THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 292, NO. 10, pp. 4350–4357, March 10, 2017

Pseudouridine Modification Inhibits Muscleblind-like 1 (MBNL1) Binding to CCUG Repeats and Minimally Structured RNA through Reduced RNA Flexibility*

Received for publication, December 2, 2016, and in revised form, January 25, 2017 Published, JBC Papers in Press, January 27, 2017, DOI 10.1074/jbc.M116.770768

Elaine deLorimier[‡1], Melissa N. Hinman[‡1], Jeremy Copperman[‡], Kausiki Datta[§], Marina Guenza[‡], and J. Andrew Berglund[‡§2]

From the [‡]Institute of Molecular Biology, Department of Chemistry and Biochemistry, University of Oregon, Eugene, Oregon 97403 and the [§]Center for NeuroGenetics, Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, Florida 32610-3010

Edited by Ronald C. Wek

Myotonic dystrophy type 2 is a genetic neuromuscular disease caused by the expression of expanded CCUG repeat RNAs from the non-coding region of the CCHC-type zinc finger nucleic acid-binding protein (*CNBP*) gene. These CCUG repeats bind and sequester a family of RNA-binding proteins known as Muscleblind-like 1, 2, and 3 (MBNL1, MBNL2, and MBNL3), and sequestration plays a significant role in pathogenicity. MBNL proteins are alternative splicing regulators that bind to the consensus RNA sequence YGCY (Y = pyrimidine). This consensus sequence is found in the toxic RNAs (CCUG repeats) and in cellular RNA substrates that MBNL proteins have been shown to bind. Replacing the uridine in CCUG repeats with pseudouridine (Ψ) resulted in a modest reduction of MBNL1 binding. Interestingly, Ψ modification of a minimally structured RNA containing YGCY motifs resulted in more robust inhibition of MBNL1 binding. The different levels of inhibition between CCUG repeat and minimally structured RNA binding appear to be due to the ability to modify both pyrimidines in the YGCY motif, which is not possible in the CCUG repeats. Molecular dynamic studies of unmodified and pseudouridylated minimally structured RNAs suggest that reducing the flexibility of the minimally structured RNA leads to reduced binding by MBNL1.

Myotonic dystrophy type 1 (DM1)[3] is a genetic neuromuscular disease caused by expression of expanded CUG repeats in the 3′ UTR of the dystrophia myotonica protein kinase (*DMPK*) gene. Similar to DM1, myotonic dystrophy type 2 (DM2) is caused by expression of expanded CCUG repeats in an intron of the CCHC-type zinc finger nucleic acid-binding protein (*CNBP*) gene. DM1 and DM2 occur when the CUG/CCUG repeats are expanded beyond 100 repeats, and patients can have up to thousands of CUG/CCUG repeats (1, 2). A primary component of the currently accepted DM1 and DM2 disease mechanism is that expanded CUG/CCUG repeats sequester RNA-binding proteins (primarily the Muscleblind-like family), which prevents these proteins from performing their functions in cells (3, 4).

The members of the Muscleblind-like family of proteins (MBNL1, MBNL2, and MBNL3) bind RNA and regulate several RNA processing pathways, including alternative splicing, pre-miRNA biogenesis, mRNA localization, alternative polyadenylation, and circular RNA generation (5–9). MBNL proteins bind to the consensus YGCY RNA sequence (6, 10). CUG and CCUG repeats are composed of YGCY motifs creating hundreds or thousands of perfect MBNL-binding sites resulting in large numbers of MBNL proteins binding to the repeats and forming nuclear foci (11). When MBNL proteins are sequestered, they are unable to regulate RNA processing events, and consequently, many DM1 and DM2 symptoms are caused by misregulated alternative splicing and potentially the loss of other MBNL activities (12). It is therefore important to understand how MBNL proteins bind to their toxic and cellular RNA substrates to develop mechanisms to alleviate MBNL sequestration in DM1 and DM2.

MBNL proteins have two sets of zinc finger domains that are proposed to bind to RNA on opposing faces of the domain (13, 14). Teplova and Patel (15) published a crystal structure of one of the zinc finger domains of MBNL1 in complex with a YGCY-containing RNA, showing that MBNL1 interacts with the Watson-Crick face of the GC dinucleotide. This structure indicates that at least this portion of the RNA must be single-stranded to interact with MBNL1, as the Watson-Crick face interacts with the opposing strand in double-stranded RNA.

Pseudouridine (Ψ), the most abundant RNA modification, is often referred to as the fifth RNA base. Ψ is found in tRNA, rRNA, and many spliceosomal RNAs, and it functions to stabilize the structure of these RNAs (16). Ψ is a 5-ribosyl isomer of uridine, with a C–C bond connecting the ribose to the base instead of a C–N and an extra H-bonding donor on the N1 (Fig.

* This work was supported by Muscular Dystrophy Association Grant 294447 (to J. A. B.), National Institutes of Health Grant AR059833 (to J. A. B.), Myotonic Dystrophy/Wyck Foundation Grant Wyck-FF-2014-0013 (to M. N.H .), and American Heart Association Pre-doctoral Fellowship 13PRE17350007 (to E. D.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Center for NeuroGenetics, Dept. of Biochemistry and Molecular Biology, University of Florida, 2033 Mowry Rd., Gainesville, FL 32610-3010. Tel.: 352-273-7888; Fax: 352-273-8284; E-mail: aberglund@ufl.edu.

[3] The abbreviations used are: DM1/2, myotonic dystrophy type 1/2; MBNL1/2/3, muscleblind-like 1/2/3; Ψ, pseudouridine; miRNA, microRNA; MD, molecular dynamics; Bicine, *N*,*N*-bis(2-hydroxyethyl)glycine.

1*A*). Nuclear magnetic resonance (NMR), X-ray crystallography, and molecular dynamics (MD) simulations have shown that Ψ induces rigidity in both single- and double-stranded RNA via base-stacking and hydrogen-bonding interactions (16–18). There is evidence that in some cases base-stacking interactions are the primary stabilizing force in RNA containing Ψ (19).

