

# Single-molecule study of the CUG repeat–MBNL1 interaction and its inhibition by small molecules

Amin Haghghat Jahromi<sup>1,2</sup>, Masayoshi Honda<sup>3</sup>, Steven C. Zimmerman<sup>1,2,\*</sup> and Maria Spies<sup>3,\*</sup>

<sup>1</sup>Center for Biophysics and Computational Biology, University of Illinois, Urbana, IL 61801, USA, <sup>2</sup>Department of Chemistry, University of Illinois, Urbana, IL 61801, USA and <sup>3</sup>Department of Biochemistry, University of Iowa Carver College of Medicine, Iowa City, IA 52242, USA

Received February 5, 2013; Revised March 28, 2013; Accepted April 9, 2013

## ABSTRACT

Effective drug discovery and optimization can be accelerated by techniques capable of deconvoluting the complexities often present in targeted biological systems. We report a single-molecule approach to study the binding of an alternative splicing regulator, muscleblind-like 1 protein (MBNL1), to (CUG)<sub>n=4,6</sub> and the effect of small molecules on this interaction. Expanded CUG repeats (CUG<sup>exp</sup>) are the causative agent of myotonic dystrophy type 1 by sequestering MBNL1. MBNL1 is able to bind to the (CUG)<sub>n</sub>-inhibitor complex, indicating that the inhibition is not a straightforward competitive process. A simple ligand, highly selective for CUG<sup>exp</sup>, was used to design a new dimeric ligand that binds to (CUG)<sub>n</sub> almost 50-fold more tightly and is more effective in destabilizing MBNL1–(CUG)<sub>4</sub>. The single-molecule method and the analysis framework might be extended to the study of other biomolecular interactions.

## INTRODUCTION

Myotonic dystrophy type 1 (DM1) is a triplet-repeat disease originating from a progressive expansion of the CTG repeat (CTG<sup>exp</sup>) in the 3'-untranslated region of the *DMPK* gene (1). The expanded *DMPK* gene produces a toxic RNA transcript (CUG<sup>exp</sup> containing RNA) that does not exit the nucleus but associates with proteins. One of these proteins, muscleblind-like 1 protein (MBNL1), is an important regulator of alternative splicing (2). Sequestration of MBNL1 in nuclear foci leads to multiple mis-spliced pre-mRNAs, incorrect protein levels and ultimately the disease (3). In a mouse model of DM1, a morpholino antisense oligonucleotide (ASO) (1), a 2'-*O*-(2-methoxyethyl) ASO (4) and a D-amino acid hexapeptide (ABP1) (5) reversed this process, rescued normal splicing

and reversed the phenotype in mice thereby validating CUG<sup>exp</sup> as a therapeutic target. Because no drugs are currently available to treat DM1, there is intense interest in finding small molecules that may function in a manner similar to the morpholino antisense oligonucleotides, but avoid the limitations inherent in the antisense therapeutic approach (6). Pentamidine (7), benzo[g]quinolone heterocycle derivatives (8), a Hoechst derivative (H1) (9) and modularly assembled Hoechst 33258 (10) are examples of bioactive CUG<sup>exp</sup> binders that are able to restore MBNL1 function in DM1 cell and animal models.

Central to the discovery of new and improved therapeutic agents for DM1 is to understand how small molecules bind CUG<sup>exp</sup> and the mechanism by which they block the MBNL1 binding. Previously, we and others have analyzed the MBNL1–CUG<sup>exp</sup>-inhibitor system using a competitive inhibition model and have used bulk solution techniques to analyze the equilibrium binding. However, these bulk solution methods usually require large sample volumes and have low sensitivity. Most importantly, these techniques do not provide the kinetic information that can be vital for accelerating drug discovery and development (11) and sometimes do not give a full thermodynamic picture. Surface-based biosensors provide kinetic information, directly in real-time with a fast response and high sensitivity, by measuring the interaction between an immobilized macromolecule and its soluble binding partner. The localized surface plasmon resonance (LSPR) technique, in particular, is developing rapidly into a powerful method (12). Nonetheless, the SPR techniques are model dependent and can have difficulties distinguishing different analytes, for example, a protein from an inhibitor (13).

The lack of a suitable method for studying the MBNL1–CUG<sup>exp</sup> interaction and its inhibition by small molecules means that the overall complexation process remains largely unexplored. The same holds true for many other systems where it is desirable to disrupt off-pathway protein–RNA interactions without affecting the

\*To whom correspondence should be addressed. Tel: +1 319 335 7932; Fax: +1 319 335 9570; Email: maria-spies@uiowa.edu  
Correspondence may also be addressed to Steven C. Zimmerman. Tel: +1 217 333 6655; Fax: +1 217 244 8024; Email: sczimmer@illinois.edu

biologically important function of the protein. To aid our drug discovery program and provide a potentially general approach to study RNA–protein inhibition, we developed a single-molecule method to analyze the MBNL1–CUG<sup>exp</sup>-inhibitor system. The major advantage of our model-independent approach is the ability to detect and measure the individual binding events in real-time under equilibrium conditions (14). Analyzing the distributions of bound and free states gives the most reliable on and off rates and a full picture of the inhibition process (15,16).

Herein, we describe the use of total internal reflection fluorescence (TIRF) microscopy to study, at the single-molecule level, the inhibition of MBNL1 binding to CUG repeat by known small molecule inhibitor **1** (Figure 1). The synthesis of dimeric inhibitor ligand **2**, based on structure **1**, is described and evaluated using the TIRF method. Ligand **2** was shown to be bioactive in DM1 cell models (data will be published elsewhere). Both compounds are found to bind RNA at the same time as MBNL1, indicating that the inhibition of MBNL1–CUG<sup>exp</sup> interaction does not follow a simple competitive mechanism. Rather than competitive steric blocking of the MBNL1 complexation site on CUG repeat, it seems the ligands are accelerating the dissociation of the protein.

## MATERIALS AND METHODS

### Instrumentation and chemicals

All reagents were purchased from commercial suppliers and were used without further purification. Anhydrous solvents were obtained from an anhydrous solvent dispensing system. For all reactions using anhydrous solvents, glassware was oven-dried, cooled under

vacuum and then purged with dry nitrogen; all reactions were conducted under dry nitrogen. Purified compounds were further dried under high vacuum (0.01–0.05 Torr) or lyophilized using a Labconco lyophilizer. Yields refer to purified and spectroscopically pure compounds. Nuclear magnetic resonance (NMR) spectra were recorded at 23°C on either Varian Unity 500 or Varian Unity Inova 500NB, operating at 500 and 125 MHz for <sup>1</sup>H and <sup>13</sup>C acquisitions, respectively. NMR spectra were processed using MestReNova software. Chemical shifts are reported in ppm and referenced to the corresponding residual nuclei in the following deuterated solvents: CDCl<sub>3</sub> (7.26 ppm <sup>1</sup>H, 77.16 ppm <sup>13</sup>C); dimethyl sulfoxide (DMSO) (2.50 ppm <sup>1</sup>H, 39.52 ppm <sup>13</sup>C); D<sub>2</sub>O (4.79 ppm <sup>1</sup>H); and CD<sub>3</sub>OD (3.31 ppm <sup>1</sup>H, 50.41 ppm <sup>13</sup>C). Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), sext (sextet), dd (doublet of doublets), ddd (doublet of doublet of doublets), td (triplet of doublets), dt (doublet of triplets), m (multiplet) and b (broad). Integration is provided, and coupling constants, *J*, are reported in Hertz (Hz). Electrospray ionization (ESI) mass spectra were recorded using the Quattro or ZMD mass spectrometer. High-resolution mass spectra (HRMS) were obtained at the University of Illinois mass spectrometry facility. All compounds described herein gave NMR and mass spectral data in accord with their structures. Ligands **1** and **2** gave an HRMS within 5 ppm of calculated values.

**9-((4-((4,6-Diamino-1,3,5-triazin-2-yl)amino)butyl)amino)-N-methylacridine-2-carboxamide (1)**. A round-bottom flask, equipped with a stir bar, was charged with **S5** (290 mg, 1.07 mmol, 1 eq) and **S1** (233 mg, 1.18 mmol, 1.1 eq) (compounds **S1** and **S5** and synthetic scheme are shown in Supplementary Figure S1). *N,N*-diisopropylethylamine (DIPEA) (305 mg, 2.36 mmol, 2.2 eq) and anhydrous *N,N*-dimethylformamide (DMF) (20 ml) were added. The solution was heated at 70°C for 5 h. The solvent was removed by rotary evaporation, and the product was purified via flash chromatography (basic alumina; DCM:methanol:NH<sub>4</sub>OH, 95:4.9:0.1 to 85:14:1) to yield a yellow solid (377 mg, 0.76 mmol, 71%). <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ 8.91 (s, 1H), 8.33 (d, *J* = 8.6 Hz, 1H), 8.02 (dd, *J* = 9.1, 1.9 Hz, 1H), 7.85 (t, *J* = 7.6 Hz, 2H), 7.75–7.68 (m, 1H), 7.42–7.35 (m, 1H), 4.03 (t, *J* = 7.2 Hz, 2H), 3.37–3.33 (m, 2H), 3.00 (s, 3H), 1.93 (p, *J* = 7.4 Hz, 2H), 1.70 (p, *J* = 7.2 Hz, 2H); *m/z* HRMS (ESI) calculated for [M + H]<sup>+</sup>: 432.2260; found 432.2267.

