Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules

Jessica L. Childs-Disney, Ewa Stepniak-Konieczna, Tuan Tran, Ilyas Yildirim, HaJeung Park, Catherine Z. Chen, Jason Hoskins, Noel Southall, Juan J. Marugan, Samarjit Patnaik, Wei Zheng, Chris P. Austin, George C. Schatz, Krzysztof Sobczak, Charles A. Thornton & Matthew D. Disney

The ability to control pre-mRNA splicing with small molecules could facilitate the development of therapeutics or cell-based circuits that control gene function. Myotonic dystrophy type 1 is caused by the dysregulation of alternative pre-mRNA splicing due to sequestration of muscleblind-like 1 protein (MBNL1) by expanded, non-coding (r(CUG)) repeats (r(CUG)exp). Here we report two small molecules that induce or ameliorate alternative splicing dysregulation. A thiophene-containing small molecule (1) inhibits the interaction of MBNL1 with its natural pre-mRNA substrates. Compound (2), a substituted naphthyridine, binds r(CUG)exp and displaces MBNL1. Structural models show that 1 binds MBNL1 in the Zn-finger domain and that 2 interacts with UU loops in r(CUG)exp. This study provides a structural framework for small molecules that target MBNL1 by mimicking r(CUG)exp and shows that targeting MBNL1 causes dysregulation of alternative splicing, suggesting that MBNL1 is thus not a suitable therapeutic target for the treatment of myotonic dystrophy type 1.
Pre-mRNA splicing defects cause a wide variety of diseases including β-thalassaemia, inherited breast cancer, fragile X-associated tremor ataxia syndrome and myotonic dystrophy types 1 and 2 (DM1 and DM2, respectively)\(^1\)\(^-\)\(^6\). In the former two cases, pre-mRNA splicing defects are caused by single-nucleotide polymorphisms within intronic regions of their mRNAs, yielding cryptic alternative splice sites. In the latter three cases, pre-mRNA splicing defects are caused by sequestration of proteins that regulate alternative pre-mRNA splicing\(^5\)\(^,\)\(^7\).

DM1 is caused by sequestration of muscleblind-like 1 protein (MBNL1) by expanded r(CUG) repeats (r(CUG)\(^\text{exp}\)) present in the 3’-untranslated region (UTR) of the dystrophia myotonica protein kinase (DMPK) mRNA\(^8\). r(CUG)\(^\text{exp}\) folds into an extended hairpin that displays multiple copies of a 5’/CUG/3’GUC motif, which are high-affinity-binding sites for MBNL1. Studies in patient-derived cell lines and in mouse models have shown that pre-mRNA splicing events are dysregulated in DM1, including alternative splicing of the insulin receptor (IR)\(^9\), cardiac troponin T (cTNT)\(^10\) and the muscle-specific chloride ion channel (CLCN1)\(^11\)\(^,\)\(^12\). These defects explain symptoms suffered by DM1 patients such as insulin insensitivity (IR splicing), myotonia (CLCN1 splicing) and cardiac defects (cTNT splicing). Moreover, translational defects of the DMPK mRNA and the presence of nuclear foci are observed in DM patients\(^13\)\(^,\)\(^14\). Both defects are due to the binding of r(CUG)\(^\text{exp}\) to various proteins, leading to poor nucleocytoplasmic transport\(^13\)\(^,\)\(^15\).

To identify small molecules that disrupt the r(CUG)\(^\text{exp}\)–MBNL1 complex, we previously developed an assay amendable for quantitative high-throughput screening\(^16\). The small molecules that comprise the Molecular Libraries Small Molecule Repository (>300,000 members) were screened, and the results have been published online at PubChem (http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=2675&loc=ea_ras) and in Chen et al.\(^16\).

The high-throughput screen identifies compounds that inhibit the r(CUG)–MBNL1 interaction by binding the RNA or the protein. Thus, a small molecule that binds r(CUG)\(^\text{exp}\) and displaces MBNL1 could theoretically improve DM1-associated pre-mRNA splicing defects (Fig. 1a). However, small molecules that bind MBNL1 and disrupt the r(CUG)\(^\text{exp}\)–MBNL1 interaction may inactivate the protein and induce DM1-like pre-mRNA splicing defects (Fig. 1a). Importantly, the screen could yield chemotypes that reverse DM1-associated pre-mRNA splicing defects as a first step towards the development of therapeutics for this devastating disease.

Herein, we describe the results from our investigation of two hits from our quantitative high-throughput screening: one that reverses DM1 pre-mRNA spliceopathy and another that induces DM1-like pre-mRNA spliceopathy. These compounds can serve as useful tools to study the effect of DM1 splicing on disease and to understand how to control pre-mRNA splicing or translation with cell-permeable small molecules.

**Results**

**High-throughput screening for lead compounds against DM1.**

Previously, our team reported a screen for inhibitors of the toxic RNA–protein interaction that causes myotonic dystrophy type 1 (DM1) (ref.16). The screen was designed to identify ligands that inhibit formation of a r(CUG)–MBNL1 complex that causes
DM1. The assay utilizes biotinylated r(CUG)_{12} and MBNL1-His$_6$. After formation of the RNA–protein complex, a ligand of interest is added followed by the addition of fluorescently labelled streptavidin (recognizes biotinylated r(CUG)$_{12}$) and a fluorescently labelled anti-His$_6$ antibody (recognizes MBNL1). When the r(CUG)$_{12}$–MBNL1 complex forms, the two fluorophores are within close enough proximity to form a fluorescence resonance energy transfer (FRET) pair, and time resolved (TR)-FRET can be measured. If, however, a small molecule inhibits r(CUG)$_{12}$–MBNL1 complex formation, no TR-FRET will be observed. Four compounds identified from the NIH Chemical Genomics Center (NCGC) screen were further investigated: (E)-4-phenyl-2-(3-(thiophen-2-yl)acrylamido)thiophene-3-carboxylic acid (1), 1,8-diamino-3,6-di(pyrrolidin-1-yl)-2,7-naphthyridine-4-carbonitrile (2), 6-hydroxy-3a,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4-carboxylic acid (3), and 3-(((4-fluorophenyl)amino)isoquinolin-1-ol (4) (Fig. 1b). These compounds inhibited formation of the r(CUG)$_{12}$–MBNL1 complex with IC$_{50}$s that range from 2–242 μM.