In our previous work, replacing uridines with Ψ in the CUG repeats induced structural stabilization, prevented MBNL1 binding, and rescued mis-splicing. X-ray crystallography data suggested that Ψ stabilized the CUG repeats through water bridging via hydrogen bonding of the N1-H of Ψ with the uridine on the opposing strand (20). In this work, we have extended our studies to determine that Ψ can stabilize CCUG repeat RNA and minimally structured RNA containing YGCY motifs and that this modification inhibits MBNL1 binding. MD simulations indicate that Ψ increases base-stacking interactions in minimally structured YGCY RNAs.

Results

Replacing Uridine with Ψ in CCUG Repeat RNA Increases Structural Stability—Previous studies showed that Ψ increased the thermal stability of CUG repeat RNA in a helical conformation (20). Here, we determined whether Ψ had a similar effect on stabilization of CCUG repeats. Although both types of repeats contain the MBNL1-binding motif YGCY, the additional cytosine in CCUG repeats renders them different from CUG repeats. Uridine is only upstream of the GC dinucleotide, so Ψ substitutions can only change the 5′ side of the YGCY-binding site unlike the CUG repeats, where uridine is on both sides. We designed a model CCUG repeat RNA containing two sets of three CCUG repeats separated by a UUCG linker (Fig. 1*C*, *RNA 1*). This linker has been shown to favor stem-loop formation (21) and was included to facilitate structural studies. The most energetically favorable secondary structure for this RNA, as predicted by several different software programs, is one in which C-U and U-C mismatched base pairs are adjacent to G-C base pairs (Fig. 1*B*). However, NMR studies suggested that RNA 1 did not adopt a single defined structure. Unmodified CCUG repeats exhibited a melting temperature (T_m) of 47 ± 1 °C (Fig. 1*D*, *RNA 1*), whereas the T_m for CCUG repeats with Ψ substitutions was 56 ± 2 °C (Fig. 1*D*, *RNA 2*). These data showed that Ψ increases the thermal stability of CCUG repeats, similar to what was found with CUG repeats. Melting temperatures were similar across a broad range of RNA concentrations, suggesting that intramolecular rather than intermolecular interactions are the dominant force determining thermal stability (Fig. 1*E* and Table 1). The same trends were observed whether T_m was measured from a curve of increasing or decreasing temperatures, although in our spectrophotometer we consistently obtained a slightly higher T_m when using a curve of increasing temperature (Table 1).

MBNL1 Has a Slightly Reduced Affinity for CCUG Repeats Modified with Ψ—Gel shift assays were used to determine whether the Ψ stabilization observed in the CCUG repeat RNAs inhibited MBNL1 binding. The dissociation constant (K_d) for MBNL1 binding to unmodified CCUG repeats is 118 ± 15 nM (Fig. 2, *A* and *B*, *RNA 1*), and modification of uridines to

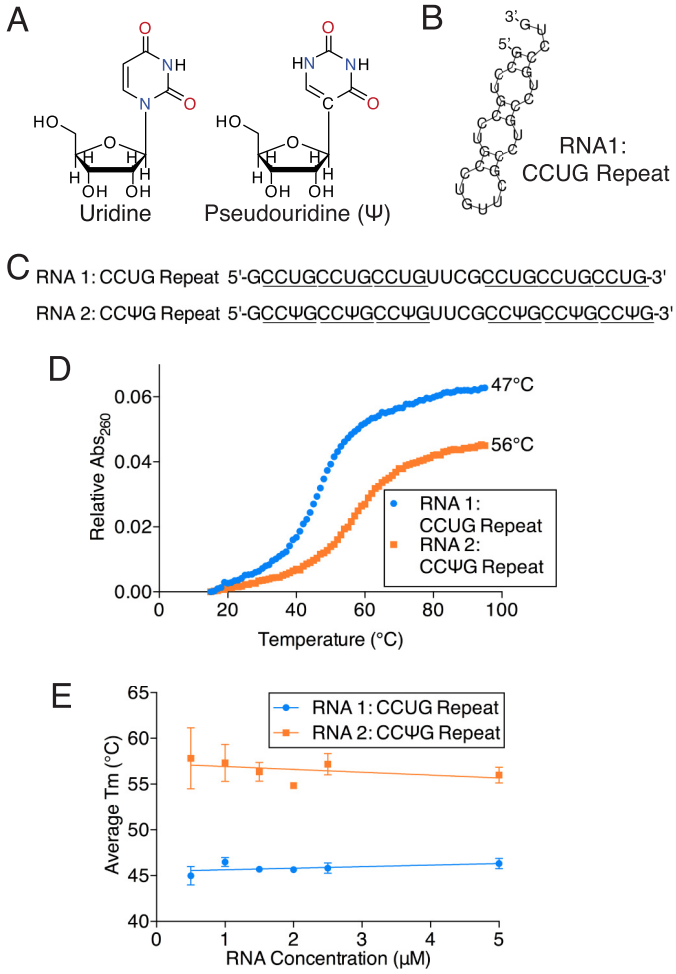

FIGURE 1. **Ψ increases thermal stability of CCUG repeat RNA.** *A*, representative images of uridine and Ψ. *B*, secondary structure of CCUG repeat RNA 1 as predicted by RNAfold. *C*, sequence of the CCUG and CCΨG repeat RNAs (*RNAs 1* and *2*) used in the thermal melt and gel-shift assays. Repeats are *underlined*. *D*, representative thermal melt curves of CCUG and CCΨG repeat RNAs. Absorbance at 260 nm was measured over the temperature range of 15–95 °C. Melting temperatures are shown to the *right* of each curve. Curves were generated using 2 μM RNA and increasing temperatures. Data are averages of three independent experiments. *E*, melting temperatures of CCUG repeat RNAs across a range of RNA concentrations. Each point represents the mean T_m for increasing (15–95 °C) and decreasing (95–15 °C) temperatures for each experiment, averaged across three experimental replicates. *Error bars* represent standard deviation. Data points for each RNA were fit to a line through linear regression analysis. Neither line had a significantly non-zero slope (RNA 1, $p = 0.14$, and RNA 2, $p = 0.30$).

Ψ in CCΨG repeats resulted in a modestly increased K_d of 168 ± 19 nM (Fig. 2, *A* and *B*, *RNA 2*). These results show that Ψ has a modest inhibitory effect on MBNL1 binding to the YGCY-binding sites in the CCUG repeats. This effect is significantly less than that observed with Ψ modification of CUG repeats (20).

