the constructs in HeLa relative to GAPDH.

Regardless of the positive qualitative results, the western blot does not fully encompass our goals. The original goal of this dosing model was to compare splicing activities between two splicing factors, one wild type and one engineered. In order to compare the activities between the two splicing factors, it is necessary to determine a relationship between the transfected plasmid amount and the concentration of construct protein expressed. However, we cannot directly cross compare protein concentrations as a function of the amount of plasmid transfected because of differences in protein expression of the plasmid system. This issue becomes apparent in the western blot

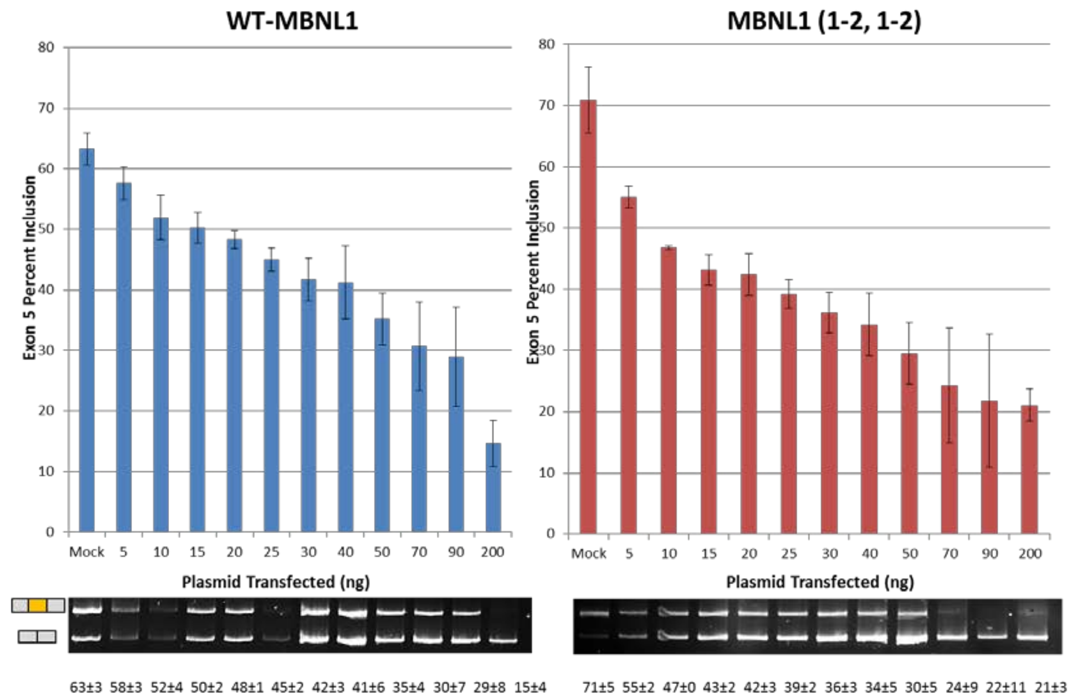
quantification of protein levels in the original overexpression assays performed. Here western blot analysis revealed that the concentration of MBNL1 (1-2, 1-2) is significantly lower (about half) than that of WT-MBNL1 in an overexpression system when the same amount of plasmid was transfected (Hale, M., unpublished) (Figure 8). Therefore, it is necessary to accurately determine the protein concentration at each "dose" for each construct.

Quantitative analysis reveals that the range of protein plasmid concentration (0-200 ng) creates an effective range of protein expression levels. According to the data, a near 100-fold range in construct protein concentration was achieved from the smallest transfected amount (5 ng) to the largest (200 ng) in each of the constructs (Figure 7). Unfortunately, the accuracy and reproducibility of the quantitative western blots pose a concern for a quantitative analysis. Performing multiple western blots under consistent conditions and with the same samples results in significantly different concentration values (represented by normalized fluorescence). In some occasions, the same sample (dose) subject to the same protocol produced up to a two-fold difference in construct protein concentrations expressed (data not shown). Therefore, technical reproducibility issues were an obstacle in western blot quantification. Due to the technical issues with the system, it is difficult to cross-compare blots. These inconsistencies within the same sample can likely be attributed to limitations inherent to quantitative western blot analysis. As it stands, the technology does not yet exist for quantitative purposes in westerns, although western blots serve qualitative purposes quite well.