***N,N'*-((Propane-1,3-diylbis(azanediyl))bis(propane-3,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide (2)**. A round-bottom flask, equipped with a stir bar, was charged with **S12** (310 mg, 0.261 mmol, 1 eq). (compound **S12** and synthetic scheme are shown in Supplementary Figures S2 and S3). trifluoroacetic acid (TFA) (30 ml) and anhydrous DCM (70 ml) was added, and the mixture was stirred at room temperature for 2 h. The solvents were removed to afford a yellow solid (437 mg, 0.261 mmol, 100%). <sup>1</sup>H NMR (500 MHz, deuterium oxide) δ 8.34 (d, *J* = 8.3 Hz, 2H), 8.17 (d, *J* = 8.3 Hz, 2H), 8.11 (d, *J* = 7.4 Hz, 2H), 7.78 (t, *J* = 7.7 Hz, 2H), 7.52 (d, *J* = 8.3 Hz, 2H), 7.48–7.39 (m, 4H), 4.07 (t, *J* = 6.8 Hz, 4H), 3.58 (t, *J* = 6.7 Hz, 4H), 3.37–3.29

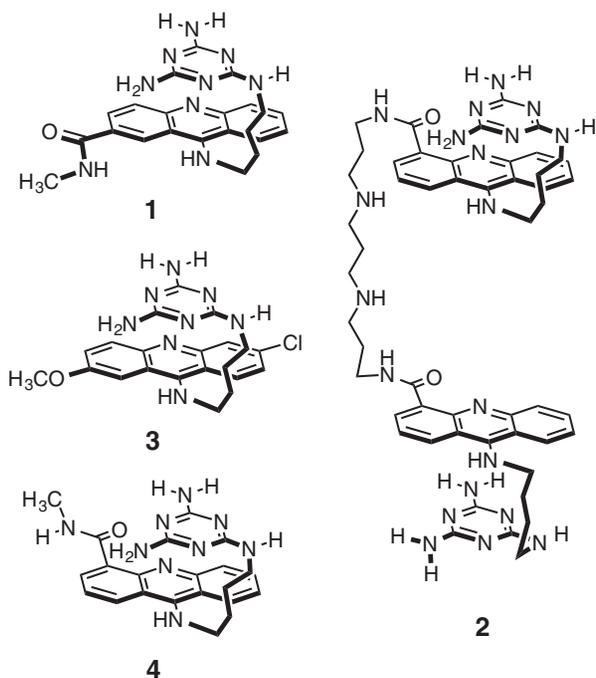


Figure 1. Compounds 1–4 used in this study.

(m, 4H), 3.26 (t,  $J = 5.9$  Hz, 8H), 2.31 (s, 2H), 2.14–2.08 (m, 4H), 1.98–1.91 (m, 4H), 1.71–1.65 (m, 4H); elemental analysis, calculated for  $C_{63}H_{70}F_{18}N_{20}O_{14}$ : C, 45.22%; H, 4.22%; F, 20.44%; N, 16.74%; found: C, 45.15%; H, 4.20%; F, 20.09%; N, 16.46%,  $m/z$  HRMS (ESI) calculated for  $[M + H]^+$ : 989.5599; found 989.5590.

### MBNL1N plasmid and RNA

The expression vector pGEX-6p-1/MBNL1N was obtained from Maurice S. Swanson (University of Florida, College of Medicine, Gainesville, FL, USA) (17). MBNL1N comprises the four zinc-finger motifs of MBNL1, the RNA-binding module of MBNL1 (17). It contains a 6xHis tag at the C-terminus and the Glutathione S-transferase (GST) tag at the N-terminus. MBNL1N binds RNA with similar affinity as the full-length MBNL1, but it does not form oligomers characteristic of the full-length protein (17). It is referred to as MBNL1 throughout this article for the sake of simplicity. All the oligonucleotides were purchased from Integrated DNA Technology and were high-performance liquid chromatography purified. The sequences and modifications for RNA constructs used in this study are shown in Supplementary Note S5.

### MBNL1N protein expression and purification

Using BL21-CodonPlus(DE3)-RP competent cells (Stratagene), the expression of MBNL1N protein was induced with 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) at  $OD_{600}$  0.6 in Lysogeny Broth (LB) media with ampicillin for 2 h at 37°C. Bacterial cells were collected by centrifugation and were then resuspended in a lysis buffer containing 25 mM Tris-Cl (pH = 8), 0.5 M NaCl, 10 mM imidazole, 2 mM beta-mercaptoethanol (BME), 5% glycerol, 0.1% Triton X-100, 2 mg/ml lysozyme, 0.1 mM phenylmethanesulfonylfluoride (PMSF), 1  $\mu$ M pepstatin and 1  $\mu$ M leupeptin and sonicated six times for 15 s each. The cell pellet was centrifuged, and the clarified lysate was collected and filtered through a 45- $\mu$ m Millex Filter (Millipore). To purify MBNL1N, Ni-Nitrilotriacetic acid (NTA) agarose (QIAGEN) was incubated with the lysate for 1 h at 4°C and washed with a washing buffer containing 25 mM Tris-Cl (pH = 8), 0.5 M NaCl, 20 mM imidazole and 0.1% Triton X-100, followed by elution with elution buffer of 25 mM Tris-Cl (pH = 8), 0.5 M NaCl, 250 mM imidazole and 0.1% Triton X-100. The eluate containing the GST fusion 6xHis-MBNL1N was dialyzed against phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM  $Na_2HPO_4$  and 2 mM  $KH_2PO_4$ , pH 7.4), for SPR studies. The molecular weight was confirmed by Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry, and the concentration was determined by Bradford assay.

### Preparation of Cy3-MBNL1 protein for TIRFM study

The GST fusion protein was incubated with Glutathione Sepharose 4B (GE Healthcare) for 1 h at 4°C. After washing with a buffer containing 25 mM Tris-Cl (pH = 8), 300 mM NaCl, 5 mM BME and 0.1% Triton

X-100, the beads were collected and incubated with PreScission Protease (GE Healthcare) overnight at 4°C. After being cleaved from the beads, the protein was collected in the flow-through of the column. Fluorescent labeling of MBNL1 was performed by coupling Cy3 mono-reactive NHS esters (GE Healthcare) to the N-terminal amine group at pH 7.0 (18–20). MBNL1 was mixed with a 12.5-fold molar excess of the Cy3 mono-reactive NHS ester in potassium phosphate buffer (62 mM  $K_2HPO_4$ , 38 mM  $KH_2PO_4$ , pH 7.05, 100 mM NaCl and 1 mM dithiothreitol (DTT)) for 10 min at room temperature. The reaction mixture was incubated for 12 h at 4°C. The labeling reaction was terminated by the addition of 50 mM Tris-HCl, pH 7.5. Cy3-labeled MBNL1 was separated from the free dye using PD SpinTrap G-25 column (GE healthcare). The ratio of dye incorporated per protein molecule was determined to be 1.1 mol Cy3 per 1 mol MBNL1.

### Surface plasmon resonance analysis

Detailed experimental procedure can be found in the Supplementary Methods.