MBNL1 Has Reduced Affinity for a Minimally Structured RNA Modified with Ψ—We next asked whether Ψ modification would inhibit the ability of MBNL1 to bind to minimally structured RNA. We chose the model MBNL1 RNA substrate ($U_4(GC)U_{11}(GC)U_4$), which was previously demonstrated to have little RNA structure (Fig. 3*A*, *RNA 3*) (14). We cannot rule out the possibility that RNA 3 forms a weak hairpin structure through non-canonical U-U base pairing (22–25). However, such a structure would be much less stable than

TABLE 1

CCUG repeat RNA melting temperatures

Melting temperatures are shown for modified and unmodified CCUG repeat RNAs across a range of RNA concentrations. Absorbance at 260 nm was measured over the temperature range of 15–95 °C and then from 95 to 15 °C. For each concentration, the T_m for increasing temperature (up) and decreasing temperature (down) is shown. Data are averages of three experimental replicates and error represents standard deviation.

	0.5 μM		1 μM		1.5 μM		2 μM		2.5 μM		5 μM	
RNA	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
RNA 1, CCUG repeat	45 ± 1	45 ± 3	47 ± 2	46 ± 2	47 ± 1	45 ± 1	47 ± 1	44 ± 1	46 ± 1	46 ± 2	48 ± 1	45 ± 2
RNA 2, CCΨG repeat	59 ± 4	57 ± 4	58 ± 3	57 ± 2	57 ± 1	55 ± 1	56 ± 2	54 ± 2	58 ± 1	57 ± 3	58 ± 2	54 ± 1

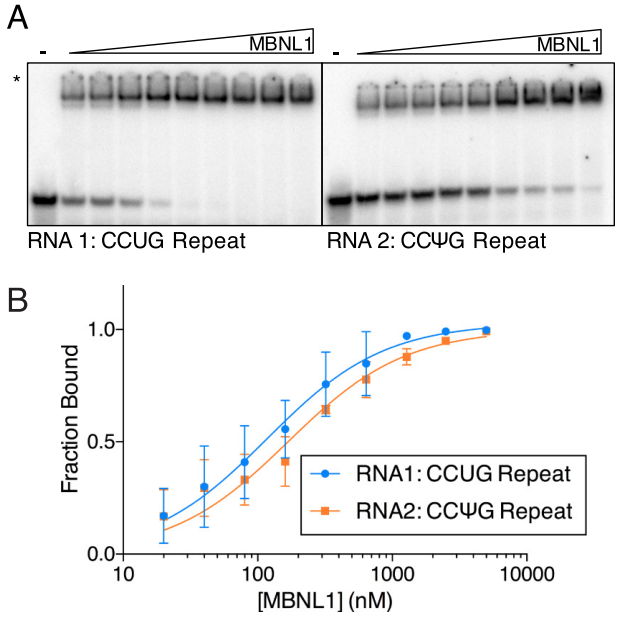

FIGURE 2. **MBNL1 has reduced affinity for pseudouridylated CCUG repeat RNA.** *A,* representative binding gels of MBNL1-CCUG repeat RNAs with and without Ψ modification. MBNL1 concentration increased in 2-fold steps from 20 nM to 5 μM. The K_d for RNA 1 is 118 ± 15 nM and the K_d for RNA 2 is 168 ± 19 nM. *Asterisk* indicates the location of the wells. *B,* binding curves for unmodified and Ψ modified CCUG repeat RNAs bound to MBNL1. They represent the mean of five independent experiments, and *error bars* represent standard error.

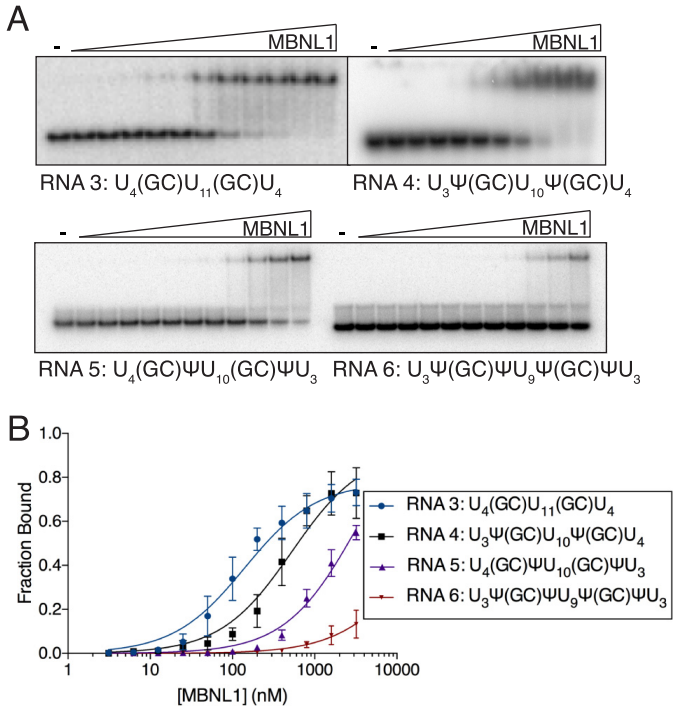

FIGURE 3. **MBNL1 has reduced affinity for minimally structured RNA containing pseudouridines.** *A,* binding gels of MBNL1-$U_4(GC)U_{11}(GC)U_4$ RNAs with and without Ψ modifications. MBNL1 concentration increased in 2-fold steps from 3.125 nM to 3.2 μM for RNAs 3, 5, and 6, and from 9.8 nM to 5 μM for RNA 4. K_d values were 0.14 ± 0.09 μM for RNA 3, 0.83 ± 0.15 μM for RNA 4, 3.2 ± 0.66 μM for RNA 5, and > 3.2 μM for RNA 6. *B,* binding curves for $U_4(GC)U_{11}(GC)U_4$ RNA and Ψ-modified RNAs incubated with MBNL1. They represent the mean of three independent experiments, and *error bars* represent standard error.

