Published in final edited form as:

Chembiochem. 2012 November 26; 13(17): 2505–2509. doi:10.1002/cbic.201200602.

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

Chun-Ho Wong^[a], Dr. Stacie L. Richardson^[a], Yen-Jun Ho^[a], Alex M. H. Lucas^[a], Prof. Dr. Tiziano Tuccinardi^[b], Prof. Dr. Anne M. Baranger^[a], and Prof. Dr. Steven C. Zimmerman^[a]

Steven C. Zimmerman: sczimmer@illinois.edu

^[a]Department of Chemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana IL 61801 (USA)

^[b]Dipartimento di Scienze Farmaceutiche, Università di Pisa via Bonanno 6, 56126 Pisa (Italy)

Keywords

DNA; drug design; inhibitors; recognition; RNA

The development of small molecules that recognize specific DNA and RNA sequences or structures remains a critical challenge. If cell-permeable, such agents might allow the regulation of genes for therapeutic purposes or the targeting of DNA or RNA sites known to play some other roles in causing disease. A few general approaches to the recognition of double-helical DNA,^[1] base mismatches,^[2] G-quartets,^[3] and other targets^[4] have emerged over the past two decades. However, many difficulties remain. For example, it is often difficult to obtain high selectivity for DNA over RNA or vice versa.^[5] We reported that ligand **1** binds CUG and CTG sites with similar high affinity and selectivity in RNA and DNA oligonucleotides, respectively.^[6] Herein we show that ligands **2** and **3**, with one or two N-methyl groups, respectively, selectively abolish RNA binding, with **3** giving a more than threefold increase in DNA affinity that could allow DNA-targeted gene therapy for myotonic dystrophy type 1 (DM1).^[7]

This work arose from the well-supported hypothesis that DM1 originates in the sequestration of the alternate splicing regulator protein, muscleblind-like 1 (MBNL1), by abnormally long CUG triplet repeats. It has been suggested that ligands able to bind selectively to these pathogenic triplet repeats might reverse the DM1 phenotype by inhibiting the MBNL1-RNA complex, allowing the splicing regulator to resume its normal function.^[8] Much less attention has focused on agents that might shorten or halt the expansion of the CTG repeats found in the *DMPK* gene but this DNA sequence has been identified as a potential target for molecular therapy.^[9] Thus, we developed ligand **1**,^[6] which selectively binds CTG and CUG sites and inhibits the MBNL1-CUG complex.^[10]

The design of ligand **1** was based on the X-ray crystal structure of r(CUG)₆ determined by Berglund,^[11] which showed the CUG repeat in a standard A-form helical structure with the U–U mismatch flanked by G–C pairs. The triaminotriazine unit was selected as a Janus

wedge^[12–14] to form a base triplet with U–U (Figure 1A). Critical to the design was the assumption that the two heterocycles would π -stack, reducing the nonspecific intercalation (Figure 1B). Insertions from the minor and major groove were considered, but the former was in line with the intercalation preference of 9-aminoacridine.^[15]

The data collected for ligand **1** was consistent with the binding model shown in Figures 1 and 2A. However, a detwinned CUG repeat structure with high-resolution hydration details^[16] led us to consider other possible binding modes. In particular, there was concern that the minor groove base-triplet model (Figure 2A) required the lengthening of the C1'–C1' distance of the U–U from 10.4 to 13.8 Å. Although the RNA minor groove is wide and shallow, 3.4 Å is a significant increase in the C1'–C1' distance (Figure 2A) and may signal a distortion in the DNA or RNA backbone that is energetically unfavorable.^[17] This distance is even longer than that of a G–A purine-purine pair (12.5 Å).^[18] In fact, repeated attempts to model r(CUG)₆·**1** with the ligand in the minor groove failed to yield a stable complex.

Modeling showed that the water molecule bridging the two uracil bases in a stretched wobble pair (Figure 2C) could be replaced by the triazine unit in the minor groove (Figure 2D).^[19] This led to a smaller structural reorganization of the original pair (Figure 2D). Further, modeling of a major groove base triplet suggested a C1'–C1' distance even closer to that in the X-ray structure (Figure 2B). In absence of direct structural information, an extended “methyl scanning” approach^[20] appeared to be ideal to distinguish between models in Figure 2A and 2D. Thus, N-methylated analogues **2–6** were synthesized and tested using isothermal titration calorimetry (ITC) with both d(CTG)₂ and r(CUG)₂ (Figure 3). The ability to inhibit the MBNL1·CUG interaction was examined by electrophoretic mobility shift assays (EMSA). These methods were supplemented with molecular dynamics (MD) simulations.