### Steady-state fluorescence-based-binding assays

To determine the equilibrium parameters for binding of **1** and **2** to CUG<sup>exp</sup>, we followed quenching of TAMRA in TAMRA-(CUG)<sub>6</sub> at various ligand concentrations. The assays were performed using a Cary Eclipse Fluorescence Spectrophotometer (Varian). TAMRA-(CUG)<sub>6</sub> was excited at 560 nm, and its emission was recorded at 590 nm. Stoichiometric titrations were carried out at 20°C in PBS, 1 $\times$  buffer. The baseline fluorescence was recorded before addition of 20 nM TAMRA-(CUG)<sub>6</sub>. Increase in the fluorescence was recorded and attributed to the fluorescence of TAMRA-(CUG)<sub>6</sub>. On addition of each aliquot of **1** or **2**, the fluorescence signal was allowed to equilibrate, recorded for 5 min and averaged. Titration was continued at a series of increasing final concentrations of the ligand until the fluorescence was completely quenched. Fluorescence intensities at different concentrations of **1** (1.95, 3.91, 7.81, 15.6, 31.4, 62.8, 94.2, 125.6, 157, 500 and 1000  $\mu$ M) and **2** (39, 90, 210, 300, 390, 480, 570, 660, 750, 840 and 1250 nM) were fit to the following equation using Kaleidagraph software:

$$F = \frac{F_{\max} - F_0}{1 + \left(\frac{K_D}{L}\right)^n} + F_0$$

where  $K_D$  is the dissociation binding constant,  $[L]$  is the ligand concentration,  $n$  is the Hill coefficient and  $F_0$  and  $F_{\max}$  are the fluorescence intensity of free and fully bound RNA, respectively. In the control experiment, ligands **1** and **2** had no effect on the fluorescence intensity of the free TAMRA.

### Reaction conditions for the single-molecule assay

Biotinylated (CUG)<sub>4</sub> or (CUG)<sub>12</sub> were immobilized on a quartz surface, which was coated with polyethylene glycol to eliminate non-specific surface adsorption of proteins. The immobilization was mediated by biotin-neutravidin

interaction between biotinylated Cy5-labeled (CUG)<sub>4</sub>, neutravidin and biotinylated polymer (polyethylene glycol (PEG), MW = 5000). Standard PBS buffer contained the oxygen scavenging system consisting of 1 mg/ml glucose oxidase, D-glucose (0.4%, w/v), 2-mercaptoethanol (1%, v/v) and 0.04 mg/ml catalase. Immobilization of 50 pM of each oligonucleotide allowed for detection of 100–600 individual molecules per slide. Cy3-labeled MBNL1 was then added and incubated for 5 min at 25°C in the standard PBS, 1× buffer in the absence or presence of the indicated concentrations of **1** and **2**. Sample excitation and data recording were initiated after all components of the MBNL1–(CUG)<sub>4</sub>-inhibitor system were allowed to equilibrate in the TIRFM sample chamber. The presence of ligands had no effect on the fluorescence intensity of the Cy3-labeled MBNL1.

### Single-molecule data acquisition and analysis

TIRFM was used to excite fluorophores present near the surface, within the evanescent field. Cy3 fluorescence was excited by a DPSS laser (532 nm, 75 mW), whereas diode laser (641 nm, 100 mW) was used for direct Cy5 excitation. The fluorescence signals originated from the Cy3 and Cy5 dyes were collected by a water immersion ×60 objective, separated by a 630 nm dichroic mirror, passed through a 550 nm long-pass filter to block out laser scattering and detected by EMCCD camera with a time resolution of 100 ms. Single-molecule fluorescence trajectories were extracted from the recorded video file by IDL software. The collected trajectories were visualized using an in-house MATLAB program and were validated for the presence of the Cy5 signal by visual inspection of the acquired data. The Cy3 excitation regions of all individual trajectories were fit to the two-state model using the QuB software. This fitting generated the dwell times in the bound (ON) and free (OFF) states of the immobilized RNA molecules, which were binned and plotted as the histograms. The resulting ON and OFF event distributions were globally fit to exponential decay functions or to models described by Equations (1) and (2) using Prism 4.0 software to obtain the kinetic parameters.

## RESULTS

### Design and synthesis of small molecules

Our previous studies validated **3** (Figure 1) as a selective CUG binder, which was able to inhibit the MBNL1–CUG interaction with an apparent IC<sub>50</sub> = 52 ± 20 μM (21). Because the ultimate RNA target is CUG<sup>exp</sup>, a logical way to increase the selectivity and potency of this lead compound is to use the multivalent effect (22). Disney and co-workers (23) successfully applied this approach to the Hoechst 33258 ligand, which was known to bind CUG with modest selectivity. Because the rationally designed ligand **3** already exhibited a high affinity to (CUG)<sub>4</sub> and showed excellent selectivity, we initially pursued a dimeric ligand.

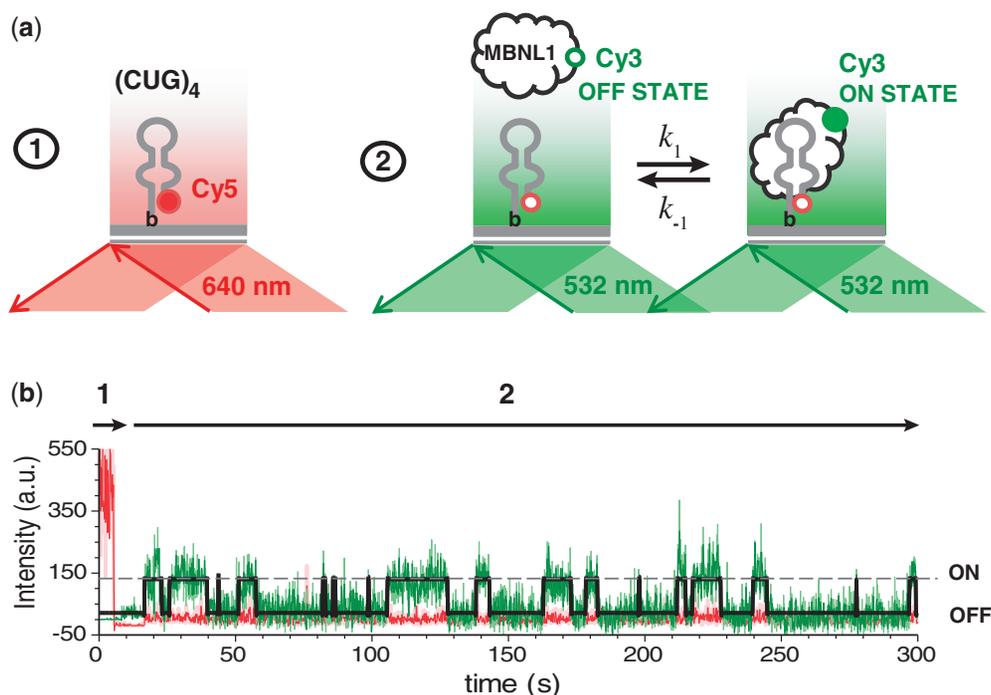
The covalent linkage of ligand **3** to itself or other compounds required a functional derivative. The three

possible sites for covalent modification of **3** are the acridine ring, the triaminotriazine recognition unit or the linking chain. The most expeditious synthetic approach used an acridine ring containing a carboxylic acid group to interconnect two ligands. It was found that the chloro- and methoxy-groups in **3** could be replaced with a 2- or 4-carboxamido group (see **1** and **4** in Figure 1, respectively) without altering its affinity for CUG or its inhibition of the MBNL1–CUG complex. Limited by the absence of structural data on the ligand–CUG<sup>exp</sup> complex, we designed and synthesized a small library of dimeric ligands that could potentially bind to consecutive CUG sites. This study, which will be published elsewhere, led us to **2** as the optimized dimeric ligand used in the current study (Figure 1). Ligand **2** was designed to target two consecutive UU mismatches, with the *N,N'*-bis(3-aminopropyl)-1,3-propanediamine linker spanning two central GC base pairs. Although the binding mode was not firmly established for **3**, it was designed to act as a ‘stacked intercalator’ with the acridine and triaminotriazine rings π-stacked while the intercalator sits between the GC base pair and the U–triaminotriazine–U base triplet (21). A recent combined experimental and computational study provides support for that binding model in the major groove of the CUG RNA (24). By positioning the linking chain and triaminotriazine recognition units on opposite sides of the acridine unit, the bivalent complex likely requires a threading mechanism for binding (25). This design element was intentional with the goal of increasing the binding affinity through a higher residence time (26,27).

### Establishing a TIRFM-based assay to monitor the MBNL1–(CUG)<sub>n</sub> interaction

To investigate binding of MBNL1 to (CUG)<sub>n</sub> and the effect of **1** and our bioactive dimeric ligand, **2**, on MBNL1–(CUG)<sub>n</sub> interaction, we developed a TIRFM-based single-molecule-binding assay (28,29). A stem-loop (5'-bio-GCUGCUGUGCGCUGCUG-3') containing two CUG pairs separated by a tetraloop (underlined), (CUG)<sub>4</sub>, was used as an RNA substrate to analyze the MBNL1 binding. This RNA construct has been previously established as the shortest CUG repeat to which MBNL1 binds with a similar affinity as to the long CUG repeats (30). The predicted stem-loop structure of (CUG)<sub>4</sub> has been confirmed by melting studies ( $T_m = 60.6 \pm 0.6$ ) (30,31). Although (CUG)<sub>4</sub> is much shorter than typical (CUG)<sub>n > 50</sub> structure, it forms a hairpin structure characteristic of (CUG)<sup>exp</sup>; thus, it is the shortest validated CUG repeat model that can be used in biochemical analysis. Moreover, it contains only one MBNL1-binding site per RNA molecule, which simplifies analysis of the association/dissociation events. Validation of the single-site binding is described in the next section later in the text.