Both WT-MBNL1 and MBNL1 (1-2, 1-2) exhibit a splicing sensitivity response

In order to determine the splicing activities of each construct at varying concentrations, an RT-PCR splicing assay was performed. Once again, HeLa cells were transfected with a construct protein plasmid titration of 0 (mock), 5, 10, 15, 20, 25, 30, 40, 50, 70, 90, and 200 nanograms, and were accompanied by the *MBNL* reporter minigene. Twelve data points of exon 5 percent exclusion, (splicing activity) corresponding to each dose, were generated to highlight the splicing response to the individual constructs. WT-MBNL1 was performed in at least a triplicate (n=4 for 0-30 ngs, n=3 for 40-200 ngs) and MBNL1 (1-2, 1-2) was performed in duplicate. Some of the bands in the lanes failed to appear. Reprocessing the RNA and re-running it on a gel allowed us to recover the “missing bands” (data not shown). Human error or reverse transcriptase complications may account for the missing bands in the original gels.

Dosing Responses



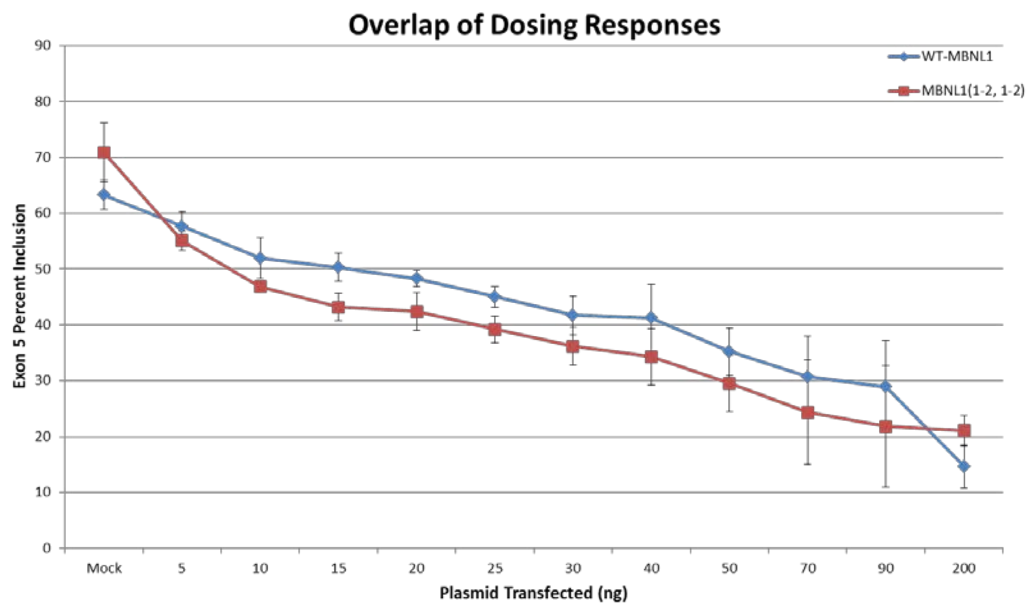


Figure 9. Splicing assay analysis. (a) Dosing response to both WT-MBNLl and MBNL1 (1-2, 1-2). The numbers below the splicing gels are actual exon 5 percent inclusion values. (b) An overlapping of both constructs' splicing response curves.