The methyl scanning method was possible because the number of hydrogen bonds differ significantly in the base triplet (Figure 2A and B) and the stretched wobble binding modes (Figure 2D). In particular, the wobble binding mode has two, possibly three, free N–H groups, whereas only a single N–H group is free in the triplet binding modes.

The possibility of base-triplet recognition through the major groove was also considered, because simple modeling indicated the C1'–C1' distance to be minimally adjusted from 10.4 to 10.1 Å (Figure 2B). Indeed, an unconstrained 10 ns MD simulation showed ligand **1** could fit satisfactorily in the major groove (Figure 4A and Video S1). The MD simulation of the minor groove stretched wobble binding was also performed (Figure 4B and Video S2). Although fewer hydrogen bonds are formed, the unconstrained simulation led to a stable structure. The average C1'–C1' distance of the U–U mismatch is moderately lengthened by 0.4 Å (11.8 vs. 10.4 Å, see Figure 5).

The d(CTG)₂ recognition by **1–6** was measured by ITC and all K_d values are collected in Table 1. Ligand **3** was found to bind more than threefold stronger than ligand **1** ($K_d = 0.12$ μ M for **3** vs. 0.39 μ M for **1**), whereas ligands **2**, **4**, and **5** each bound d(CTG)₂ but progressively more weakly than **1**.^[21] Ligand **6**, carrying three methyl groups, showed no measurable binding to d(CTG)₂. Because of the limited aqueous solubility, we were only able to assign a lower limit of $K_d > 200$ μ M. The combined ITC and modeling results are consistent with Watson–Crick-type recognition of the stretched wobble pair (Figure 2C). An unconstrained 10 ns MD simulation further suggested the weakly bound thymine could be flipped-out from the dsDNA (Video S3 and Figure S3). In contrast to the results with d(CTG)₂, ligand **3** did not bind under similar conditions to r(CUG)₂ (Figure 6, Table 1) and, indeed, the single N-methyl group in **2** was enough to eliminate ligand binding. The only

CUG-binding, N-methylated ligand was **4**, and even its K_d was seven times weaker than that of **1**. All methylated ligands also showed reduced affinity towards herring sperm DNA (hsDNA). The ability of the recognition unit to stack on the intercalator is enhanced by methylation which in turn reduced the nonspecific intercalative binding to duplex DNA.^[22]

To assess the potential of these ligands for drug development, their binding affinity to various DNA and RNA duplexes was studied by ITC (Table S1). All methylated ligands showed reduced affinity toward the duplexes, including hsDNA. The ability to complex other mismatches was also measured. In general, these ligands exhibited high selectivity (up to 1600-fold) for DNA over RNA mismatches. The ligands were also found to bind weakly to purine-purine mismatches. Interestingly, all the methylated ligands bind quite strongly ($K_d = 0.13$ – $2.0 \mu\text{M}$) to d(CCG)₂, a trinucleotide repeat sequence associated with Fragile XE syndrome^[23] and chronic lymphocytic leukemia,^[24] thus suggesting a potential molecular therapeutic approach to these diseases that warrants additional study.

The ability of ligands **1**–**6** to inhibit MBNL1 binding to r(CUG)₁₂ was measured by EMSA, and the results paralleled the ITC data. Thus, only ligands **1** and **4** showed inhibition (Figure 7). Although nothing can be said about the groove preference from these experiments, the MD simulations favor major groove triplet formation (see above), and it is known that loop and mismatch structures can enhance the accessibility of the major groove.^[25]