We followed the binding of Cy3-labeled MBNL1 to, and its dissociation from, individual (CUG)<sub>4</sub> RNA molecules. Biotinylated and Cy5-labeled (CUG)<sub>4</sub> was immobilized on a PEG-coated surface of the TIRFM flow chamber via biotin–neutravidin interaction (Figure 2a, left). First, locations of the surface-tethered



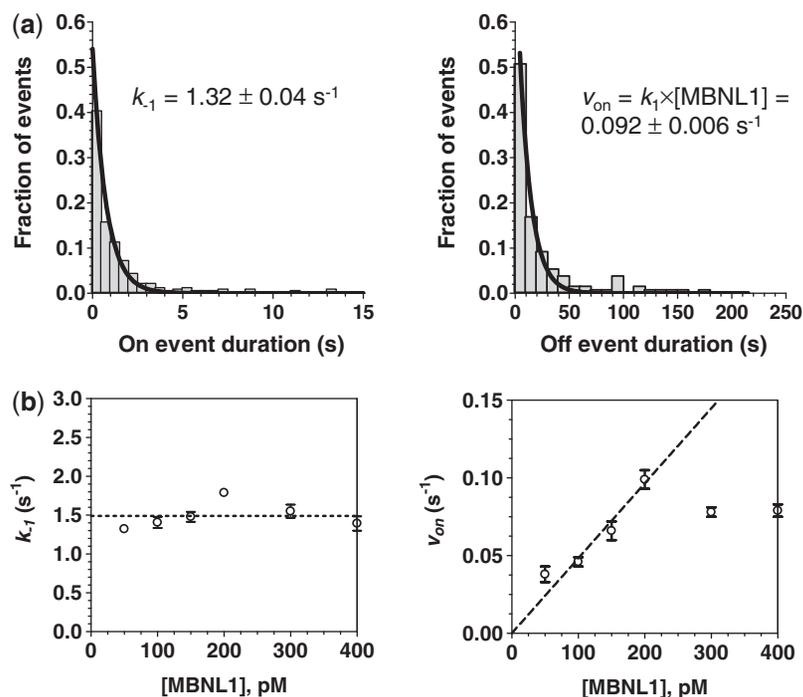
**Figure 2.** Single-molecule analysis of MBNL1 binding to  $(\text{CUG})_4$ . (a)  $(\text{CUG})_4$  was immobilized on a PEG-coated surface of the slide via biotin–neutravidin interaction. The experiment was carried out in two steps: (i) positions of  $(\text{CUG})_4$  molecules on the slide were triangulated by exciting Cy5 label of  $(\text{CUG})_4$  with a red laser and (ii) MBNL1– $(\text{CUG})_4$  interaction was monitored using green TIR illumination. On binding to  $(\text{CUG})_4$ , Cy3 label on MBNL1 is excited, as it is sequestered within the evanescent field. (b) Representative Cy3 (green) and Cy5 (red) fluorescence intensity trajectories from an individual  $(\text{CUG})_4$  molecule. Arrows 1 and 2 depict periods of red and green excitation, respectively. ON and OFF events are indicated on the right.

$(\text{CUG})_4$  molecules were identified by illuminating the slide with red laser (641 nm), which selectively excited the Cy5 fluorophore present on  $(\text{CUG})_4$ . Then, the laser was switched to green (532 nm) to monitor binding of Cy3-labeled MBNL1 to  $(\text{CUG})_4$  (Figure 2a, right). We followed MBNL1– $(\text{CUG})_4$  interaction in real-time, with a time resolution of 100 ms. In this experimental scheme, the surface-tethered  $(\text{CUG})_4$  and Cy3-labeled MBNL1 are at equilibrium during the length of the measurement. TIR-generated evanescent field excites only the Cy3-labeled MBNL1 molecules that reside near the surface longer than two frames of the camera, which only occurs when MBNL1 is bound to the surface-tethered  $(\text{CUG})_4$ . This mode of excitation eliminates the background fluorescence of Cy3-MBNL1 in solution and thereby allows monitoring MBNL1– $(\text{CUG})_4$  interactions over a broad range of MBNL1 concentrations.

Each event of MBNL1 binding to the surface-tethered  $(\text{CUG})_4$  was observed as the appearance of a Cy3 signal at the location where  $(\text{CUG})_4$  resided. Conversely, dissociation of MBNL1 from  $(\text{CUG})_4$  resulted in the disappearance of the Cy3 signal. Fluorescence trajectories (Figure 2b) showed multiple events of a two-state (ON and OFF) association and dissociation process. The presence of only two states indicates that each observed event corresponded to binding of a single MBNL1 to a single  $(\text{CUG})_4$ . To confirm that potential non-specific interaction of MBNL1 with the surface does not interfere with our analysis, we substituted Cy5-labeled  $(\text{CUG})_4$  with Cy5-labeled streptavidin. Only a few fluorescence trajectories showing

both Cy5 and Cy3 signals were observed in the presence of 300 pM Cy3-MBNL1 [ $\sim 5\%$  of what was observed in the presence of Cy5-labeled  $(\text{CUG})_4$ ], confirming that the vast majority of binding events described earlier in the text resulted from specific MBNL1– $(\text{CUG})_4$  interaction. Analysis of the Cy3-MBNL1– $(\text{CUG})_4$  interaction in the presence of unlabeled MBNL1 confirmed that MBNL1-binding properties were not affected by the Cy3 conjugation (Supplementary Figure S9).

The Cy3 (green laser excitation) regions of the trajectories were fit to a two-state model using QuB software (32) (Supplementary Figure S4) yielding dwell times in the OFF (dissociated) and ON (bound) states. We determined kinetic parameters of MBNL1 binding to the surface-tethered  $(\text{CUG})_4$  by globally analyzing distributions of ON and OFF dwell times in the presence of different concentrations of MBNL1 protein. All protein concentrations (50, 100, 150, 200, 300 and 400 pM) were used to determine the dissociation rate constant, whereas the association rate constant was determined using only 50–200 pM range of concentrations, which gave a linear dependence of the observed association rate on MBNL1 concentration. Figure 3a and b show individually fit ON and OFF time distributions obtained in the presence of 50 and 200 pM MBNL1, respectively. Global analyses of the dwell distributions yielded the dissociation rate constant for MBNL1– $(\text{CUG})_4$  complex,  $k_{-1} = 1.44 \pm 0.3 \text{ s}^{-1}$ , and association rate constant,  $k_1 = (4.6 \pm 0.2) \times 10^8 \text{ s}^{-1} \text{ M}^{-1}$ , from which we calculated the equilibrium dissociation constant,  $K_{D1} = k_{-1}/k_1 = 3.1 \pm 0.1 \text{ nM}$ .



**Figure 3.** (a) Distributions of all ON (left) and OFF (right) event durations detected in the 50 trajectories of individual  $(\text{CUG})_4$  molecules are fit to single exponentials to yield  $k_{-1}$  ( $[\text{MBNL1}] = 50 \text{ pM}$ ) and  $v_{\text{on}}$  ( $[\text{MBNL1}] = 200 \text{ pM}$ ), respectively. (b) Effect of MBNL1 concentration on the  $k_{-1}$  and  $v_{\text{on}}$  of the observed events is shown.  $k_{-1}$  (dissociation rate constant) was independent of  $[\text{MBNL1}]$ , suggesting the absence of functional cooperation between monomers. Initially,  $v_{\text{on}}$  (association rate) increased with increasing  $[\text{MBNL1}]$ . It was saturated at the MBNL1 concentrations  $>200 \text{ pM}$ .

### Stoichiometry of the MBNL1– $(\text{CUG})_n$ interaction

Biotinylated  $(\text{CUG})_{12}$  can simultaneously accommodate multiple Cy3-labeled MBNL1 molecules. Its Cy3 trajectories contained five states characterized by different Cy3 intensities (Figure 4a). An intensity histogram showed that there are three major Cy3 intensity populations consistent with the binding of one ( $\sim 196 \text{ a.u.}$ ), two ( $\sim 332 \text{ a.u.}$ ) and three ( $\sim 458 \text{ a.u.}$ ) MBNL1 proteins to the  $(\text{CUG})_{12}$ . A small fourth peak ( $\sim 632 \text{ a.u.}$ ) may represent more than one Cy3 label present on a fraction of MBNL1 molecules (Cy3 labeling efficiency of MBNL1 was 110%). Stoichiometry determination assay using SPR confirmed a stoichiometric ratio of MBNL1: $(\text{CUG})_{12}$  close to 3:1 (Supplementary Figure S10).