that formed through canonical base pairing in CCUG repeat RNA 1 (Fig. 1*B*). The Ψ modification was placed 5′ or 3′ of the GC dinucleotide and flanking both GC dinucleotides in the $U_4(GC)U_{11}(GC)U_4$ RNA (Fig. 3*A*, *RNAs 4–6*). The dissociation constant of MBNL1 binding to unmodified minimally structured RNA is 0.14 ± 0.09 μM (Fig. 3*A*, *RNA 3*). The addition of Ψ modifications 5′ to the GC dinucleotides increased the K_d to 0.83 ± 0.15 μM (Fig. 3*A*, *RNA 4*), whereas Ψ placement 3′ to the GC dinucleotides increased the K_d to 3.2 ± 0.66 μM (Fig. 3*A*, *RNA 5*). When four pseudouridines flank the two GC dinucleotides, the fraction of RNA bound does not reach 0.5 in these assays, so the dissociation constant is >3.2 μM (Fig. 3*A*, *RNA 6*). Binding curves used to calculate these values are shown in Fig. 3*B*. These results show that MBNL1 binding was generally inhibited by Ψ incorporation. The nucleotide 3′ to the GC dinucleotide had a more pronounced effect on MBNL1 binding than the nucleotide in the 5′ position, as the K_d for RNA 5 with the 3′ Ψ modifications is ~4-fold greater than for RNA 4 with 5′ modifications (Fig. 3).

Ψ Increases the Thermal Stability of Minimally Structured YGCY RNAs—Melting studies were performed to determine whether the Ψ modifications would increase the thermal stability of the MBNL1 model minimally structured substrate, RNA 3. Previous reports have shown that Ψ can induce base-stacking interactions in single-stranded RNAs, thereby predicting that increased base stacking might be observed with this RNA (17). RNA 3 melted at a temperature of 29 ± 1 °C, which is presumably due primarily to the unstacking of the bases in the minimally structured RNA, but could also reflect melting of a hairpin containing non-canonical base pairs (Fig. 4*A*). The RNA with the two pseudouridines 5′ of the GC dinucleotides melted at 39 ± 1 °C (Fig. 4*A*, *RNA 4*). The RNA containing pseudouridines 3′ to the dinucleotide melted at 37 ± 2 °C (Fig. 4*A*, *RNA 5*). The RNA containing Ψ substitutions flanking the GC dinucleotides melted at 45 ± 1 °C (Fig. 4*A*, *RNA 6*). These data indicate that Ψ stabilizes minimally structured RNAs, possibly by increasing base stacking. Consistent with this interpretation is that additional Ψ substitutions increased the melting temperature (Fig. 4*A*). An increase in thermal stability of these minimally structured RNAs is likely due to strand rigidifying via

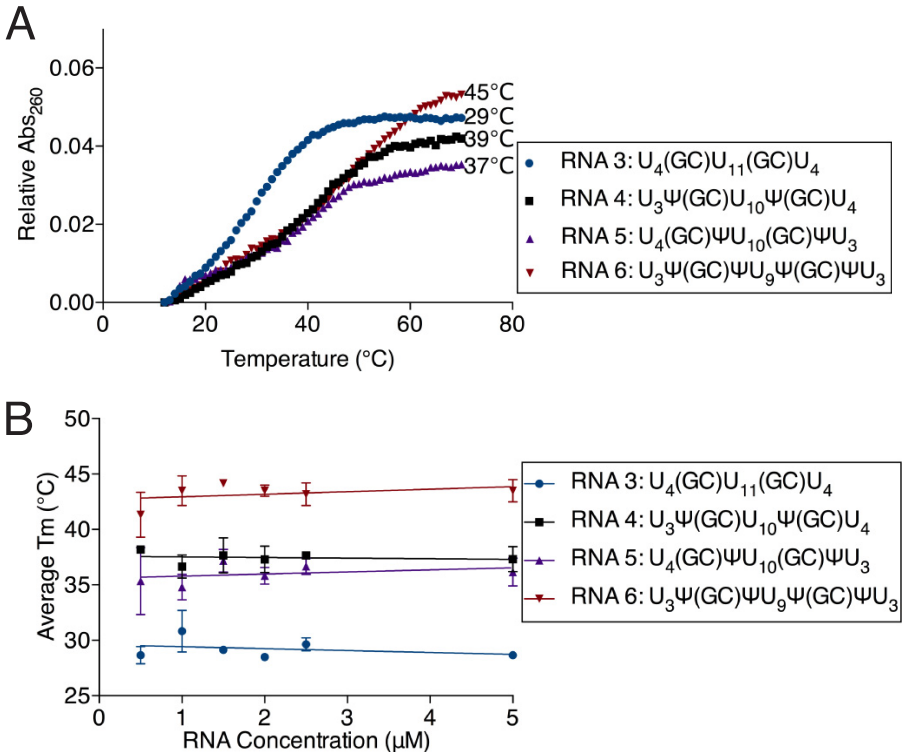

FIGURE 4. **Pseudouridine increases the thermal stability of minimally structured YGCY RNAs.** *A,* melting curves for $U_4(GC)U_{11}(GC)U_4$ RNA and RNA modified with Ψ as depicted in the legend. Absorbance at 260 nm was measured over the temperature range of 12–70 °C. Melting temperatures are shown to the *right* of each curve. Curves were generated using 2 μM RNA and increasing temperatures. *B,* melting temperatures of $U_4(GC)U_{11}(GC)U_4$ RNAs across a range of RNA concentrations. Each point represents the mean of the T_m for increasing (12–70 °C) and decreasing (70–12 °C) temperatures for each experiment, averaged across three experimental replicates. *Error bars* represent standard deviation. Data points for each RNA were fit to a line through linear regression analysis. None of the lines had significantly non-zero slopes (RNA 3 $p = 0.35$; RNA 4 $p = 0.74$; RNA 5 $p = 0.46$; and RNA 6 $p = 0.31$).

base-stacking interactions. Melting temperatures were similar across a broad range of RNA concentrations, suggesting that intramolecular interactions were the primary drivers behind changes in thermal stability (Fig. 4*B* and Table 2). Like what was observed with CCUG repeat RNAs, minimally structured RNA melting temperatures measured from a curve of increasing temperatures were consistently slightly higher than those from a curve of decreasing temperatures, likely due to an artifact of our measuring equipment (Table 2).