**Improvement of DM1-associated translational defects.** As an initial test for bioactivity, a previously described cellular model system that mimics the DMPK translational defects observed in DM1 was used (Fig. 2a, ref. 17). In this model, firefly luciferase mRNA containing 800 r(CUG) repeats (r(CUG)$_{800}$) in the 3′-UTR is stably expressed in C2C12 cells. Formation of the r(CUG)$_{800}$–MBNL1 complex retains the firefly luciferase mRNA in the nucleus, disallowing nucleocytoplasmic transport. If a small molecule enters the nucleus and displaces MBNL1 from r(CUG)$_{800}$, then the luciferase mRNA may be translocated to the cytoplasm and translated, resulting in increased luciferase activity. An increase in luciferase activity should be observed if compounds bind to r(CUG)$_{800}$ or MBNL1, as both inhibit the formation of the r(CUG)$_{800}$–MBNL1 complex.

All four compounds shown in Fig. 1b were tested in this system at 20 μM (Fig. 2a). Compounds 1 and 2 stimulated luciferase production most effectively, with increases in luciferase activity of ~700% and ~1,600%, respectively. (Please note that no change in luciferase activity is equal to 0%) Compounds 3 and 4 also increase luciferase activity but to a lesser extent (ca. 125%). Control experiments were performed using an analogous cellular model system that stably expresses firefly luciferase mRNA lacking r(CUG)$_{800}$. A slight increase in luciferase activity was observed for compounds 1, 2 and 3, although the increase is small compared to the increase observed in the presence of r(CUG)$_{800}$. In the case of compound 4, however, we observed similar increases in luciferase activity in cell lines with and without r(CUG)$_{800}$, indicating a nonspecific effect. Therefore, 4 was no longer pursued as a lead compound. Owing to the enhanced biological activity of 1 and 2, these two compounds were further investigated.

**Affinity and selectivity of 1 and 2 for r(CUG)$_{800}$ and MBNL1.** The binding of 1 and 2 to a mimic r(CUG)$_{800}$, r(CUG)$_{12}$, and MBNL1 were determined by competition dialysis$^{18}$ and fluorescence anisotropy, respectively. Compound 2 nonspecifically adheres to dialysis membranes, and thus is incompatible with competition dialysis.

Competition dialysis shows that 1 binds specifically to MBNL1 over both r(CUG)$_{12}$ and a protein control, bovine serum albumin (BSA) (Fig. 2b). Almost no detectable partitioning of 1 was observed into dialysis tubes that contained BSA or r(CUG)$_{12}$. In contrast, a large amount of partitioning was observed in tubes that contained MBNL1, affording a K$_D$ of 470 ± 20 nM. These data suggest that 1 inhibits r(CUG)$_{800}$–MBNL1 complex formation by binding to MBNL1 (Fig. 1a).

Fluorescence anisotropy measurements with 2 show that it binds specifically to r(CUG)$_{12}$ with a K$_D$ of 125 nM and a stoichiometry of ~6 small molecules per RNA (Table 2). No change in anisotropy was observed when up to two times the concentration of MBNL1 (650 nM) and ~35 times the concentration of BSA (12.5 μM) were added to 2 (325 nM), indicating weak protein binding. The affinity of 2 for another repeating RNA, r(CAG)$_{12}$ (Supplementary Fig. S1), and a fully paired RNA (AU, Supplementary Fig. S1) were also determined. Compound 2 binds r(CAG)$_{12}$ with a dissociation constant of 1,400 nM and a 1:1 stoichiometry while no binding was observed to the AU RNA (K$_D$ >> 25,000 nM) (Table 2).

The binding of 2 to an RNA containing a single 1 × 1 nucleotide UU internal loop (1 × 1 UU, Supplementary Fig. S1) was also measured. The compound binds 1 × 1 UU with a 1:1 stoichiometry and a K$_D$ of 500 nM. There is an ~4-fold difference in affinity of 2 for r(CUG)$_{12}$ and 1 × 1 UU. These data suggest that the ligand binds with positive cooperativity to longer RNAs. Such properties are favourable for selectively recognizing r(CUG)$_{800}$ in vivo$^{19}$ as many other genomic RNA targets contain single 1 × 1 nucleotide UU internal loop, notably the A-site in human ribosomes.

**Selectivity of 1 and 2 for inhibiting RNA–MBNL1 complexes.** As 1 binds to MBNL1, its potency for disrupting RNA–MBNL1 interactions should be independent of the RNA’s sequence. In contrast, 2 (binds specifically to r(CUG)$_{800}$) should more potently inhibit r(CUG)$_{12}$–MBNL1 complex formation as compared with complexes formed with other RNAs that bind MBNL1 (r(CAG)$_{12}$, r(CCUG)$_{12}$ and r(CGG)$_{12}$, for example refs 20, 21). (r(CAG)$_{800}$–MBNL1 and r(CUG)$_{800}$–MBNL1 complexes cause pre-mRNA splicing defects in Huntington disease patient-derived cell lines and in a myotonic dystrophy type 2 cellular model,

| Compound | IC$_{50}$ (μM) | Complex | IC$_{50}$ of 1; selectivity for r(CUG)$_{800}$-MBNL1 | IC$_{50}$ of 2; selectivity for r(CUG)$_{800}$-MBNL1 |
|----------|---------------|---------|---------------------------------|---------------------------------|
| 1        | 52 ± 12       | r(CUG)$_{12}$-MBNL1 | 54 ± 2; -                        | 2 ± 0.4; -                      |
| 2        | 2 ± 0.4       | r(CAG)$_{12}$-MBNL1 | 66 ± 13; 1.2                     | 37 ± 5; 20                     |
| 3        | 242 ± 1       | r(CGG)$_{12}$-MBNL1 | 54 ± 4; 1.0                      | 45 ± 5; 24                     |
| 4        | 12 ± 1        | r(CCUG)$_{12}$-MBNL1 | 75 ± 6; 1.4                      | 29 ± 3; 15                     |

*Secondary structures of the RNAs are provided in Supplementary Fig. S1.
The RNA and small molecule were pre-incubated followed by addition of MBNL1.
For studies with compound 1, the ligand was pre-incubated with the RNA followed by addition of MBNL1.