### Revealing the mechanism of the MBNL1– $(\text{CUG})_n$ interaction inhibition by 1 and 2

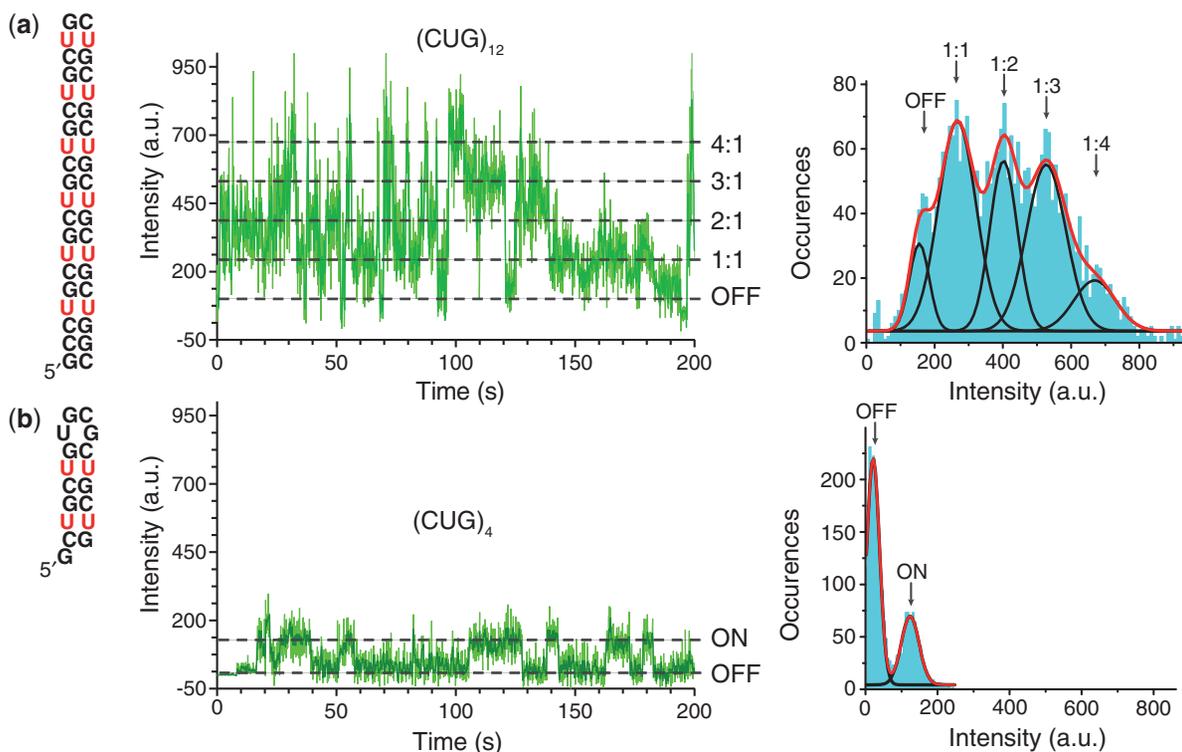
We then carried out MBNL1– $(\text{CUG})_4$ -binding studies in the presence of increasing concentrations of **1** and **2**. Before starting the recording, the Cy3-labeled MBNL1, and the ligands were incubated for at least 5 min in the TIRFM reaction chamber to ensure that all components of the system are at equilibrium. The difference in the dwell-time distributions for the MBNL1– $(\text{CUG})_4$  complex in the absence and presence of **1** and **2**, and therefore the difference in the apparent  $k_{\text{off}}$  of the MBNL1– $(\text{CUG})_4$  complex, suggested that both **1** and **2** do not act as simple competitive inhibitors, which should only affect the on-rate (Supplementary Figures S5–S8). Based on the proposed modes of  $(\text{CUG})_4$  binding by the ligand and

MBNL1, it is possible that both can coexist on the same  $(\text{CUG})_4$  molecule: although the ligand is expected to interact with the U–U mismatch (21), MBNL1 specifically binds to consecutive GC nucleotides (33). The observed inhibition likely stems from the different affinity of the MBNL1 protein for naked and ligand-bound  $(\text{CUG})_4$ . Most of this effect originates from an increased off-rate of the  $(\text{CUG})_4$ –MBNL1–ligand complex compared with  $(\text{CUG})_4$ –MBNL1. This situation can be described by a closed scheme of linked equilibria (Table 1a).

Assumptions derived from our experimental design and conditions (described in detail in the Supplementary Note S1) allowed us to globally fit distributions of the time intervals between binding events, obtained at a range of ligand concentrations, to Equation (1). Equation (1) assumes a simplified open scheme (Table 1b) and contains a double exponential

$$Y = \text{Const} \times [K_{D2} \times \exp(-k_1 \times [\text{MBNL1}] \times t) + [I] \times \exp(-k_3 \times [\text{MBNL1}] \times t)] \quad (1)$$

where the decay constants correspond to kinetic association constants for MBNL1 binding to  $(\text{CUG})_4$ ,  $k_1$ , and  $(\text{CUG})_4$ –I complex,  $k_3$ , respectively. The two exponentials are weighted by the abundance of each manifold under given conditions, which are defined by the equilibrium between free and ligand-bound  $(\text{CUG})_4$ . The experiments were carried out in the presence of 0, 4, 20, 100 and 200  $\mu\text{M}$  of ligand **1** or in the presence of 0, 0.5, 1, 2, 4, 6 and 8  $\mu\text{M}$  of ligand **2**.



**Figure 4.** Each  $(\text{CUG})_4$  can accommodate a single-MBNL1 protein, whereas three MBNL1 proteins can be simultaneously bound to  $(\text{CUG})_{12}$ . (a) Schematic representation of  $(\text{CUG})_{12}$  (left). (b) Schematic representation of  $(\text{CUG})_4$ . Representative fluorescence intensity trajectories for  $(\text{MBNL1})_3$ - $(\text{CUG})_{12}$  (a) and  $\text{MBNL1}$ - $(\text{CUG})_4$  (b) are shown in the middle (light green line shows the raw data collected at 100 ms time resolution, dark green line shows intensity averaged for nine data points). The fluorescence intensity histogram compiled from the trajectory is shown on the right (aqua) and is fit to five Gaussian peaks (black lines), corresponding to stepwise increase in  $(\text{CUG})_{12}$  occupancy by MBNL1 (a), whereas only a single 1:1 binding state is observed for  $\text{MBNL1}$ - $(\text{CUG})_4$  (b).

The closed scheme in Table 1a was used to analyze the ON time distributions. These distributions were globally fit to Equation (2).

$$Y = \text{Const} \times [K_{D2} \times \exp(-(k_{-1} - k_{-4}[I]) \times t) + [I] \times \exp(-(k_{-3} - k_{-4}) \times t)] \quad (2)$$

$\text{Const} \times K_{D2}$  is the weight of manifold (1), whereas  $\text{Const} \times [I]$  is the weight of manifold (2) adjusted by the number of observed events.

These analyses yielded  $K_{D2}$ ,  $k_4$  and  $(k_{-3} - k_{-4})$ . Parameters  $k_{-3}$  and  $k_{-4}$  are linked as  $(k_{-3} - k_{-4})$  and, therefore, cannot be determined individually by fitting but can be calculated from the linked equilibria as  $K_{-3} = \frac{(K_{-3} - K_{-4}) \times K_3 \times K_{D1}}{K_3 \times K_{D1} - K_4 \times K_{D2}}$ .

The results are summarized in the Table 1.

Indicative of its function as a bivalent inhibitor, the binding affinity of **2** for  $(\text{CUG})_4$  was almost 50-fold higher than that of **1**. Because neither **1** nor **2** behave as typical competitive inhibitors, but instead allow formation of the  $(\text{CUG})_4$ -MBNL1-ligand complex, the traditionally defined  $K_I$  is replaced by  $K_{D2}$ , which represents the affinity of the ligand for  $(\text{CUG})_4$ . We can, however, define an apparent  $\text{IC}_{50}$  as the concentration of ligand at which 50% of MBNL1 is free (Supplementary Note S2). It is notable, however, that this apparent  $\text{IC}_{50}$  depends on the concentration of the MBNL1 and  $(\text{CUG})_4$  (Figure 5)

making each inhibitor to be effective in a relatively narrow concentration range.