The combined experimental and computational approach described herein suggests that this class of ligands bind DNA and RNA by significantly different modes. Thus, it is proposed that ligands **1** and **4** recognize the U–U mismatch in RNA through formation of a major groove base triplet (Figure 8A) whereas ligands **1**–**5** bind the T–T mismatch in DNA through the stretched wobble pair (Figure 8B). These binding modes also explain the decreased binding to U–U by **4** ($K_d = 14 \mu\text{M}$ vs. $2.1 \mu\text{M}$ for **1**) and to T–T by **5** ($K_d = 50 \mu\text{M}$ for **5** vs. $0.39 \mu\text{M}$ for **1**). Thus, the C–NH(Me) bond rotation^[26] leads to unfavorable binding for one or more of the rotamers.^[27]

In conclusion, we have used the methyl scanning method combined with MD simulations to indirectly investigate the possible binding modes by which ligands **1**–**5** recognize CUG and CTG sites in RNA and DNA, respectively. Beyond informing on the binding mode, the ability to substitute the amino groups of ligand **1** suggests these as sites for further modifications that might enhance the selectivity and efficacy of these lead compounds in treating DM1.

More significant was the unexpected discovery of DNA-selective ligands for CTG. Whereas ligand **1** binds both CUG and CTG with similar strength, ligand **3** showed a more than threefold increase in affinity to CTG with negligible binding to RNA. It was also found that these ligands showed reduced binding toward DNA and RNA duplexes (i.e., without mismatches) upon methylation, potentially opening an avenue for a DNA-targeted molecular therapy of DM1.^[7] Competitive binding to the corresponding CUG transcript would be expected to reduce the effectiveness of this approach and could also complicate efforts to assess the DNA-targeted approach. More broadly, this study increases our knowledge of how small molecules can selectively recognize nucleic acids.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (RO1AR058361). C.H.W. is a recipient of a Croucher Foundation Scholarship (HK) and a Pines Graduate Fellowship.