Table 1 | Potencies and selectivities of compounds 1–4.*
Therefore, we studied the potencies of compounds 1 and 2 for preventing RNA–MBNL1 complex formation (Table 1). In these experiments, 1 was pre-incubated with MBNL1, followed by addition of RNA while 2 was pre-incubated with the RNA of interest followed by addition of MBNL1.

As expected, our results indicate that compound 1 is a general inhibitor of RNA–MBNL1 complexes with an average IC_{50} of 62 ± 10 µM (Table 1). In contrast, the IC_{50}'s for 2 vary from 2 µM (r(CUG)_{12}) to 45 µM (r(CGG)_{12}). Compound 2 is at least 10-fold more potent for inhibiting the r(CUG)_{12}–MBNL1 interaction than for inhibiting r(CAG)_{12}+, r(CCG)_{12} and r(CGG)_{12}–MBNL1 interactions. These data are in good agreement with the results of the binding measurements summarized in Table 2 and Fig. 2b and further support the hypothesis that the potencies of 1 and 2 are derived from binding MBNL1 and r(CUG)^{exp}, respectively.

The contrasting modes of in vitro inhibition and binding data support the results observed in the luciferase assay. The luciferase model system cannot discriminate between compounds that disrupt r(CUG)^{exp}–protein complexes by binding r(CUG)^{exp} from those that disrupt the complexes by binding to proteins. If a small molecule binds to either protein or RNA, then the complex will be disrupted, allowing the mRNA to be translocated to the cytoplasm and translated. Our previously described competition dialysis, anisotropy and potency data suggest that compounds 1 and 2 could have very different effects on pre-mRNA splicing. Compound 1, which binds MBNL1, could prevent MBNL1 from binding to its natural pre-mRNA substrates, inducing DM1-like splicing dysregulation in non-DM1 cells (that is, like that observed when MBNL1 is sequestered by r(CUG)^{exp} (Fig. 1a). Compound 2, which binds r(CUG)^{exp}, may displace MBNL1, freeing it to regulate pre-mRNA splicing as though r(CUG)^{exp} had been removed (Fig. 1a).

Compound 1 induces splicing defects. MBNL1 regulates the alternative splicing of many pre-mRNAs^{7,9–12,23,24}. Numerous splicing defects are observed in DM1 patients including the IR and cTNT pre-mRNAs^{9,10,23,24}. It has been shown that co-transfection of HeLa cells with a DM1 mini-gene that expresses 960 interrupted r(CUG) repeats causes dysregulation of mini-genes regulated by MBNL1; in particular splicing defects in the IR and cTNT pre-mRNAs have been observed by reverse transcription PCR (RT–PCR)^{10,20,25,26}. The percentage of exon 11 included in the IR mini-gene transcripts decreases from ≈97% in the absence of r(CUG)^{exp} to ≈75% in the presence of r(CUG)^{exp} (Fig. 3). Regarding the alternative splicing of the cTNT pre-mRNA, the percentage of exon 5 inclusion increases from ≈75% in the absence of r(CUG)^{exp} to ≈90% in the presence of r(CUG)^{exp} (Fig. 3).

**Table 2 | Affinity of 2 for various biomolecules.**

| Biomolecule | \(K_d\) (nM); stoichiometry | Biomolecule | \(K_d\) (nM); stoichiometry |
|-------------|-----------------------------|-------------|-----------------------------|
| MBNL1       | >>2,000                      | r(CAG)_{12} | 1400 ± 0.2; 1.2 ± 0.1 |
| BSA         | >>2,000                      | 1 × 1UU     | 530 ± 0.1; 1.0 ± 0.1 |
| r(CUG)_{12} | 125 ± 0.052; 6 ± 0.4        | AU          | >>25,000                    |

**Figure 2 | Compounds 1–4 improve translational defects caused by r(CUG)^{exp}.** Improvement of translational could occur if a compound binds to r(CUG)^{exp} (compound 2) or MBNL1 (compound 1), as both modes of action disrupt the r(CUG)^{exp}–MBNL1 complex and could allow for more efficient nucleocytoplasmic export. (a, top) A schematic of the luciferase reporter system that models the DM1 translational defect. The presence of r(CUG)^{exp} in the 3’-UTR of firefly luciferase reduces nucleocytoplasmic transport and thereby suppresses luciferase expression. However, if a small molecule binds r(CUG)^{exp} and displaces proteins, nucleocytoplasmic transport is improved and luciferase activity increases. Likewise, if a small molecule binds proteins and displaces them from r(CUG)^{exp}, nucleocytoplasmic transport is also improved and luciferase activity increases. (a, bottom) Effects on luciferase activity when cells are dosed with 20 µM compound of interest. ‘+’ r(CUG)^{exp} indicates that the cell line expresses luciferase with r(CUG)^{exp} in the 3’-UTR, ‘−’r(CUG)^{exp} indicates that the cell line expresses luciferase without r(CUG)^{exp}. Results are expressed as the percentage increase of luciferase activity relative to untreated cells, where a value of ‘0’ denotes no change in activity. Experiments were completed in triplicate. Values shown in the plot are the averages of those experiments, and the errors reported are the s.d. values. (b) Competition dialysis data for binding of 1 to BSA, MBNL1 and r(CUG)_{12}. The data clearly show that the preferred target is MBNL1 as there is no measureable binding to r(CUG)_{12} or to BSA. Experiments were completed in duplicate. The values reported are the averages of those experiments, and the errors reported are the s.d. values.
Treatment of HeLa cells that do not express r(CUG)$^{\text{exp}}$ with compound 1 causes a shift in the alternative splicing patterns of the IR and cTNT mini-genes towards a DM1-like phenotype (Fig. 3). Dysregulation of IR pre-mRNA splicing is observed when cells are treated with 500 and 1,000 μM of 1. Likewise, alternative splicing of the cTNT mini-gene is shifted towards DM-like phenotype when cells are dosed with compound 1, and a dose-response is observed from 50 to 560 μM. Statistically significant changes in cTNT alternative splicing are only observed at 560 μM when compared with cells that do not express r(CUG)$^{\text{exp}}$.