At micromolar concentrations, many small molecules self-associate into colloidal aggregates that non-specifically inhibit protein activity. Aggregate-based inhibition is sensitive to detergent (34); therefore, we confirmed that the inhibition by **2** was not affected by the presence of Triton X-100, eliminating the possibility of ligand aggregation.

To ensure that the observed inhibition mode does indeed stem from the  $(\text{CUG})_4$ -MBNL1-ligand complex formation and not an experimental artifact, we analyzed the  $(\text{CUG})_4$ -MBNL1 binding in the presence of the unlabeled MBNL1 protein, which should act a competitive inhibitor of Cy3-MBNL1 binding to the immobilized  $(\text{CUG})_4$ . Addition of unlabeled MBNL1 had no effect on  $k_{-1}$  for the  $(\text{CUG})_4$ -MBNL1 complex; in contrast  $k_1$  decreased proportionally to the fraction of Cy3-MBNL1 (Supplementary Figure S9). Not only did this control experiment confirm that we can clearly distinguish the different modes of inhibition but also that Cy3-labeling does not interfere with  $(\text{CUG})_4$ -MBNL1 interaction.

#### Binding affinity and MBNL1 inhibition potency of ligand **2** are nearly 100-fold that of **1** in bulk experiments

To measure the binding constants of **1** and **2** to CUG repeat by a more traditional approach, a steady-state

**Table 1.** Parameters from global fitting of the dwell-time distributions to the closed (a) and open (b) schemes of linked equilibria describing state interconversion in the (CUG)<sub>4</sub>-MBNL1–ligand system

**(a)**

**(b)**

Parameter	Significance	Value	
<b>Manifold 1</b>			
$k_1, \text{s}^{-1} \text{M}^{-1}$	Rate constant for MBNL1 associating to naked (CUG) <sub>4</sub>	$(4.6 \pm 0.2) \times 10^8$	
$k_{-1}, \text{s}^{-1}$	Rate constant for MBNL1 dissociating from naked (CUG) <sub>4</sub>	$1.44 \pm 0.3$	
$K_{D1} = k_{-1}/k_1, \text{nM}$	Affinity of MBNL1 for (CUG) <sub>4</sub>	$3.1 \pm 0.1$	
<b>Manifold 2</b>			
$K_{D2}, \mu\text{M}$	Affinity of ligand for (CUG) <sub>4</sub>	$0.45 \pm 0.02$	$22 \pm 3$
$(k_{-3}-k_{-4}), \text{s}^{-1}$		$4.9 \pm 0.3$	$2.7 \pm 0.1$
$k_{-3}, \text{s}^{-1}$	Rate constant for MBNL1 dissociating from (CUG) <sub>4</sub> -MBNL1-I	$5.4 \pm 0.3$	$2.9 \pm 0.1$
$k_{+3}, \text{s}^{-1}\text{M}^{-1}$	Rate constant for MBNL1 associating to (CUG) <sub>4</sub> -I	$(2.7 \pm 0.2) \times 10^8$	$(2.8 \pm 0.2) \times 10^8$
$K_{D3} = \frac{K_{-3}}{K_{+3}}, \text{nM}$	Affinity of MBNL1 for ligand-bound (CUG) <sub>4</sub>	$20 \pm 2^a$	$10.2 \pm 0.8^a$
$\frac{K_{D3}}{K_{D1}}$	Affinity of MBNL1 for ligand-bound relative to naked (CUG) <sub>4</sub>	$0.16 \pm 0.01$	$0.30 \pm 0.06$
$k_{-4}, \text{s}^{-1}$	Rate constant for ligand dissociating from (CUG) <sub>4</sub> -MBNL1-I complex	$0.55 \pm 0.03$	$0.12 \pm 0.01$
$k_{+3}, \text{s}^{-1}\text{M}^{-1}$	Rate constant for ligand associating to (CUG) <sub>4</sub> -MBNL1	$(1.93 \pm 0.03) \times 10^5$	$(1.7 \pm 0.7) \times 10^3$
$K_{D4} = \frac{K_{-4}}{K_{+4}}, \mu\text{M}$	Affinity of ligand for MBNL1-bound (CUG) <sub>4</sub>	$2.9 \pm 0.2^a$	$71 \pm 27^a$

$k_i$  is the association rate constant,  $k_{-i}$  is the dissociation rate constant and  $K_{Di} = k_{-i}/k_i$  is the equilibrium dissociation constant for the respective steps in the scheme. I, inhibitor (**1** or **2**).

Unless indicated, the errors shown are standard errors from fitting the data.

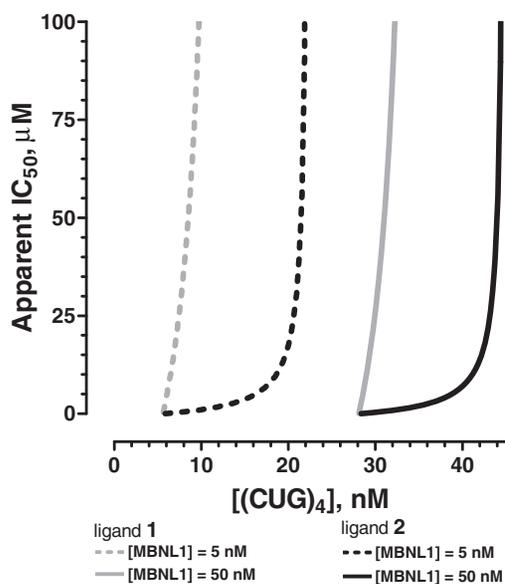
<sup>a</sup>Errors were calculated as  $\Delta(K_D) = K_D \times \sqrt{\left(\frac{\Delta K_{off}}{K_{off}}\right)^2 + \left(\frac{\Delta K_{on}}{K_{on}}\right)^2}$

fluorescence titration method with TAMRA-(CUG)<sub>6</sub> was used. It is known that guanosine quenches the TAMRA fluorescence through photoinduced electron transfer (35,36). Therefore, it is possible that binding of a ligand to the UU mismatch close to 5'-TAMRA makes a structural change that can lead to quenching of the 5'-TAMRA by the 3'-G in 5'-TAMRA-(CUG)<sub>6</sub>-3'. Ligands **1** and **2** were titrated into TAMRA-(CUG)<sub>6</sub> solution.

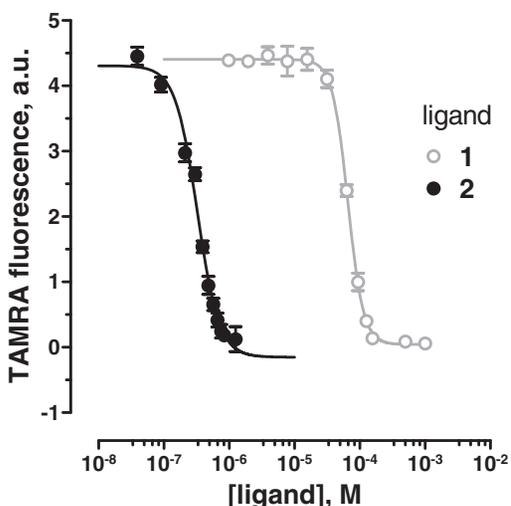
On increasing the ligand concentration, TAMRA fluorescence intensity gradually decreased as a result of fluorophore quenching by the bound ligand. A plot of normalized fluorescence intensity versus increased concentrations of each ligand yielded a binding isotherm with  $K_D$  of  $66 \pm 2 \mu\text{M}$  and  $318 \pm 15 \text{nM}$  for **1** and **2**, respectively (Figure 6). These  $K_D$  values are similar to what we derived from the single-molecule study (Table 1). To confirm that the observed change in TAMRA fluorescence is a direct result of the ligand binding and is not caused by RNA aggregation, we compared the absorbance of the

RNA–ligand solution before and after it was filtered through 0.02- $\mu\text{m}$  (Whatman Anotop 6809-1002) filter. No loss of RNA because of aggregation or precipitation was observed in the presence of the two assayed ligands (data not shown).

To provide a more conventional confirmation of single-molecule results, we used SPR to study the inhibition of MBNL1 binding to (CUG)<sub>4</sub> and (CUG)<sub>12</sub>. To distinguish binding of ligands from MBNL1 to the immobilized RNA constructs, the injection of MBNL1 was preceded by ligand injection. The experiments were done in the presence of a large excess of competitor tRNA (1  $\mu\text{M}$  yeast tRNA) to confirm the specificity of ligands for CUG repeat. The maximum Response Unit (RU) on MBNL1 injection at various concentrations of each inhibitor was recorded. These values were normalized and plotted versus increasing concentration of each ligand to yield a binding isotherm. Apparent IC<sub>50</sub> values against (CUG)<sub>4</sub> and (CUG)<sub>12</sub> for **1** were  $174 \pm 12$  and



**Figure 5.** Theoretical curves for apparent IC<sub>50</sub> values for **1** (gray) and **2** (black) at [MBNL1] = 5 (dashed lines) and 50 nM (solid lines). The concentration of ligand required to observe half of MBNL1 in the free form, IC<sub>50</sub>, depends on actual concentrations of MBNL1 and (CUG)<sub>4</sub>. The curves were generated based on the experimentally determined affinities shown in Table 1.