References

1. a) Dervan PB, Edelson BS. *Curr Opin Struct Biol.* 2003; 13:284–299. [PubMed: 12831879] b) Dervan PB. *Bioorg Med Chem.* 2001; 9:2215–2235. [PubMed: 11553460]
2. Nakatani K. *Bull Chem Soc Jpn.* 2009; 82:1055–1069.
3. Haider SM, Neidle S, Parkinson GN. *Biochimie.* 2011; 93:1239–1251. [PubMed: 21635933]
4. a) Ong HC, Arambula JF, Ramisetty SR, Baranger AM, Zimmerman SC. *Chem Commun.* 2009:668–670. b) Nakatani K, Horie S, Murase T, Hagihara S, Saito I. *Bioorg Med Chem.* 2003; 11:2347–2353. [PubMed: 12713847] c) Sankaran NB, Sato Y, Sato F, Rajendar B, Morita K, Seino T, Nishizawa S, Teramae N. *J Phys Chem B.* 2009; 113:1522–1529. [PubMed: 19175344]
5. a) Rzuczek SG, Pilch DS, Liu A, Liu L, LaVoie EJ, Rice JE. *J Med Chem.* 2010; 53:3632–3644. [PubMed: 20359224] b) Collie G, Reszka AP, Haider SM, Gabelica V, Parkinson GN, Neidle S. *Chem Commun.* 2009:7482–7484. c) Wilson WD, Ratmeyer L, Cegla MT, Spychala J, Boykin D, Demeunynck M, Lhomme J, Krishnan G, Kennedy D, Vinayak R, Zon G. *New J Chem.* 1994; 18:419–423.
6. Arambula JF, Ramisetty SR, Baranger AM, Zimmerman SC. *Proc Natl Acad Sci USA.* 2009; 106:16068–16073. [PubMed: 19805260]
7. a) Castel AL, Cleary JD, Pearson CE. *Nat Rev Mol Cell Biol.* 2010; 11:165–170. [PubMed: 20177394] b) Mirkin SM. *Nature.* 2007; 447:932–940. [PubMed: 17581576] c) Gomes-Pereira M, Monckton DG. *Mutat Res.* 2006; 598:15–34. [PubMed: 16500684]
8. Cooper TA. *N Engl J Med.* 2006; 355:1825–1827. [PubMed: 17065646]
9. a) Gomes-Pereira M, Monckton DG. *Nucleic Acids Res.* 2004; 32:2865–2872. [PubMed: 15155855] b) Hashem VI, Pytlos MJ, Klysik EA, Tsuji K, Khajav M, Ashizawa T, Sinden RR. *Nucleic Acids Res.* 2004; 32:6334–6346. [PubMed: 15576360]
10. For efforts by other groups, see García-López A, Llamusi B, Orzáez M, Pérez-Payá E, Artero RD. *Proc Natl Acad Sci USA.* 2011; 108:11866–11871. [PubMed: 21730182] Childs-Disney JL, Tsitovich PB, Disney MD. *Chem-BioChem.* 2011; 12:2143–2146. Disney MD, Lee MM, Pushechnikov A, Childs-Disney JL. *ChemBioChem.* 2010; 11:375–382. [PubMed: 20058255] Warf MB, Nakamori M, Matthys CM, Thornton CA, Berglund JA. *Proc Natl Acad Sci USA.* 2009; 106:18551–18556. [PubMed: 19822739] Pushechnikov A, Lee MM, Childs-Disney JL, Sobczak K, French JM, Thornton CA, Disney MD. *J Am Chem Soc.* 2009; 131:9767–9779. [PubMed: 19552411] Lee MM, Pushechnikov A, Disney MD. *ACS Chem Biol.* 2009; 4:345–355. [PubMed: 19348464] Gareiss PC, Sobczak K, McNaughton BR, Palde PB, Thornton CA, Miller BL. *J Am Chem Soc.* 2008; 130:16254–16261. [PubMed: 18998634]
11. Mooers BHM, Logue JS, Berglund JA. *Proc Natl Acad Sci USA.* 2005; 102:16626–16631. [PubMed: 16269545]
12. Branda N, Kurz G, Lehn JM. *Chem Commun.* 1996:2443–2444.
13. a) Chen D, Meena, Sharma SK, McLaughlin LW. *J Am Chem Soc.* 2004; 126:70–71. [PubMed: 14709064] b) Chen H, Meena, McLaughlin LW. *J Am Chem Soc.* 2008; 130:13190–13191. [PubMed: 18783217] c) Shin D, Tor Y. *J Am Chem Soc.* 2011; 133:6926–6929. [PubMed: 21495708] d) Zeng Y, Pratumyot Y, Piao X, Bong D. *J Am Chem Soc.* 2012; 134:832–835. [PubMed: 22201288]
14. Replacing the triazine by pyrimidine unit led to a change in mismatch selectivity, presumably due to the difference in pK_a values, see Wong CH, Fu Y, Ramisetty SR, Baranger AM, Zimmerman SC. *Nucleic Acids Res.* 2011; 39:8881–8890. [PubMed: 21768123]
15. a) Adams A, Guss JM, Collyer CA, Denny WA, Wakelin LPG. *Biochemistry.* 1999; 38:9221–9233. [PubMed: 10413496] b) Woodson SA, Crothers DM. *Biochemistry.* 1988; 27:8904–8914. [PubMed: 3233212]
16. Kiliszek A, Kierzek R, Krzyzosiak WJ, Rypniewski W. *Nucleic Acids Res.* 2009; 37:4149–4156. [PubMed: 19433512]