In order to determine if compound 1 generally affects the regulation of alternative splicing, additional experiments were completed with three mini-genes, the alternative splicing of which are not regulated by MBNL1 (Supplementary Fig. S2a-c) (refs 5,26–28). These include pleckstrin homology domain containing, family H member 2 (PLEKHH2) ref. 26, survival of motor neuron 2 (SMN2) (ref. 28) and B-cell lymphoma x (Bcl-x)$^{\text{exp}}$, and calcium/calmodulin-dependent protein kinase II gamma (CAMK2G). Splicing abnormalities (towards the DM-like phenotype) were observed when cells were treated with 250 or 500 μM of 1 for all five MBNL1-dependent exons (Fig. 4a,b). The strongest effects were observed for MBNL1 and MBNL2 exons, in which splicing dysregulation is the same as in DM1 affected fibroblasts. The DM1-like splicing shift in NCOR2, NFIX and CAMK2G was also evident upon 500 μM 1, although more subtle than in the case of MBNL1s (Fig. 4a,b).

Notably, efficient short interfering RNA (siRNA) knockdown of MBNL1 (by ~70% relative to untreated controls) in normal fibroblasts triggered only slightly more pronounced DM1-like splicing shift in MBNL1-dependent exons than treatment with 1 or than DM1 affected fibroblasts (Fig. 4b, Supplementary Fig. S3A). This suggests that 500 μM of 1 is not a saturating concentration required to induce a maximum possible effect on the splicing activity of MBNL1. To explore whether compound 1-induced DM-like splicing alterations can be further exacerbated, we next knocked-down MBNL1 in normal human fibroblasts using siRNA and then, 24 h post siMBNL1

Figure 3 | Compound 1 induces DM1-like splicing defects in HeLa cells. HeLa cells were co-transfected with the mini-gene of interest and a plasmid that expresses r(CUG)$^{\text{exp}}$ or an empty vector. Varying concentrations of 1 were added in growth medium post transfection, and total RNA was harvested 16–20 h later. The percentage of each splicing isoform was determined by RT–PCR using a radioactively labelled forward PCR primer. (a, top) Schematic of the pre-mRNA splicing isoforms observed for the IR mini-gene. (a, middle) representative gel image of IR alternative splicing in the presence of varying concentrations of 1. (a, bottom) Quantification of exon inclusion for IR alternative splicing in the presence or absence of r(CUG)$^{\text{exp}}$ and 1 (n ≥ 3). Error bars are the s.d. values in the measurements. (*) indicates P ≤ 0.05 and *** indicates P ≤ 0.001 as determined by a two-tailed Student’s t-test.

(b, top) Schematic of the pre-mRNA splicing isoforms observed for the cTNT mini-gene. (b, middle) representative gel image of cTNT alternative splicing in the presence of varying concentrations of 1. (b, bottom) Quantification of exon inclusion for cTNT alternative splicing in the presence or absence of r(CUG)$^{\text{exp}}$ and 1 (n ≥ 3). Error bars are the s.d. values in the measurements. * indicates P ≤ 0.05 and *** indicates P ≤ 0.01 as determined by a two-tailed Student’s t-test.

| Exon 4 | Exon 5 | Exon 6 |
|-------|-------|-------|
| [1] (μM) | [1] (μM) | [1] (μM) |
| 0 | + | + |
| 50 | – | + |
| 100 | – | + |
| 560 | – | + |

| Exon 10 | Exon 11 | Exon 12 |
|---------|---------|---------|
| ~25% | ~75% | ~0% |

| Exon 4 | Exon 5 | Exon 6 |
|-------|-------|-------|
| [1] (μM) | [1] (μM) | [1] (μM) |
| 0 | + | + |
| 50 | – | + |
| 100 | – | + |
| 560 | – | + |

**Figure 3 | Compound 1 induces DM1-like splicing defects in HeLa cells.** HeLa cells were co-transfected with the mini-gene of interest and a plasmid that expresses r(CUG)$^{\text{exp}}$ or an empty vector. Varying concentrations of 1 were added in growth medium post transfection, and total RNA was harvested 16–20 h later. The percentage of each splicing isoform was determined by RT–PCR using a radioactively labelled forward PCR primer. (a, top) Schematic of the pre-mRNA splicing isoforms observed for the IR mini-gene. (a, middle) representative gel image of IR alternative splicing in the presence of varying concentrations of 1. (a, bottom) Quantification of exon inclusion for IR alternative splicing in the presence or absence of r(CUG)$^{\text{exp}}$ and 1 (n ≥ 3). Error bars are the s.d. values in the measurements. (*) indicates P ≤ 0.05 and *** indicates P ≤ 0.001 as determined by a two-tailed Student’s t-test.

(b, top) Schematic of the pre-mRNA splicing isoforms observed for the cTNT mini-gene. (b, middle) representative gel image of cTNT alternative splicing in the presence of varying concentrations of 1. (b, bottom) Quantification of exon inclusion for cTNT alternative splicing in the presence or absence of r(CUG)$^{\text{exp}}$ and 1 (n ≥ 3). Error bars are the s.d. values in the measurements. * indicates P ≤ 0.05 and *** indicates P ≤ 0.01 as determined by a two-tailed Student’s t-test.
transfection, treated the cells with 1. Combined treatment with siMBNL1 (which reduced the level of MBNL1 mRNA by \( \sim 70\% \) relative to untreated controls) and compound 1 allowed us to test whether reducing the pool of available MBNL1 exerts any effects on the splicing dysregulation induced by 1 and to assess whether 1 induces DM1-like splicing shift via MBNL1-dependent mechanism. Importantly, compound 1-induced dysregulation of pre-mRNA alternative splicing of MBNL1-regulated exons (MBNL1, MBNL2, NCOR2 and NFIX) was significantly exacerbated by MBNL1 knockdown, strongly suggesting that the effects of 1 are indeed MBNL1-dependent (Fig. 5). For all analysed MBNL1-regulated exons, splicing changes induced by highest concentration of 1 (that is, 500 \( \mu \text{M} \)) combined with an siRNA against MBNL1 were much more pronounced than those induced...
by 1 alone (Figs 4 and 5). We hypothesize that knocking-down MBNL1 with siRNA makes it possible to 'saturate' the remaining pool of MBNL1 with compound 1, and thus augments the effect of 1 on the alternative splicing dysregulation of MBNL1-regulated exons.