**Figure 6.** Fluorescence titrations of TAMRA-(CUG)<sub>6</sub> with **1** and **2**. Fluorescence intensity of TAMRA-(CUG)<sub>6</sub> in the presence of increasing concentrations of **1** (gray open circles) and **2** (black filled circles). TAMRA was excited at 560 nm, and its emission was recorded at 590 nm. Error bars represent mean ± standard deviation of three replicates.

293 ± 19 μM, respectively, whereas for **2**, the values were 1.3 ± 0.2 and 1.1 ± 0.1 μM, respectively (Supplementary Note S3 and Supplementary Figure S11). Because of the non-competitive inhibition mechanism, IC<sub>50</sub> values should depend on the concentration of immobilized (CUG)<sub>n</sub> on the SPR sensor chip, which is not measurable. Nonetheless, obtaining smaller IC<sub>50</sub> value for **2**, compared with **1**, is in agreement with the single-molecule results. It is noteworthy that **1** and **2** showed

no binding to MBNL1 (Supplementary Note S4 and Supplementary Figure S12).

## DISCUSSION

RNA is emerging as an attractive drug target because of its central role in transcription, translation and an ever-growing number of key cellular processes (37). Most traditional RNA targeted drug discovery has focused on antibiotics and anti-HIV drug candidates (38). However, non-coding RNA, pre-mRNA and mRNA as a drug target have recently gained significant attention (39). In many diseases, such as DM1, targeting RNA is the most appropriate option, as the RNA is the causative agent of the pathology, whereas the protein function should not be altered (40). Despite their potential health benefits, the development of small molecules that specifically target RNA is still in its infancy. Previous studies, including our own, have identified small molecule inhibitors of the MBNL1-(CUG)<sup>exp</sup> interaction. In each case, it was assumed that the identified small molecule acts as a conventional competitive inhibitor. To further develop this approach to drug discovery, it is vital to establish the exact mechanism of the inhibition and ideally to provide a full kinetic and thermodynamic picture of the process (11). Thus, we undertook the first study of its kind to provide a full description of the mechanism by which a protein-RNA interaction is inhibited at the single-molecule level.

We developed and used a TIRFM-based single-molecule analysis to study the interaction of MBNL1, a key regulatory protein in alternative splicing, and (CUG)<sub>4</sub>, a validated model of CUG<sup>exp</sup> as well as the inhibition of this interaction by the small molecule RNA-binding ligands. This model system represents a single independent unit of RNA-protein or RNA-ligand-protein interaction. Our analysis revealed the affinity of MBNL1 for (CUG)<sub>4</sub> to be 3.1 ± 0.1 nM, whereas using Electrophoretic Mobility Shift Assay (EMSA), we (21) and others (30) reported K<sub>D</sub> values of 26 ± 4 and 170 ± 20 nM, respectively. This was not surprising because the single-molecule measurements were carried out under true equilibrium conditions, whereas EMSA can underestimate affinity because the short-lived complexes may dissociate while being resolved. Additionally, only productive (native) interactions between (CUG)<sub>4</sub> and MBNL1 molecules were detected by the single-molecule measurements, thereby canceling out any possible errors resulting from heterogeneity of MBNL1 or (CUG)<sub>4</sub>.

Single-molecule analyses of the MBNL1-(CUG)<sub>4</sub> interaction in the presence of two inhibitors allowed us to determine the affinities of the two tested ligands for (CUG)<sub>4</sub>, as well as how they affect MBNL1-(CUG)<sub>4</sub> complex formation. Notably, the equilibrium dissociation constants for the ligand-(CUG)<sub>n</sub> binding determined in bulk and at the single-molecule level were similar. Converging of the binding affinity of the **1**-(CUG)<sub>n</sub> and **2**-(CUG)<sub>n</sub> complexes (*n* = 4 or 6) measured by the fluorescence and single-molecule methods to the same value validated the single-molecule method. Our single-molecule analysis was

model independent and suggested an inhibition mode that has not been considered previously: instead of competing for the same binding site, the ligands and MBNL1 can form a ternary complex on (CUG)<sub>4</sub>.

In the more physiologically relevant buffer used in this study, ligand **1** was a relatively weak inhibitor, which bound to (CUG)<sub>4</sub> with a low affinity (i.e.  $K_D = 22 \pm 3 \mu\text{M}$ ). Moreover, it allowed MBNL1 to bind to the 1-(CUG)<sub>4</sub> complex with very high affinity ( $K_D = 10.2 \pm 0.8 \text{ nM}$ ). This interesting finding prompted us to design more effective small molecules. We pursued a bivalent ligand with the idea that it would: (i) bind to CUG<sup>exp</sup> with a higher affinity and ideally weaken the MBNL1 interaction with ligand-CUG<sup>exp</sup>; (ii) exhibit a longer occupancy time and increase the RNA ds stabilization, if MBNL1 indeed binds preferentially to the ss form of RNA as has been suggested (41,42); and (iii) bind a longer segment of RNA to achieve more effective coverage and steric blocking of CUG<sup>exp</sup>.

To reach the drug development goals outlined earlier in the text, we designed a library of dimeric ligands and found **2** to be the most potent inhibitor of CUG repeat-MBNL1 interaction. The affinity of **2** for (CUG)<sub>4</sub> was measured by the single-molecule method to be  $K_D = 450 \pm 20 \text{ nM}$ , a value that is almost 50-fold lower than that measured for **1**. Despite this large increase in affinity, MBNL1 can still bind to the 2-(CUG)<sub>4</sub> complex. This unexpected observation of a non-competitive inhibition mode for both **1** and **2** is the key finding of this study. Its implication for drug design is obvious. The 2-(CUG)<sub>4</sub> complex reduces the MBNL1 affinity only 6-fold relative to naked (CUG)<sub>4</sub> and 2-fold relative to that measured for the 1-(CUG)<sub>4</sub> complex. Thus, tight binding is an insufficient criterion for effective inhibition because it may shift the overall equilibrium into manifold **2** of Table 1a (i.e. low  $K_{D2}$ ) without necessarily affecting MBNL1 binding. Highly effective inhibition will result from agents that exhibit both tight and selective binding to CUG<sup>exp</sup> (low  $K_{D2}$ ) and high  $K_{D3}$  values.

Another significant result is that the apparent IC<sub>50</sub> values for the ligands depend markedly on both (CUG)<sub>4</sub> and MBNL1 concentrations (Figure 5). Thus, the CUG<sup>exp</sup> length and cellular copy number and the MBNL1 concentration will determine the drug effectiveness. Because the CUG repeat continues to expand over the patient's life, an ideal therapeutic agent should be able to inhibit binding of MBNL1 to CUG<sup>exp</sup> over a broad range of nuclear CUG<sup>exp</sup> copy number and repeat lengths. By single-molecule and bulk methods (Figure 5, Supplementary Figure S11 and Table 1), we found that, given the same concentration of MBNL1, **2** effectively inhibits (CUG)<sub>n</sub>-MBNL1 ( $n = 4$  or 12) interaction to a greater extent than **1** and over a broader range of (CUG)<sub>n</sub> concentrations, a trend that can be further improved in the future ligands.

In summary, this study revealed that inhibition of (CUG)<sub>4</sub>-MBNL1 interaction by small molecules does not occur by a simple competitive mechanism. These findings aided in the design of a more potent (CUG)<sub>4</sub>-MBNL1 inhibitor and are guiding our current optimization efforts. More generally, the increasingly frequent discovery of key roles for RNA outside its

established role in protein synthesis suggests new opportunities for RNA-targeted therapeutics. Indeed, RNA is now a validated, yet underused drug target, with less off-pathway binding. As the development of inhibitors for protein-RNA interactions becomes more widespread, we believe the inhibition model revealed in this study will need to be considered. It is likely that the dynamic and versatile structure of RNA allows it to form ternary ligand-protein-RNA complexes, although with lower stability (43). Finally, the single-molecule methodology described herein may prove to be a powerful method to unravel the inhibition mechanism for any biomacromolecular interaction.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–12, Supplementary Notes 1–5, Supplementary Methods, Supplementary References [21,43–45] and NMR spectra of compounds.