Molecular Dynamics Simulations Indicate That Ψ Increases Base Stacking in Minimally Structured YGCY RNAs—We conducted MD simulations to study whether Ψ induces increased base-stacking interactions in a fully atomistic explicit solvent model of minimally structured YGCY RNA. MD simulations allowed us to computationally investigate the dynamics of minimally structured $U_5(GC)U_5$ (RNA 7), $U_4Ψ(GC)U_5$ (RNA 8), $U_5(GC)ΨU_4$ (RNA 9), and $U_4Ψ(GC)ΨU_4$ (RNA 10). There were many differences in the conformational ensemble of the unmodified RNA *versus* the pseudouridylated RNAs. Overall, the pseudouridylated structures were much more compact and stable than the ensemble of unmodified RNA structures. To quantify base-stacking interactions, we calculated the relative free energy of the distance between the U/Ψ5-G6 and the G6-C7 bases throughout the simulation (Fig. 5, *A* and *B*). We define the distance between the bases to be the distance between the C2 atom on the G6 base and the C2 (C4) atom on the U5 (Ψ5) base.

U5-G6 bases had a global minimum in the stacked configuration but a very small barrier of 1.0 kcal/mol separating a range of open states of nearly equal free energy to the stacked state (Fig. 5*B*, *RNA 7*). Overall, the range of open conformations, defined as structures with a U/Ψ5-G base separation of greater than 7 Å, was about four times more probable than closed conformations (Fig. 5*B*, *RNA 7*). The widespread distribution of base separation distances in the $U_5(GC)U_5$ simulation indicates that the bases are dynamic, without a strong preference for base stacking. Furthermore, analysis of the ensemble of open structures of the $U_5(GC)U_5$ from the simulation show that G6 is commonly unstacked from both neighboring bases (Fig. 5, *A* and *B*, *RNA 7*). These open conformations can be easily docked into the MBNL1-binding site similar to the bound conformation observed in the crystal structure (Fig. 5*C*), supporting the model that effective MBNL1 binding requires the presence of flexible conformations with the G of the YGCY-binding site unstacked from both neighbors.

Although some pseudouridylated constructs displayed slightly increased flexibility at small separations, open configurations with a stacking distance to the guanine base of greater than 7 Å were highly disfavored in comparison with the $U_5(GC)U_5$ construct for both constructs containing Ψ upstream of G6 (Fig. 5*B*, *RNAs 8* and *10*). At the U-G base separation distance of 10.7 Å, as measured in the crystal structure of bound MBNL1 and RNA, the difference in free energy between the unmodified RNA and fully pseudouridylated RNA

TABLE 2

Minimally structured model RNA melting temperatures

Melting temperatures are shown for modified and unmodified $U_4(GC)U_{11}(GC)U_4$ RNAs across a range of RNA concentrations. Absorbance at 260 nm was measured over the temperature range of 12–70 °C and then from 70 to 12 °C. For each concentration, the T_m for increasing temperature (up) and decreasing temperature (down) is shown. Data are averages of three experimental replicates and error represents standard deviation.

	0.5 μM		1 μM		1.5 μM		2 μM		2.5 μM		5 μM	
RNA	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
RNA 3, $U_4(GC)U_{11}(GC)U_4$	28 ± 1	29 ± 2	32 ± 2	29 ± 2	31 ± 1	28 ± 1	29 ± 1	28 ± 1	31 ± 1	28 ± 1	30 ± 1	28 ± 1
RNA 4, $U_3(GC)U_{10}\psi(GC)U_4$	39 ± 1	37 ± 1	38 ± 1	35 ± 1	39 ± 3	36 ± 1	39 ± 1	36 ± 1	39 ± 1	36 ± 1	39 ± 2	35 ± 1
RNA 5, $U_4(GC)\psi U_{10}(GC)\psi U_3$	34 ± 3	37 ± 3	36 ± 3	33 ± 1	39 ± 2	36 ± 2	37 ± 2	35 ± 1	39 ± 1	35 ± 2	37 ± 1	35 ± 2
RNA 6, $U_3\psi(GC)\psi U_9\psi(GC)\psi U_3$	40 ± 3	42 ± 2	42 ± 1	45 ± 3	46 ± 1	43 ± 1	45 ± 1	42 ± 1	45 ± 1	42 ± 1	44 ± 2	43 ± 1

studied in the MD simulations was 1.9 kcal/mol (Fig. 5*B*, *RNAs 7* and *10*). This could be significant in a conformational selection model of binding. Pseudouridylation of only U5 resulted in increased base stacking with G6 but a modest decrease in base stacking between G6 and C7 (Fig. 5, *A* and *B*, *RNAs 7* and *8*). Conversely, pseudouridylation of the 3′ U8 alone led to increased base-stacking between G6 and C7 but increased U5-G6 base separation (Fig. 5, *A* and *B*, *RNAs 7* and *9*). Pseudouridylation of both U5 and U8 led to increased base stacking on both sides of G6 (Fig. 5, *A* and *B*, *RNAs 7* and *10*). These data suggest that Ψ increases base-stacking interactions in these minimally structured YGCY RNAs.

Discussion

Replacing the six uridines in CCUG repeats with pseudouridines (RNAs 1 and 2) increased the thermal stability of this RNA by 9 °C (Fig. 1*D*). Surprisingly, this stabilization only resulted in a modest ~1.5-fold inhibition of MBNL1 binding (Fig. 2*A*). In contrast, stabilization of CUG repeats with two or four Ψs resulted in a larger increase in thermal stability (13 or 19 °C for two or four Ψ modifications, respectively) (20) and displayed a more dramatic effect on MBNL1 binding (>20-fold loss of binding) compared with CCUG repeat binding. A possible explanation for the difference in MBNL1 binding observed between CUG and CCUG repeats with Ψ modification is that CCUG repeats are relatively less stable compared with CUG repeats (the observed T_m of $(CUG)_4$ was ~59 °C (20) and that of CCUG repeat RNA 1 was ~47 °C) and the increased stability provided by the Ψ modification to the CCUG repeats was not sufficient to strongly inhibit MBNL1 binding. Another possibility is that modifying both pyrimidines of the YGCY motif is necessary for robust inhibition of MBNL1 binding.