In agreement with our hypothesis, the combined effect of siMBNL1 and 500 µM 1 in normal fibroblast resulted in a roughly little as 125 µM of 2 to cells that express r(CUG)exp hairpin structure formation (and thus the amount of active MBNL1) causes phenotypic variations31. Taken together, our data strongly suggest that 1 specifically affects the function of MBNL1 protein in vivo.

Compound 2 improves splicing defects. In contrast, compound 2 improves DM1-associated pre-mRNA splicing defects in HeLa cells transfected with the DM1 mini-gene (Fig. 6). Addition of as little as 125 µM of 2 to cells that express r(CUG)exp restores IR pre-mRNA splicing back to levels those observed in the absence of r(CUG)exp. Similarly, improvement in CTNT alternative splicing is also observed. When cells are treated with 300 µM of 2, pre-mRNA splicing levels are statistically indistinguishable from those observed in the absence of r(CUG)exp and thus are statistically different from cells that express r(CUG)exp.

Fluorescence in situ hybridization was employed to determine if 2 decreases the number or intensity of nuclear foci. Although foci are still present, they are much more dispersed when cells are treated with 2, and it appears that there is less MBNL1 within the foci as determined by immunostaining (MB1a antibody (Supplementary Fig. S9 and Supplementary Note 1) (ref. 32).
To test the specificity of splicing modulation by compound 2, we examined its effect on exons not regulated by MBNL1. HeLa cells were co-transfected with a DM1 mini-gene and splicing reporter constructs for PLEKH2, SMN2 and Bcl-x, the alternative splicing of which have been shown previously not to be regulated by MBNL1 (refs 5,26–28). Treatment with 500 μM 2 had no effect on the alternative splicing of PLEKH2, SMN2 or Bcl-x mini-genes (Supplementary Fig. S2a–c), regardless of whether cells expressed r(CUG)exp. The alternative splicing of endogenous mRNAs not regulated by MBNL1, CAMKK2 (calcium/calmodulin-dependent protein kinase kinase 2) and TTC8 (tetratricopeptide repeat domain 8) (ref. 26), were also not affected (Supplementary Fig. S2d,e). Thus, the effect of compound 2 appears to be specific to its ability to bind r(CUG)exp and displace MBNL1.

Next, we determined if 2 also improves defects associated with MBNL2 sequestration by r(CUG)exp as they contribute to toxicity in the central nervous systems/brains of DM1 patients29. For this purpose, the HEK cell line was used as it expresses MBNL proteins at very low levels33. HEK cells were co-transfected with the DM1 and cTNT mini-genes and increasing amounts of a plasmid encoding MBNL1 or MBNL2. As expected, increased expression levels of MBNLs suppress cTNT exon 5 inclusion10,23, confirming that cTNT alternative splicing is regulated by both proteins in our model system (Supplementary Fig. S10). Dysregulation of cTNT alternative splicing is improved when HEK cells that express r(CUG)exp and MBNL1 or MBNL2 are treated with 300 μM of 2, although the effect is somewhat less for splicing regulated by MBNL2 (Supplementary Fig. S11). (Dysregulation of splicing is not improved when cells are treated with 75 μM or 150 μM of 2.) Taken together, 2 inhibits r(CUG)exp–MBNL1 and r(CUG)exp–MBNL2 complexes.

Molecular modelling and docking studies. A set of docking studies support our experimental evidence that 1 binds to MBNL1 and 2 binds to r(CUG) repeats. Figure 7a shows the results of the molecular modelling of 1 and MBNL1. The model of the complex shows extensive interactions between MBNL1 and 1 that occurs in the RNA-binding pocket. The ethenyl group and the phenyl and thiofuran rings of 1 are stabilized by extensive cation–π interactions with Arg186, Arg195 and Arg201. The cation–π interaction between 1 and Arg195 is reminiscent of the interaction between MBNL1 and RNA34. In addition, compound 1’s carbonyl group forms a bifurcated hydrogen bond with backbone amino groups of Arg186 and Glu187, as well as a hydrogen bond with side chain of Arg186. Noticeably, the carbonyl oxygen of 1 is within the coordination distance of a bound Zn ion. This suggests that the oxygen provides the fifth coordination site, completing the trigonal bipyramidal geometry of the Zn ion.

We previously reported the structure of r(CCGCUGCGG)2, which was determined by NMR spectroscopy35. These studies showed that the UU mismatch stacks within the helix, maintaining an overall A-form geometry. The UU mismatch is dynamic, however, and can interconvert between structures that contain 0, 1 or 2 hydrogen bonds35. This dynamic nature could provide a means for small molecule binding. Docking studies of 2 and a r(CUG) repeat were completed using the structure of r(CCGCUGCGG)2. (Because of the dynamic nature of the RNA, a new method to simulate the binding of 2 to r(CUG) repeats was developed (Supplementary Methods).) The lowest free energy conformation of the 2–r(CUG) complex is shown in Fig. 7b. The complex is stabilized by stacking of 2 on the 5′ and 3′ closing GC pairs (shown in red and blue, respectively; Fig. 7b) and formation of three hydrogen bonds with the uridine bases. Both of 2’s amino groups hydrogen bond with one uridine base: one hydrogen bond with the uridine’s N1 while the other hydrogen bonds with a non-bridging oxygen. Compound 2’s nitrile N hydrogen bonds with another U’s H2'.

Non-canonical 1 × 1 UU base pairs have weaker hydrogen bonds compared with Watson–Crick base pairs, which makes them more sensitive to the environment such as in the presence of 2. The pathway for the binding mode shown in Fig. 7b is via the minor groove side as no major groove side binding was seen in the molecular dynamics simulation trajectories.
Initial contact between 2 and the UU mismatch occurs at the minor groove side where 2’s amino groups interact with the free carboxyl group of the uridine bases. While compound 2 tries to stack within the helix, the weak hydrogen bonds of the UU mismatch break, resulting in one of the uridines unstacking from the helix. Later, the second uridine unstacks from the helix and yields the final binding mode where 2 is fully stacked between the flanking GC base pairs (Fig. 7b). The simulation indicates that the complex is dynamic (Supplementary Figs S13 and S14) and that stacking is the predominant interaction driving complex formation.