According to the splicing results, a direct correlation between splicing activity as measured by exon 5 percent inclusion and protein levels as detected by western blots was observed. As the amount of plasmid transfected increases, the percent of exon 5 inclusion decreases – corresponding to an increase in splicing activity (Figure 9). In response to overexpression (the maximum amount of protein plasmid, 200 ng) and in the absence of protein plasmid (mock), the exon 5 exclusion event of *MBNL* shows similar activity between WT-MBNL1 and MBNL1 (1-2, 1-2). For both constructs, the splicing activity appears to saturate near 16-20 percent inclusion, 14.63% ±3.84 for WT-MBNL1 and 21.09% ±2.70 for MBNL1 (1-2, 1-2).The endogenous (mock) splicing activity sits near 65 percent inclusion, 63.30% ±2.63 for WT-MBNL1 and 70.90% ±5.35 for MBNL1 (1-2, 1-2). These data coincide with the original overexpression assays (Hale, M., unpublished). Additionally, this titration dosing system offers ten extra data points, in which one can compare the splicing activity response sensitivity between both constructs. For each of the constructs, the percent inclusion of exon 5 is plotted against the amount of the plasmid transfected in an effort to compare efficacy of the proteins. Figure 9 depicts the degree of sensitivity by the slope of the response. According to Figure 9, splicing activity is at least the same or greater for MBNL1 (1-2, 1-2) at the same amount of transfected protein plasmid in comparison to WT-MBNL1.

A comparison of EC_{50} – effective amount of protein plasmid at which half of splicing activity (halfway point of percent exon inclusion between percent exon inclusion at 0 ng and 200 ng) is achieved – values between the two constructs suggests that MBNL1 (1-2, 1-2) regulates the exclusion of exon 5 much more effectively than WT-MBNL1. MBNL1 (1-2, 1-2) has an EC_{50} value of approximately 10 to 15

nanograms of plasmid versus an EC_{50} value of approximately 40 to 50 nanograms for wild type. However, as mentioned previously, a direct comparison of activity between the two constructs that only considers the transfected plasmid amount and neglects the actual protein concentration is inaccurate. Across constructs, this comparison fails to adequately address the discrepancies between plasmid amount and actual protein expression levels. Without a direct correlation between the two concentrations (plasmid and protein), the activities of the constructs cannot be accurately compared as a function of protein concentration like we hoped to achieve.

DISCUSSION

Western blot analysis and splicing assay results confirm successful dosing simulation

The main goal of this research was to develop a simple model in which a splicing activity response can be measured across a wide range of protein concentrations. Utilizing a titration of protein plasmid produced a measurable splicing activity response with two different proteins, WT-MBNL1 and MBNL1 (1-2, 1-2). This achievement is observable in both the western blot and the splicing assay results. The western blot analysis reveals that increasing the amount of transfected protein plasmid translates to an increase in protein expression levels. With the range of plasmid concentration transfected (0-200 ng), we were able to generate a 100-fold increase in protein expression (5 ng to 200 ng), which offers a generous range to analyze the sensitivity in splicing activity response. This titration created ten additional data points that contributed to a more thorough analysis in comparison to an overexpression assay. For both proteins studied, an increase in transfected protein plasmid correlates to an increase in splicing activity (measured by a decrease in exon 5 % inclusion of the *MBNL1* transcript).

Applications and advantages of this dosage simulation model (transfection titration)

The model outlined in this experiment expands upon an overexpression assay. Limitations of an overexpression assay, which only generate a single data point, are addressed through multiple transfections of varying protein plasmid amounts. Of note,

the concentrations selected for this experiment were optimized for a single splicing event, exon 5 exclusion of *MBNL1*. These same concentrations (0, 5, 10, 15, 20, 25, 30, 40, 50, 70, 90, and 200 nanograms) are not fixed and can easily be manipulated to optimize a range specialized to any splicing event. Each of these doses denote a reproducible and specific splicing activity (exon 5 percent inclusion), which is illustrated by small standard deviation values – a maximum standard deviation of 10.88, which would likely reduce had the experiment been performed in triplicate for MBNL1 (1-2, 1-2) (as opposed to only two independent splicing experiments).