## ACKNOWLEDGEMENTS

The authors thank Yuan Fu and Kali Miller for their assistance with MBNL1 expression, purification and the synthetic procedures and Dr Tom Meek (GSK) for valuable discussions.

## FUNDING

National Institutes of Health (NIH) [R01AR058361 to S.C.Z.]; Howard Hughes Medical Institute (Early Career Scientist Award to M.S.). Funding for open access charge: NIH and University of Iowa start-up funds.

*Conflict of interest statement.* None declared.

## REFERENCES

- Wheeler, T.M., Sobczak, K., Lueck, J.D., Osborne, R.J., Lin, X., Dirksen, R.T. and Thornton, C.A. (2009) Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science*, **325**, 336–339.
- O'Rourke, J.R. and Swanson, M.S. (2009) Mechanisms of RNA-mediated disease. *J. Biol. Chem.*, **284**, 7419–7423.
- Echeverria, G.V. and Cooper, T.A. (2012) RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity. *Brain Res.*, **1462**, 100–111.
- Wheeler, T.M., Leger, A.J., Pandey, S.K., MacLeod, A.R., Nakamori, M., Cheng, S.H., Wentworth, B.M., Bennett, C.F. and Thornton, C.A. (2012) Targeting nuclear RNA for *in vivo* correction of myotonic dystrophy. *Nature*, **488**, 111–115.
- Garcia-Lopez, A., Llamusi, B., Orzaez, M., Perez-Paya, E. and Artero, R.D. (2011) *In vivo* discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. *Proc. Natl Acad. Sci. USA*, **108**, 11866–11871.
- Foff, E.P. and Mahadevan, M.S. (2011) Therapeutics development in myotonic dystrophy type 1. *Muscle Nerve*, **44**, 160–169.
- Warf, M.B., Nakamori, M., Matthys, C.M., Thornton, C.A. and Berglund, J.A. (2009) Pentamidine reverses the splicing defects associated with myotonic dystrophy. *Proc. Natl Acad. Sci. USA*, **106**, 18551–18556.

8. Ofori, L.O., Hoskins, J., Nakamori, M., Thornton, C.A. and Miller, B.L. (2012) From dynamic combinatorial 'hit' to lead: *in vitro* and *in vivo* activity of compounds targeting the pathogenic RNAs that cause myotonic dystrophy. *Nucleic Acids Res.*, **40**, 6380–6390.
9. Parkesh, R., Childs-Disney, J.L., Nakamori, M., Kumar, A., Wang, E., Wang, T., Hoskins, J., Tran, T., Housman, D., Thornton, C.A. *et al.* (2012) Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching. *J. Am. Chem. Soc.*, **134**, 4731–4742.
10. Pushechnikov, A., Lee, M.M., Childs-Disney, J.L., Sobczak, K., French, J.M., Thornton, C.A. and Disney, M.D. (2009) Rational design of ligands targeting triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3. *J. Am. Chem. Soc.*, **131**, 9767–9779.
11. Holdgate, G.A. and Gill, A.L. (2011) Kinetic efficiency: the missing metric for enhancing compound quality? *Drug Discov. Today*, **16**, 910–913.
12. Blow, N. (2009) Proteins and proteomics: life on the surface. *Nat. Methods*, **6**, 389–393.
13. Navratilova, I. and Myska, D.G. (2006) Investigating biomolecular interactions and binding properties using SPR biosensors. *Surf. Plasmon Reson. Based Sens.*, **4**, 155–176.
14. Elenko, M.P., Szostak, J.W. and van Oijen, A.M. (2009) Single-molecule imaging of an *in vitro*-evolved RNA aptamer reveals homogeneous ligand binding kinetics. *J. Am. Chem. Soc.*, **131**, 9866–9867.
15. Moffitt, J.R., Chemla, Y.R. and Bustamante, C. (2010) Methods in statistical kinetics. *Methods Enzymol.*, **475**, 221–257.
16. Karunatilaka, K.S., Solem, A., Pyle, A.M. and Rueda, D. (2010) Single-molecule analysis of Mss116-mediated group II intron folding. *Nature*, **467**, 935–939.
17. Yuan, Y., Compton, S.A., Sobczak, K., Stenberg, M.G., Thornton, C.A., Griffith, J.D. and Swanson, M.S. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.*, **35**, 5474–5486.
18. Galletto, R., Amitani, I., Baskin, R.J. and Kowalczykowski, S.C. (2006) Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature*, **443**, 875–878.
19. Kim, Y.T., Tabor, S., Churchich, J.E. and Richardson, C.C. (1992) Interactions of gene 2.5 protein and DNA polymerase of bacteriophage T7. *J. Biol. Chem.*, **267**, 15032–15040.
20. Honda, M., Park, J., Pugh, R.A., Ha, T. and Spies, M. (2009) Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. *Mol. Cell*, **35**, 694–703.
21. Arambula, J.F., Ramisetty, S.R., Baranger, A.M. and Zimmerman, S.C. (2009) A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding. *Proc. Natl Acad. Sci. USA*, **106**, 16068–16073.
22. Mammen, M., Choi, S.K. and Whitesides, G.M. (1998) Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.*, **37**, 2755–2794.
23. Childs-Disney, J.L., Hoskins, J., Rzuczek, S.G., Thornton, C.A. and Disney, M.D. (2012) Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem. Biol.*, **7**, 856–862.
24. Wong, C.H., Richardson, S.L., Ho, Y.J., Lucas, A.M., Tuccinardi, T., Baranger, A.M. and Zimmerman, S.C. (2012) Investigating the binding mode of an inhibitor of the MBNL1. RNA complex in myotonic dystrophy Type 1 (DM1) leads to the unexpected discovery of a DNA-selective binder. *ChemBioChem*, **13**, 2505–2509.
25. Zimmerman, S.C., Lamberson, C.R., Cory, M. and Fairley, T.A. (1989) Topologically Constrained bifunctional intercalators—DNA intercalation by a macrocyclic bisacridine. *J. Am. Chem. Soc.*, **111**, 6805–6809.
26. Fechter, E.J., Olenyuk, B. and Dervan, P.B. (2004) Design of a sequence-specific DNA bisintercalator. *Angew. Chem. Int. Ed. Engl.*, **43**, 3591–3594.
27. Holman, G.G., Zewail-Foote, M., Smith, A.R., Johnson, K.A. and Iverson, B.L. (2011) A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant. *Nat. Chem.*, **3**, 875–881.
28. Van den Wildenberg, S.M., Prevo, B. and Peterman, E.J. (2011) Single molecule analysis. *Methods*, **783**, 81–99.
29. Reck-Peterson, S.L., Derr, N.D. and Stuurman, N. (2010) Imaging single molecules using total internal reflection fluorescence microscopy (TIRFM). *Cold Spring Harb. Protoc.*, **2010**, pdb.top1173.
30. Warf, M.B. and Berglund, J.A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA*, **13**, 2238–2251.
31. 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.
32. Qin, F., Auerbach, A. and Sachs, F. (1996) Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. *Biophys. J.*, **70**, 264–280.
33. 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.
34. Feng, B.Y. and Shoichet, B.K. (2006) A detergent-based assay for the detection of promiscuous inhibitors. *Nat. Protoc.*, **1**, 550–553.
35. Vamosi, G., Gohlke, C. and Clegg, R.M. (1996) Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: multiple conformers of single-stranded and double-stranded dye-DNA complexes. *Biophys. J.*, **71**, 972–994.
36. Qu, P., Chen, X.D., Zhou, X.X., Li, X. and Zhao, X.S. (2009) Fluorescence quenching of TMR by guanosine in oligonucleotides. *Sci. China, Ser. B*, **52**, 1653–1659.
37. Xu, Y. and Komiyama, M. (2012) Structure, function and targeting of human telomere RNA. *Methods*, **57**, 100–105.
38. Vicens, Q. and Westhof, E. (2003) RNA as a drug target: the case of aminoglycosides. *ChemBioChem*, **4**, 1018–1023.
39. Vicens, Q. (2009) RNA's coming of age as a drug target. *J. Incl. Phenom. Macro.*, **65**, 171–188.
40. Cooper, T.A., Wan, L. and Dreyfuss, G. (2009) RNA and disease. *Cell*, **136**, 777–793.
41. 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.
42. Laurent, F.X., Sureau, A., Klein, A.F., Trouslard, F., Gasnier, E., Furling, D. and Marie, J. (2012) New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats. *Nucleic Acids Res.*, **40**, 3159–3171.
43. Stelzer, A.C., Frank, A.T., Kratz, J.D., Swanson, M.D., Gonzalez-Hernandez, M.J., Lee, J., Andricioaei, I., Markovitz, D.M. and Al-Hashimi, H.M. (2011) Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol.*, **7**, 553–559.