Discussion

Only very few compounds have been previously reported to target the protein in an RNA–protein complex. A notable example includes the development of T5626448, a small molecule that inhibits the formation of the toll-like receptor 3–dsRNA complex. This interaction activates toll-like receptor 3 and inhibits the formation of the toll-like receptor 3–dsRNA complex. This interaction activates toll-like receptor 3 and inhibits the formation of the toll-like receptor 3–dsRNA complex.

Compound 2 is more selective than previously reported small molecules that bind to expanded repeats. It binds with ∼11-fold selectivity for an RNA that displays 12 5’CGU/3’GUC motifs when compared with an RNA that displays 12 5’CAG/3’GAC and >200-fold selectivity over a fully paired RNA (Table 2). A bis-benzimidazole that also improves DM1-associated defects shows at best threefold selectivity for RNAs with a single 5’CGU/3’GUC (ref. 37). A 4’,6-diamidino-2-phenylindole derivative binds to r(CAG)exp and improves associated pre-mRNA splicing defects. This compound has modest specificity for the 5’CAG/3’GAC motif displayed by r(CAG)exp, ranging from ∼2.5-fold selectivity over RNAs that display other 1 × 1 nucleotide mismatches and >10-fold selectivity over fully paired RNAs.

Compounds 1 and 2 have utility as probes and tools. The ability to reverse or induce a disease-like state chemically can enable studies to determine the repertoire of pre-mRNAs whose splicing function is dysregulated in DM1 or other diseases by using RNA-seq. Studies could also be completed in animal models of DM1, such as in Drosophila, to determine if 1 can induce DM1-like pre-mRNA splicing defects. The compound can then be withdrawn to measure recovery back to wild type. Lastly, these compounds could have a use in the development of artificial gene circuits. Traditionally, such circuits have been controlled through aptamer-small molecule interactions in which the aptamer has been selected or excised from a riboswitch.
necessary to ensure complete dissociation of 1 from the RNA or protein, as binding can affect its spectroscopic properties. The total concentration of 1 (C) within the dialysis unit was determined spectrophotometrically using an appropriate absorbance wavelength and extinction coefficient. A dilution factor was applied to account for the addition of SDS.

The free concentration of 1 (C) was determined from an aliquot of the dialysate solution, which did not vary appreciably from the initial 2.5 μM. The bound ligand concentration (C₅₀) was then determined using (1):

\[ C_{50} = C - C_{f} \]  

(1)

where \( C_{b} \), \( C \), and \( C_{f} \) are concentration of 1 in bound, total, and free states, respectively. The dissociation constant was determined using (2) (ref. 18):

\[ K_{d} = \frac{C_{b}}{C_{0} \times [MBNL]} \]  

(2)

where \( K_{d} \) is the apparent association constant (equal to 1/\( K_{d} \)) and [MBNL] is the concentration of MBNL.

Fluorescence binding assays. Dissociation constants of 2 to various 1 x 1 nucleotide internal loops were measured by an in-solution, fluorescence-based assay. RNA was folded in 1 x PBS at 60 °C for 5 min and allowed to slowly cool to room temperature. 2 was then added to a final concentration of 1,250 nM. This solution was then serially diluted (2 x) into 1 x PBS solution containing 1,250 nM of 2. The solutions were incubated at room temperature for 10 min and then transferred to a 96-well plate. Anisotropy signal was recorded using a SpectraMax M5 plate reader (excitation 300 nm; emission: 470 nm; cutoff: 420 nm). The change in anisotropy signal as a function of RNA concentration was fitted to a standard ligand binding for one site saturation for r(CUG)₁₂, 1 x 1 UU and AU RNAs (3):

\[ \Delta r = \frac{B_{max} \times \text{abs}([RNA])}{K_{a} + \text{abs}([RNA])} \]  

(3)

where \( \Delta r \) is the change in anisotropy signal, \( B_{max} \) represents the maximum number of binding sites, [RNA] is the RNA concentration and \( K_{a} \) is the dissociation constant. The dissociation constant for binding to r(CUG)₁₂ was determined via a Hill plot.

RNA-MBNL1 displacement assays. The RNA-MBNL1 binding studies were completed as previously described²⁹,³⁹. Briefly, 5’-biotinylated r(CUG)₁₂ (folded by heating at 60 °C for 5 min then slowly cooling to room temperature) or MBNL1-His₂₀ was incubated with compound for 30 min at room temperature. Next, r(CUG)₁₂ was added to MBNL1-containing samples or vice versa, and the samples were incubated for 15 min at room temperature. Streptavidin-XL665 (binds to the biotinylated RNA oligonucleotide) and Tb-Anti-His₂₀ (binds to MBNL1) were then added. After 1 h, TR-FRET was measured using a Molecular Devices SpectraMax M5 plate reader. IC₅₀’s were determined by fitting the corresponding curve to SigmaPlot’s four parameter logistic curve fit.

Assessment of DM1-associated translational defects. The model systems used to study the DM1-associated translational defect have been previously described¹⁷. Briefly, a C2C12 cell line that stably expresses the firefly luciferase gene containing a (CTG)₉₀₀ expansion in the 3’-UTR was employed. r(CUG)₉₀₀ causes nuclear retention of the luciferase mRNA and thus decreased expression of luciferase. Compounds that disrupt the r(CUG)₉₀₀-MBNL1 interaction, therefore, have the potential to allow for nucleocytoplasmic export, which is correlated to the luciferase activity in cell extracts. Cells were treated with compound for 16 h and then lysed. Luciferase activity was measured using a Luciferase Assay Kit (Promega) as previously described³⁷.