This model also leaves room for additional data points. Although our experiments were performed in twelve-well plates, generating twelve data points (including endogenous splicing activity and overexpression splicing activity), a slight change in protocol may accommodate plates with a greater number of wells thus producing more than twelve data points. The greater the number of data points analyzed, the better the representation of splicing activity response. Therefore, this model imitates an inducible gene expression system that can be tightly controlled. However, the model outlined in this experiment is not the only inducible gene expression system that exists.

Recently, an inducible gene expression system has been developed that takes advantage of the Tet-On advanced transactivator and the P_{Tight} inducible reporter. In the presence of doxycycline (dox), Tet-On Advanced binds to the tetracycline response element (TRE_{Mod}) in P_{Tight}, and produces high-level transcription of the downstream gene of interest[17]. As a result, this system allows for a method to control gene expression as a function of dox concentration. Some of the benefits of this system

include extremely tight regulation, high specificity, high inducibility, and fast response times. However, one pitfall of this system is the time it takes to generate a stable, inducible cell line that is required for a Tet-On Advanced induction system. A functioning inducible cell line can take upwards of a few months to create[18], whereas the system outlined in this experiment can be completed from start to finish in just short of a week (around 5-6 days). Despite the dox dosing system boasting greater precision and tighter control, both methods, although they take different approaches, achieve similar results. This transfection titration method however, can produce results in a fraction of the time. Therefore, the model outlined in this experiment is useful under time constraints and allows for quick splicing sensitivity analysis. This model can also be used in conjunction with the dox system to provide further support of results.

Challenges faced in developing the transfection titration model

Unfortunately, the experiments did not definitively answer one of our questions: Can we engineer an MBNL1 protein with increased RNA-binding affinity and splicing activity in comparison to wild type? Firstly, an overexpression assay fails to answer this question because it is possible that MBNL1 (1-2, 1-2) regulates alternate splicing at lower concentrations than WT-MBNL1. In order to answer this question, we need to compare the splicing activities of the constructs at varying concentrations under the assumption that a concentration of transfected plasmid corresponded to the same concentration of active protein. What we failed to anticipate was that transfecting the same amount of protein plasmids resulted in vastly different expression levels between the constructs, as shown in previous western blot quantifications of the overexpression assays.

Ideally, we could eliminate this inconsistency by determining a relationship between the amount of plasmid transfected and the concentration of protein expressed for each construct. Instead of comparing the splicing activities as a function of plasmid transfected, we could compare the splicing activities as a function of actual protein concentration with the previously mentioned relationship. Determining this relationship would provide us with an accurate comparison of splicing activities between the constructs. Western blot quantification is necessary to determine such a relationship. However, the effort to determine the relationship between plasmid concentration and protein concentration was thwarted by accuracy limitations inherent in western blot quantification. A lack of reproducibility within the same samples poses an obstacle in accurately defining a relationship because in some cases the relative fluorescence between technical replicates can differ by up to a factor of two. Until western blot quantification accuracy and reproducibility improve, it is not possible to accurately determine protein concentrations for quantitative analysis. This model successfully depicts the sensitivity of the splicing event to a specific construct as a function of plasmid concentration. However, without an accurate correlation between plasmid concentration and active protein concentration, this model ultimately fails to cross-compare construct splicing activities.

Another smaller issue that this experiment encountered was “missing bands.” In some experiments, either very faint or absent bands were observed in some of the lanes. This issue suggests there was an insufficient quantity or concentration of DNA loaded on the gel. A lack of DNA can likely be attributed to a failure of the reverse transcriptase enzyme (without conversion of the RNA into cDNA, there is no DNA to

PCR amplify thus resulting in no bands) or perhaps human mixing error. The missing bands could be recovered by subjecting the same RNA to another round of processing (DNAse then RT-PCR). Recovered bands produced percent inclusion values in agreement with other independent splicing experiments.

FUTURE DIRECTIONS AND REFINEMENT

In short, this novel model is simple to use, versatile, hosts a wide range of applications, and simulates an inducible gene expression system. The model outlined in this experiment can also be performed relatively quickly, especially in comparison to other inducible gene expression systems such as the dox system. Importantly, this model expands on an overexpression system and provides a more accurate representation of splicing activity sensitivity by providing more than one data point of comparison. As previously mentioned, this model falls short in comparing activities between proteins due to limitations in western blot quantification.