17. a) Hikishima S, Minakawa N, Kuramoto K, Fujisawa Y, Ogawa M, Matsuda A. *Angew Chem.* 2005; 117:602– 604. *Angew Chem Int Ed.* 2005; 44:596–598. b) Minakawa N, Kojima N, Hikishima S, Sasaki T, Kiyosue A, Atsumi N, Matsuda A. *J Am Chem Soc.* 2003; 125:9970–9982. [PubMed: 12914460]
18. Prive GG, Heinemann U, Chandrasegaran S, Kan LS, Kopka ML, Dickerson RE. *Science.* 1987; 238:498– 504. [PubMed: 3310237]
19. This type of singly hydrogen-bonded U-U mismatches is also observed in a recent NMR study, see Parkesh R, Fountain M, Disney MD. *Biochemistry.* 2011; 50:599– 601. [PubMed: 21204525]
20. Pirrung MC, Liu Y, Deng L, Halstead DK, Li Z, May JF, Wedel M, Austin DA, Webster NJG. *J Am Chem Soc.* 2005; 127:4609– 4624. [PubMed: 15796526]
21. The result here is different from the observation in ref. [13 d] that a N,N-dimethylated triazine peptide inhibited the triplex formation.
22. a) Gung BW, Emenike BU, Alvarez CN, Rakovan J, Kirschbaum K, Jain N. *Tetrahedron Lett.* 2010; 51:1648–1650. [PubMed: 20209117] b) Zadorozhnaya AA, Krylov AI. *J Phys Chem A.* 2010; 114:2001– 2009. [PubMed: 20055394]
23. a) Cummings CJ, Zoghbi HY. *Annu Rev Genomics Hum Genet.* 2000; 1:281–328. [PubMed: 11701632] b) Pearson CE, Sinden RR. *Curr Opin Struct Biol.* 1998; 8:321– 330. [PubMed: 9666328]
24. Auer RL, Jones C, Mullenbach RA, Syndercombe-Court D, Milligan DW, Fegan CD, Cotter FE. *Blood.* 2001; 97:509– 515. [PubMed: 11154230]
25. Weeks KM, Crothers DM. *Science.* 1993; 261:1574– 1577. [PubMed: 7690496]
26. Rotamers were observed for those ligands in both ^1H and ^{13}C NMR spectroscopy (see the Supporting Information).
27. a) Ghiviriga I, Oniciu DC. *Chem Commun.* 2002:2718–2719. b) Mathias JP, Simanek EE, Zerkowski JA, Seto CT, Whitesides GM. *J Am Chem Soc.* 1994; 116:4316– 4325.

Figure 1.

A) Ligand **1** contains a triaminotriazine unit designed to selectively target T–T and U–U mismatches. B) Proposed binding mode of ligand **1** to CUG (or CTG) repeats.

Figure 2.

Three potential modes of binding: A) Minor groove triplet binding. B) Major groove triplet binding. C) A stretched U-U wobble pair with a bridging water molecule. D) Minor groove stretched wobble binding. R = CH₃ (DNA) or H (RNA).

Figure 3.

A series of N-methylated versions (2–6) of ligand 1, and RNA and DNA duplexes containing a r(CUG)₂ and d(CTG)₂ motifs, respectively, for structure–activity relationship (SAR) studies.

Figure 4. Snapshots of minimized structures (Amber) from the 10 ns MD simulations of A) the major groove triplet binding and B) the minor groove stretched wobble binding (see also the Supporting Information).

Figure 5.

Analysis of the C1'-C1' distance of 10 ns MD simulations of A) the major groove triplet binding and B) the minor groove stretched wobble binding. The average C1'-C1' distance for all the base pairs is reported in black, whereas C1'-C1' distance of the U-U pair involved in the binding with the triaminotriazine ring is reported in gray.

Figure 6. ITC profiles for complexation of ligand **3** with duplex containing the A) T-T and B) U-U mismatches.

Figure 7.

EMSA screening of ligands **1–6** showing only ligands **1** and **4** inhibit MBNL1·CUG interaction. [r(CUG)₁₂] = 0.2 nM, [MBNL1] = 460 nM, [ligand] = 100 μM, 10 % DMSO (except lane 2), Tris·borate (pH 8.3). The structure of r(CUG)₁₂ used in the assay is shown on the right.

Figure 8.
Proposed binding modes for the A) U–U and B) T–T mismatches.

Table 1

Dissociation constants [μM] of ligands **1–6** with r(CUG)₂ and d(CTG)₂ containing U–U and T–T mismatches, respectively.

Ligand	r(CUG) ₂	d(CTG) ₂	hsDNA ^[b]
1	2.1 $\pm$ 0.2	0.39 $\pm$ 0.08	55 $\pm$ 16
2	n.b. ^[a]	0.8 $\pm$ 0.4	n.b. ^[a]
3	n.b. ^[a]	0.12 $\pm$ 0.04	n.b. ^[a]
4	14 $\pm$ 5	2.2 $\pm$ 0.5	n.b. ^[a]
5	n.b. ^[a]	50 $\pm$ 30	n.b. ^[a]
6	n.b. ^[a]	n.b. ^[a]	n.b. ^[a]

^[a] n.b.: no measurable binding and a lower limit of $K_d > 200 \mu\text{M}$ is assigned.

^[b] Herring sperm DNA.