Assessment of pre-mRNA splicing defects in cellular models. DM1-associated pre-mRNA splicing defects using a HeLa model system were completed as previously described³⁷ using a DM1 mini-gene construct expresses 96 CUG repeat r(CUG) repeat⁵. Briefly, HeLa cells were co-transfected with the mini-gene of interest and a plasmid that expresses r(CUG)⁹⁰ or an empty vector. Varying concentrations of compound were added in growth medium post transfection, and total RNA was harvested 16–20 h later. The percentage of each splicing isoform was determined by RT–PCR. Statistical significance was determined using a two-tailed Student’s t-test.

HEK 293T cells were grown as monolayers in 96-well plates to ~95% confluency in growth medium (1 x DMEM, 10% FBS, and 1 x GlutaMax (Invitrogen)). In order to show that the cTNT mini-gene was responsive to MBNL1 and MBNL2 in the HEK model system, cells were co-transfected with plasmids expressing the DM1 mini-gene (100 ng), the cTNT mini-gene (100 ng) and MBNL1 or MBNL2 (varying amounts: 0–60 ng) using Lipofectamine 2000 reagent as per the manufacturer’s protocol. The effects of compound 2 on MBNL1- and MBNL2-dependent alternate splicing were then determined using 100 ng DM1 mini-gene plasmid, 100 ng cTNT mini-gene plasmid and 3 ng of MBNL1 plasmid. Approximately 5 h post transfection, the transfection cocktail was replaced with growth medium containing various concentrations of 2 (0, 75, 150 and 300 μM).

After 16–24 h, the cells were lysed, and total RNA was collected using a GElute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich). RT–PCR analysis was completed as previously described²⁷,²⁶,³⁷.

DM1 patient-derived cell lines and compound 1 treatment. The following cell lines were obtained from the Coriell Cell repositories: GM07492—fibroblasts from unaffected individual (normal fibroblasts); GM03987, GM04033 and GM03989—fibroblasts from DM1 probands expressing mutant DMPK transcript containing G500, ~1,000 and ~2,000 CUG repeat tracts, respectively (referred to as 500CUG, 1000CUG and 2000CUG, respectively). Cells were grown in a high-glucose EMEM (Lonza) supplemented with 10% fetal bovine serum, 1 x antibiotic-antimycotic (Gibco) and 1 x non-essential amino acids solution (Sigma), in a humidified incubator containing 5% CO₂ at 37 °C. Appropriate amounts of compound 1 were mixed with medium and added directly to 70–80% confluent cells. All splicing changes were analysed 48 h post treatment.

siRNA knockdown of MBNL1. siRNA oligos targeting human MBNL1 were obtained from Future Synthesis, and the sequences were as follows: MBNL1_AS 5’-CGCUUCAACAUUCCUAGGUTG-3’, MBNL1_BS 5’-GACCCUGAAAGUAUGAAGAdTdT-3’. At 80% confluence, fibroblasts were transfected with 25 nM annealed siRNA directed against MBNL1 transcript (referred to as siMBNL1) using Lipofectamine 2000 (Invitrogen). All splicing changes were analysed 48 h post siMBNL1 treatment. For combined siMBNL1 and compound 1 treatment, normal human fibroblasts were successively treated with 25 nM siMBNL1 and 125, 250 or 500 μM compound 1. Cells were first transfected with 25 nM siMBNL1 at 70–80% confluence. Then, appropriate amounts of compound 1 were added 24 h post siMBNL1 treatment. The following controls were used for combined siMBNL1 and 1 treatment: lipofectamine-treated cells (mock), siMBNL1 only, and compound 1 only treated cells. All splicing changes were analysed 48 h post compound 1 treatment (72 h post siMBNL1 treatment).

RNA isolation and RT–PCR. RNA was isolated 48 h post 1 treatment or siRNA treatment using TRI reagent (Sigma-Aldrich) per the manufacturer’s recommended protocol. DNA was synthesized using SuperscriptIII IV Reverse Transcriptase kit (Invitrogen) according to the manufacturer’s protocol, and diluted with 50 μl water before PCR. Standard RT–PCR reactions were performed using 2 μl diluted cDNA and GoTaq Flex DNA Polymerase kit (Promega). PCR primer sequences are listed in Supplementary Table S1. PCR products were separated on a 2% agarose gel stained with 0.5 μg/ml 1 ethidium bromide, photographed using a Bio-Genie (Bio-Tech) and quantified with the use of GeneTools Software (SynGene). Quantified splicing changes were plotted on a bar graph as a percentage of the respective exon spliced in (% of alternative exon inclusion), which in the case of MBNL1-dependent exons represents splicing isoforms characteristic of DM1.

Experimental set-up and statistical analyses. The effect of 1 on normal and DM1 (CUG500) human fibroblasts was analysed in duplicate biological repeats (n = 2 for each given concentration of 1). As controls, we used n = 5 each for untreated DM1 (CUG500) fibroblasts and for the untreated DM1 (CUG500) fibroblasts collected from each of the untreated CUG1000 and CUG2000 human fibroblasts. The effects of siMBNL1 knockdown as well as combined siMBNL1 and compound 1 treatment were assessed in duplicate biological repeats. All RT–PCR reactions were performed and analysed at least twice to verify results. The fractions of alternatively spliced isoforms (exon inclusions representing the DM1-like splicing isoform) were calculated by dividing the intensity of a PCR product band corresponding to the splicing variant representing exon inclusion (upper band) by the total intensity of both splicing isoforms (upper and lower band). Error bars represent s.d. (± s.d.) obtained from a single experiment with at least two independent biological repeats. Statistical significance was determined by a two-tailed Student’s t-test using Microsoft Excel (* for P ≤ 0.05, ** for P ≤ 0.01 and *** for P < 0.001).

Docking of 1 and MBNL1. Compound 1 was prepared for docking simulations using LigPrep v2.5 (Schrodinger, LLC)⁴⁴, affording a single ligand pose. PDB entry 3DS5 (MBNL1) was prepared for molecular modelling by using Maestro’s protein preparation wizard (v9.3; Schrodinger, LLC). Bound RNA and water molecules were removed, the bond order was assigned, and hydrogen atoms were added. The resultant states of Asn and Gln residues, and the charge states of His residues were also optimized. The docking grid was generated around the Zn coordinate with a box size of 15 x 15 x 15 Å³. No constraints were used for grid generation. Compound 1 was then docked into MBNL1 ref. 34 using Glide SP v5.6 (Schrodinger, LLC). Energy minimization was completed using Prime v3.1 under the QM/MM interface (Schrodinger, LLC). Figure 7a was generated using PyMol (Schrodinger, LLC).