A refinement of this model could theoretically erase the reproducibility limitations of western blots. For example, the overall trend of increasing plasmid concentration to increasing protein expression levels held within each western blot, but the magnitudes of the values were difficult to reproduce. For this reason, I propose performing the same experiment but instead first harvesting the RNA (for the splicing assay) and protein (for western blot analysis) from the same HeLa cells in a single experiment for both WT-MBNL1 and MBNL1 (1-2, 1-2). Then, unlike before, I plan to perform a western blot analysis of both constructs at the same time. This ensures that both of the constructs are subject to the same sources of error. Also dissimilar to previous experiments, I will perform a western blot analysis and splicing assay from the same cells (previously performed from separate cell passages), which will further eliminate sources of inconsistency.

Theoretically, quantifying both constructs in the same blot should eliminate the inconsistencies inherent in western blot quantification. A relationship between plasmid

concentration and active protein concentration could then be more reliably determined. Finally, with the relationship determined, we can cross compare the splicing activities of our proteins of interest.

REFERENCES

1. Bonnal S, Vigevani L, Valcárcel J. The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov*. 2012 Nov;11(11):847-59. doi: 10.1038/nrd3823.

2. Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*. 2003;72:291-336.

3. Pascual M, Vicente M, Monferrer L, Artero R. The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation*. 2006 Mar;74(2-3):65-80.

4. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008 Nov 27;456(7221):470-6.

5. López-Bigas N, Audit B, Ouzounis C, Parra G, Guigó R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett*. 2005 Mar 28;579(9):1900-3.

6. Bird, T.D. (2013). Myotonic dystrophy type 1. GeneReviews. http://www.ncbi.nlm.nih.gov/books/NBK1165/

7. Goes, E.S., Purcell, J., Voelker, R.B., Gates, D.P., and Berglund, J.A. (2010). MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res* 38(7):2467-84.

8. Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. 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 Dec 15;13(24):3079-88.

9. Ho TH, Charlet-B N, Poulos MG, Singh G, Swanson MS, Cooper TA. Muscleblind proteins regulate alternative splicing. *EMBO J*. 2004 Aug 4;23(15):3103-12.

10. Sen S, Talukdar I, Liu Y, Tam J, Reddy S, Webster NJ. Muscleblind-like 1 (Mbnl1) promotes insulin receptor exon 11 inclusion via binding to a downstream evolutionarily conserved intronic enhancer. *J Biol Chem*. 2010 Aug 13;285(33):25426-37.

11. Kino Y, Washizu C, Oma Y, Onishi H, Nezu Y, Sasagawa N, Nukina N, Ishiura S. MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res*. 2009 Oct;37(19):6477-90.

12. Gates DP, Coonrod LA, Berglund JA. Autoregulated splicing of muscleblind-like 1 (MBNL1) Pre-mRNA. *J Biol Chem*. 2011 Sep 30;286(39):34224-33.

13. Warf MB, Diegel JV, von Hippel PH, Berglund JA. The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad Sci USA*. 2009 Jun 9;106(23):9203-8.

14. Klug A, Schwabe JW. Protein motifs 5. Zinc fingers. *FASEB J*. 1995 May;9(8):597-604.

15. Teplova M, Patel DJ. Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat Struct Mol Biol*. 2008 Dec; 15(12):1343-51.

16. Purcell J, Oddo JC, Wang ET, Berglund JA. Combinatorial mutagenesis of MBNL1 zinc fingers elucidates distinct classes of regulatory events. *Mol Cell Biol*. 2012 Oct;32(20):4155-67.

17. Zhou X, Vink M, Klaver B, Berkhout B, Das AT. Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther*. 2006 Oct;13(19):1382-90.

18. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994). Current Protocols in Molecular Biology (New York: *Greene Publishing Associates and Wiley-Interscience*).