A model minimally structured RNA was used to probe the difference between the two pyrimidine positions in the YGCY motif. The modification studies with the minimally structured RNA revealed that the Ψ substitution 3′ to the GC dinucleotide had a much more pronounced effect on MBNL1 binding affinity compared with the 5′ substitution (Fig. 3, *RNAs 4* and *5*). This difference in binding affinity does not appear to be due to thermal stability as the 5′ and 3′ modifications were similarly stabilized by Ψ modification (T_m of 39 and 37 °C, respectively; Fig. 4*A*). This contrasts with findings in the context of duplex RNA, in which the position of Ψ within the duplex and the identity of adjacent Watson-Crick base pairs influence thermodynamic stability (26, 27). The difference in MBNL1 binding inhibition at the 5′- and 3′-pyrimidine positions in the model minimally structured RNA may be due to contacts made by the zinc fingers of MBNL1 with the 3′-pyrimidine observed in the crystal structure (15). The 3′ Ψ modification likely adopts conformations with bases in stacked structures that are less compatible with MBNL1 binding. The double Ψ modification that nearly eliminated MBNL1 binding (Fig. 3, *RNA 6*) again favors RNA conformations with single strand base stacking. These studies are consistent with the crystal structure that showed that the zinc fingers of MBNL1 interact primarily with the bases in an unstacked conformation.

Consistent with the melting studies, molecular dynamics simulations showed that Ψ modification of the 5′- or 3′-uridine increases the base-stacking interactions at the YGCY MBNL1-binding site of minimally structured RNAs. Although the structure published by Teplova and Patel (15) in 2008 shows that the 3′-uridine in the YGCY motif makes contact with the protein, this structure also shows that when the RNA is bound to the protein, the G of the GC dinucleotide stacks with the amino acids of the protein (Fig. 5*C*). These results impose a model in which MBNL1 is able to form interactions with YGCY RNA when it is flexible enough for the G to unstack with its neighboring pyrimidine. The MD simulations indicate that Ψ causes the RNA to prefer conformations in which the G is stacked with its neighbor. Based on the distribution of base separation distances, unmodified RNA also tends to form stacking interactions, but the population of those conformations is much less abundant than the range of open configurations. These data and our previous work (20) suggest that Ψ modifications can generally be used to shift YGCY motifs into conformations with increased base stacking that reduce binding by MBNL proteins.

Experimental Procedures

Thermal Melt Assays—Thermal melts were carried out in 20 mM PIPES buffer at a pH of 7.0 and 150 mM NaCl. PIPES buffer was chosen because of its minimal change in pH with increasing temperature (−0.0085 ΔpK_a/°C) (28). The buffer was de-gassed before diluting RNA to a concentration of 0.5, 1, 1.5, 2, 2.5, or 5 μM. The temperature was raised by 1 °C per min over a range of 12–70 °C for minimally structured model RNAs and 15–95 °C for CCUG repeat RNAs, held at the highest temperature for 2 min, and then lowered by 1 °C per min to the lowest temperature. Absorbance at 260 nM was measured once per min by a Cary UV-visible spectrophotometer. Absorbance was normalized by subtracting the absorbance at the lowest temperature measured. The melting temperature (T_m) of CCUG repeat RNAs was calculated as the temperature at which the deriva-

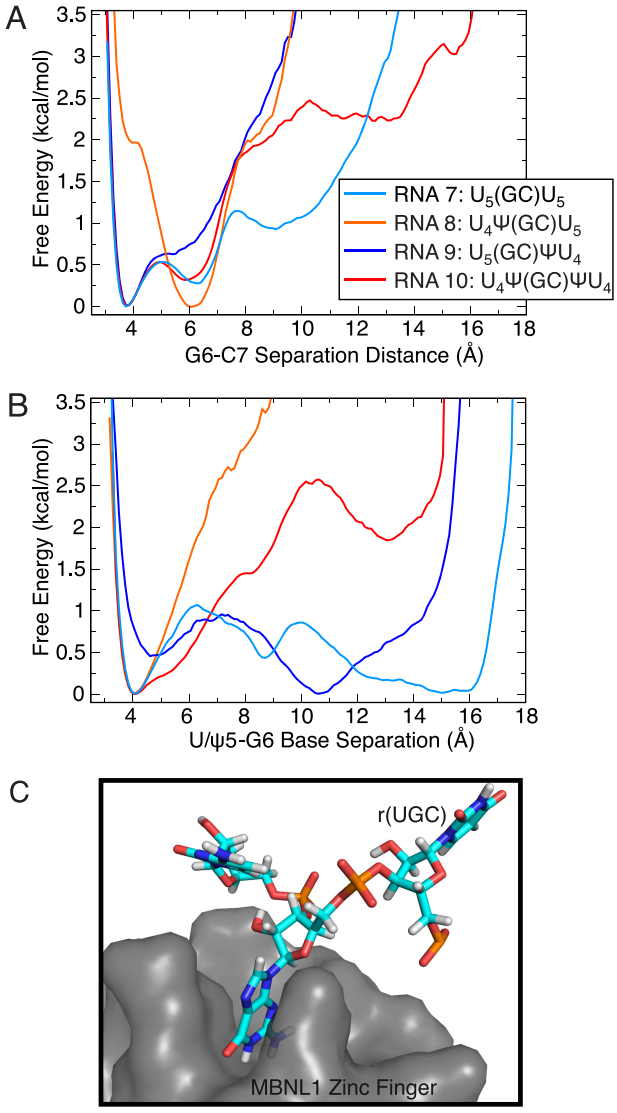

FIGURE 5. **Molecular dynamics simulations indicate that Ψ increases base stacking in minimally structured YGCY RNAs.** *A,* relative free energy of constructs at different G6-C7 separation distances in $U_5(GC)U_5$ RNA modified with Ψ as depicted in the legend. Unmodified RNA has energy minima at larger G6-C7 separation distances, whereas pseudouridylation of the 8th position in the RNA highly favors small G6-C7 separation distances. *B,* free energy of constructs at different U/Ψ5-G6 separation distances. Unmodified RNA has energy minima over a large range of U/Ψ5-G6 separation distances, whereas the U5 pseudouridylated constructs only have favorable conformations at short U/Ψ5-G6 separation distances. *C,* typical open structure of the unmodified $U_5(GC)U_5$ RNA from the simulations fit to the conformation of the guanine of the bound RNA in the crystal structure of Teplova *et al.* (15). The protein surface, which is not present in the simulations, is rendered in *light gray*, to illustrate the proposed model that effective MBNL1 binding to RNA requires the availability of conformations with the G in the YGCY-binding site unstacked from both neighbors.

tive of the absorbance was the highest. The T_m of a minimally structured RNA was calculated as the temperature at the halfway point between the minimum and maximum absorbance.