Docking of 2 and r(CUG) repeats. Please see the Supplementary Methods and Supplementary Tables S2–S5 for details related to molecular modelling of 2 and r(CUG) repeats.
References

1. Faustino, N. A. & Cooper, T. A. Pre-mRNA splicing and human disease. *Genes Dev.* 17, 419–437 (2003).
2. Savkur, R. S. et al. Insulin receptor splicing alteration in myotonic dystrophy type 2. *Am. J. Hum. Genet.* 74, 1309–1313 (2004).
3. Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T. & Thornton, C. A. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* 13, 3079–3088 (2004).
4. Fugier, C. et al. Misregulated alternative splicing of BIV1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat. Med.* 17, 720–725 (2011).
5. Sellier, C. et al. Sam68 sequestration and partial loss of function are associated with alternative splicing in FXTAS patients. *EMBO J.* 29, 1248–1261 (2010).
6. Liquori, C. L. et al. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293, 864–867 (2001).
7. Du, H. et al. Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.* 17, 187–193 (2010).
8. Orengo, J. P., Ward, A. J. & Cooper, T. A. Alternative splicing dysregulation type 1 RNA via an RNA motif-ligand database & chemical similarity searching. *J. Am. Chem. Soc.* 134, 4731–4742 (2012).
9. Teplova, M. & Patel, D. J. Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL. *Nat. Struct. Mol. Biol.* 15, 1343–1351 (2008).
10. Parkes, R. et al. Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database & chemical similarity searching. *J. Am. Chem. Soc.* 134, 4731–4742 (2012).
11. Cardani, R., Mancinelli, E., Rotondo, G., Sansone, V. & Meola, G. Muscleblind-like protein 1 nuclear sequestration is a molecular pathology marker of DM1 and DM2. *Eur. J. Histochim.* 50, 177–182 (2006).
12. Chen, C. Z. et al. Two high-throughput screening assays for aberrant RNA-protein interactions in myotonic dystrophy type 1. *Anal. Bioanal. Chem.* 402, 1889–1898 (2012).
13. Childs-Disney, J. L., Hoskins, J., Reuczik, S. G., Thornton, C. A. & Disney, M. D. Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem. Biol.* 7, 856–862 (2012).
14. Chaires, J. B., Ragazzon, P. A. & Garbert, N. C. A competition diayliss for the study of structure-selective ligand binding to nucleic acids. *Carr. Protoc. Nucleic Acid Chem. Chapter 8.* 8, 8.3.1–8.3.8 (2003).
15. Kumar, A. et al. Myotonic dystrophy type 1 RNA crystal structures reveal heterogeneous 11-nucleotide UU internal loop conformations. *Biochemistry* 50, 9928–9935 (2011).
16. Varf, M. & Berglund, J. A. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA* 13, 2283–2291 (2007).
17. Lee, M. M., Pursechnikov, A. & Disney, M. D. Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS Chem. Biol.* 4, 345–355 (2009).
18. Mykowska, A., Sobczak, K., Wojciechowska, M., Kozlowski, P. & Krzyzosiak, W. J. CAG repeats mimic CUG repeats in the misregulation of alternative splicing. *Nucleic Acids Res.* 39, 8938–8951 (2011).
19. Tsai, Y. H. et al. Muscleblind proteins regulate alternative splicing. *EMBO J.* 23, 3103–3112 (2004).
20. Danshithong, W., Paul, S., Comai, L. & Reddy, S. MBNL1 is the primary determinant of focus formation and aberrant insulin receptor splicing in DM1. *J. Biol. Chem.* 280, 5773–5780 (2005).
21. Orento, J. P., Bundman, D. & Cooper, T. A. A bichromatic fluorescent reporter for cell-based screens of alternative splicing. *Nucleic Acids Res.* 34, e148 (2006).
22. Warf, M. B., Nakamori, M., Matthys, C. M., Thornton, C. A. & Berglund, J. A. Pentamidine reverses the splicing defects associated with myotonic dystrophy. *Proc. Natl Acad. Sci. USA* 106, 18551–18556 (2009).
23. Paronetto, M. P., Achsel, T., Massiello, A., Chaffant, C. E. & Sette, C. The RNA-binding protein Sam68 modulates the alternative splicing of Cdx-1. *J. Cell. Biol.* 176, 929–939 (2007).

Acknowledgements

We thank Don Mackenzie from the Marigold Foundation for his support of this research project in particular and DM1 research in general and Amit Kumar for preliminary studies; and Professor Glenn E. Morris (Watson Centre for Inherited Neuromuscular Disease) for generously providing the MB1a antibody (anti-MBNL1). Funding for this work was provided by the Scripps Research Institute, the Muscular Dystrophy Association (158552 to MDD), the National Institutes of Health (3R01GM079235-02S1 and 1R01GM079235-01A2 to MDD; AR409077 and U54NS48843 to CAT), the Molecular Libraries Initiative of the NIH Roadmap for Medical Research, the Marigold Foundation postdoctoral fellowship, the FS-OC Center of the NIH/NCI (grant 1U54CA14869-01; I.Y. and G.C.S.), and the Foundation for Polish Science—TEAM programme co-financed by the European Union within the European Regional Development Fund (E-S-K and K.S.).

Author contributions

J.L.C.-D and M.D.D. wrote the manuscript, which was also edited by co-authors. J.L.C.-D, E.S.-K, T.T., I.Y., H.P., C.Z.C., J.H., N.S., J.J.M., S.P., and W.Z. completed the study of structure-selective ligand binding to nucleic acids. Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Childs-Disney, J. L. et al. Induction and reversal of myotonic dystrophy type 1 pre-mRNA splicing defects by small molecules. *Nat. Commun.* 4:2044 doi: 10.1038/ncomms3044 (2013).

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms3044 ARTICLE

11

© 2013 Macmillan Publishers Limited. All rights reserved.