Protein Expression and Purification—GST-tagged MBNL1 (amino acids 2–260) was expressed in BL21 Star *Escherichia coli* cells (Thermo Fisher Scientific), grown to an OD of 0.6–0.7, induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, and shaken for 2 h at 37 °C. Bacteria were pelleted by centrifugation and lysed for 15 min at 4 °C with 4 ml/g B-PER™ bacterial protein extraction reagent. 0.1 mg/ml lysozyme, 5 units/ml DNase I, 0.5 mM $CaCl_2$, and 2.5 mM $MgCl_2$ were added to the lysate followed by incubation for 30 min at room temperature. An equal volume of 1× PBS was added, and cells were lysed for an additional 30 min on ice, followed by centrifugation at 40,000 × *g* for 15 min. Supernatant was incubated for 2 h at 4 °C with 10 ml of a 50% slurry of glutathione-agarose (Sigma) in 1× PBS, followed by washing two times with 25 ml of 40 mM Bicine, pH 8.3, and 50 mM NaCl; two times with 20 ml of 20 mM Bicine, pH 8.3, and 1 M NaCl; and three times with 20 ml of 16 mM Bicine, pH 8.3, and 20 mM NaCl. Protein was eluted by incubating three times with 5 ml of 16 mM Bicine, pH 8.3, 20 mM NaCl, and 10 mM glutathione. Eluates were combined, run through a 0.2-μm filter, and concentrated to 250 μl using an Amicon Ultra-15 centrifugal filter unit with Ultracel-50 membrane. 10 ml of protein storage buffer (500 mM NaCl, 25 mM Tris-HCl, pH 7.5, 5 mM β-mercaptoethanol, and 50% glycerol) was added to the protein followed by concentration using an Amicon Ultra 0.5-ml 50K centrifugal filter, flash freezing, and storage at −80 °C.

Gel Shift Binding Assays—RNA synthesized by Dharmacon or IDT was radiolabeled with [α-^{32}P]ATP with a polynucleotide kinase and stored at −20 °C in low TE (10 mM Tris and 1 mM EDTA). RNA oligonucleotides were diluted in buffer (20 mM Tris, pH 7.5, 75 mM NaCl, and 5 mM $MgCl_2$) and folded by incubation at 95 °C for 2 min followed by incubation on ice for 5 min. RNA (final concentration 1 nM) was incubated together for 30 min at room temperature with increasing concentrations of GST-MBNL1 (from 20 nM to 5 μM in 2-fold steps for RNAs 1 and 2, from 3.125 nM to 3.2 μM for RNAs 3, 5, and 6, and from 9.8 nM to 5 μM for RNA 4) in 115 mM NaCl, 20 mM Tris, pH 7.5, 1 mM β-mercaptoethanol, 0.01 mM EDTA, 10% glycerol, 5 mM $MgCl_2$, 0.1 mg/ml heparin, 2 mg/ml BSA, and 0.02% xylene cyanol. RNA-protein complexes were separated from free RNA on a 6% native acrylamide gel. Complex formation was quantified by exposure on a phosphorimager screen followed by analysis with ImageQuant from GE Healthcare. Affinity constants were then calculated with the following equation: $f_{bound} = f_{max}([MBNL1]/([MBNL1 + K_d))$. The images of the gels were enhanced equally across each gel using Adobe Photoshop. No specific feature was enhanced.

Molecular Dynamics Simulations—Atomistic explicit solvent simulations of minimally structured $U_5(GC)U_5$, $U_4Ψ(GC)U_5$, $U_5(GC)ΨU_4$, and $U_4Ψ(GC)ΨU_4$ RNA constructs were performed using the GROMACS (Groningen Machine for Chemical Simulations (29–32) molecular dynamics package on the local ACISS (Applied Computational Instrument for Scientific Synthesis) cluster at the University of Oregon, using an AMBER (Assisted Model Building with Energy Refinement) (33) force field optimized for stem-loop RNA structures with modified force-field parameters for the Ψ base (34). All simulations were performed in explicit solvent utilizing the spc/e water model and 150 mM NaCl with excess sodium ions to neutralize the system, Ewald summation for the electrostatics, and standard force-field cutoffs and parameters. The RNA was initially in an A-form structure that was energy-minimized after solvation. A 1-fs time step was utilized in the simulation, which was equilibrated over the course of a 5-ns period. An additional 10-ns simulation was performed in the production NVT ensemble at 300 K, from which 10

random snapshots were taken to seed 10 production runs of 25 ns each from which to collect statistics.

RNA Secondary Structure Prediction—The sequence of CCUG repeat RNA 1 was run through the following secondary structure prediction software programs using default settings: RNAfold (minimum free energy and centroid), IPknot (NUPACK model), Mfold, RNAstructure (FOLD and MaxExpect), and Sfold (Ensemble Centroid and MFE). Each predicted the same structure as most energetically favorable.

Author Contributions—E. D. conducted the experiments, analyzed the results, and wrote the manuscript for the first submission. M. N. H. conducted the experiments for Figs. 1, 2, and 4 and Tables 1 and 2, analyzed the results, and revised the manuscript for resubmission. K. D. contributed to experimental design, data analysis, and editing of the manuscript. J. C. performed and analyzed data from the molecular dynamics simulations. M. G. contributed to the development of the experimental plan, data analysis, and editing of the manuscript. J. A. B. conceived the idea for the project, developed the experimental plan, and wrote the manuscript. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments—We thank Dr. Art Pardi and Sabrina Hunt for the NMR experiments; Dr. Danielle Cass for previous work with minimally structured YGCY RNAs; Melissa Hale for assistance with protein purification; Jared Richardson for RNA preparation; Dr. Peter Von Hippel for generously providing access to equipment for collecting thermal melt data; members of the Berglund laboratory; and Dr. Karen Guillemin for input, support, and advice.

References

1. O'Rourke, J. R., and Swanson, M. S. (2009) Mechanisms of RNA-mediated disease. *J. Biol. Chem.* **284,** 7419–7423
2. Cho, D. H., and Tapscott, S. J. (2007) Myotonic dystrophy: emerging mechanisms for DM1 and DM2. *Biochim. Biophys. Acta* **1772,** 195–204
3. Lee, J. E., and Cooper, T. A. (2009) Pathogenic mechanisms of myotonic dystrophy. *Biochem. Soc. Trans.* **37,** 1281–1286
4. Miller, J. W., Urbinati, C. R., Teng-Umnuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., and Swanson, M. S. (2000) Recruitment of human muscleblind proteins to $(CUG)_n$ expansions associated with myotonic dystrophy. *EMBO J.* **19,** 4439–4448
5. Rau, F., Freyermuth, F., Fugier, C., Villemin, J. P., Fischer, M. C., Jost, B., Dembele, D., Gourdon, G., Nicole, A., Duboc, D., Wahbi, K., Day, J. W., Fujimura, H., Takahashi, M. P., Auboeuf, D., *et al.* (2011) Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.* **18,** 840–845
6. Ho, T. H., Charlet-B, N., Poulos, M. G., Singh, G., Swanson, M. S., and Cooper, T. A. (2004) Muscleblind proteins regulate alternative splicing. *EMBO J.* **23,** 3103–3112
7. Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014) circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* **56,** 55–66
8. Wang, E. T., Cody, N. A., Jog, S., Biancolella, M., Wang, T. T., Treacy, D. J., Luo, S., Schroth, G. P., Housman, D. E., Reddy, S., Lécuyer, E., and Burge, C. B. (2012) Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150,** 710–724
9. Batra, R., Charizanis, K., Manchanda, M., Mohan, A., Li, M., Finn, D. J., Goodwin, M., Zhang, C., Sobczak, K., Thornton, C. A., and Swanson, M. S. (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* **56,** 311–322
10. Goers, E. S., Purcell, J., Voelker, R. B., Gates, D. P., and Berglund, J. A. (2010) MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.* **38,** 2467–2484
11. Mankodi, A., Urbinati, C. R., Yuan, Q. P., Moxley, R. T., Sansone, V., Krym, M., Henderson, D., Schalling, M., Swanson, M. S., and Thornton, C. A. (2001) Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.* **10,** 2165–2170
12. Machuca-Tzili, L., Brook, D., and Hilton-Jones, D. (2005) Clinical and molecular aspects of the myotonic dystrophies: a review. *Muscle Nerve* **32,** 1–18
13. Fu, Y., Ramisetty, S. R., Hussain, N., and Baranger, A. M. (2012) MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation. *Chembiochem* **13,** 112–119
14. Cass, D., Hotchko, R., Barber, P., Jones, K., Gates, D. P., and Berglund, J. A. (2011) The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements. *BMC Mol. Biol.* **12,** 20
15. Teplova, M., and Patel, D. J. (2008) Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat. Struct. Mol. Biol.* **15,** 1343–1351
16. Charette, M., and Gray, M. W. (2000) Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49,** 341–351
17. Davis, D. R. (1995) Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res.* **23,** 5020–5026
18. Durant, P. C., and Davis, D. R. (1999) Stabilization of the anticodon stem-loop of tRNALys,3 by an A+-C base-pair and by pseudouridine. *J. Mol. Biol.* **285,** 115–131
19. Yarian, C. S., Basti, M. M., Cain, R. J., Ansari, G., Guenther, R. H., Sochacka, E., Czerwinska, G., Malkiewicz, A., and Agris, P. F. (1999) Structural and functional roles of the N1- and N3-protons of ψ at tRNA's position 39. *Nucleic Acids Res.* **27,** 3543–3549
20. deLorimier, E., Coonrod, L. A., Copperman, J., Taber, A., Reister, E. E., Sharma, K., Todd, P. K., Guenza, M. G., and Berglund, J. A. (2014) Modifications to toxic CUG RNAs induce structural stability, rescue mis-splicing in a myotonic dystrophy cell model and reduce toxicity in a myotonic dystrophy zebrafish model. *Nucleic Acids Res.* **42,** 12768–12778
21. Molinaro, M., and Tinoco, I., Jr. (1995) Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: thermodynamic and spectroscopic applications. *Nucleic Acids Res.* **23,** 3056–3063
22. Lietzke, S. E., Barnes, C. L., Berglund, J. A., and Kundrot, C. E. (1996) The structure of an RNA dodecamer shows how tandem U-U base pairs increase the range of stable RNA structures and the diversity of recognition sites. *Structure* **4,** 917–930
23. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA. *Biochemistry* **34,** 3204–3211
24. SantaLucia, J., Jr, Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A.C, C.C, G.G, U.C, and U.U mismatches in RNA internal loops: evidence for stable hydrogen-bonded U.U and C.C.+ pairs. *Biochemistry* **30,** 8242–8251
25. Sheng, J., Gan, J., Soares, A. S., Salon, J., and Huang, Z. (2013) Structural insights of non-canonical U*U pair and Hoogsteen interaction probed with Se atom. *Nucleic Acids Res.* **41,** 10476–10487
26. Hudson, G. A., Bloomingdale, R. J., and Znosko, B. M. (2013) Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. *RNA* **19,** 1474–1482
27. Kierzek, E., Malgowska, M., Lisowiec, J., Turner, D. H., Gdaniec, Z., and Kierzek, R. (2014) The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res.* **42,** 3492–3501
28. Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S., and Singh, R. M. (1966) Hydrogen ion buffers for biological research. *Biochemistry* **5,** 467–477
29. Berendsen, H. J., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91,** 43–56
30. Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Mol. Model. Annu.* **7,** 306–317
31. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26,** 1701–1718

32. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4,** 435–447
33. Zgarbová, M., Otyepka, M., Sponer, J., Mládek, A., Banáš, P., Cheatham, T. E., 3rd, and Jurečka, P. (2011) Refinement of the Cornell *et al.* nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.* **7,** 2886–2902
34. Aduri, R., Psciuk, B. T., Saro, P., Taniga, H., Schlegel, H. B., and SantaLucia, J. (2007) AMBER force field parameters for the naturally occurring modified nucleosides in RNA. *J. Chem. Theory Comput.* **3,** 1464–1475

Pseudouridine Modification Inhibits Muscleblind-like 1 (MBNL1) Binding to CCUG Repeats and Minimally Structured RNA through Reduced RNA Flexibility

Elaine deLorimier, Melissa N. Hinman, Jeremy Copperman, Kausiki Datta, Marina Guenza and J. Andrew Berglund

J. Biol. Chem. 2017, 292:4350-4357.
doi: 10.1074/jbc.M116.770768 originally published online January 27, 2017