Published in final edited form as:

Biochemistry. 2020 September 22; 59(37): 3463–3472. doi:10.1021/acs.biochem.0c00472.

Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors

Lauren D. Hagler^{‡,1}, Long M. Luu^{‡,1}, Marco Tonelli², JuYeon Lee¹, Samuel M. Hayes³, Sarah E. Bonson¹, J. Ignacio Vergara¹, Samuel E. Butcher³, Steven C. Zimmerman^{1,*}

¹Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

²National Magnetics Resonance Facility at Madison, Biochemistry Department, University of Wisconsin-Madison, Madison, WI 53706, USA

³Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

Abstract

There are few methods available for the rapid discovery of multi-target drugs. Herein, we describe the template-assisted, target-guided discovery of small molecules that recognize d(CTG) in the expanded d(CTG-CAG) sequence and its r(CUG) transcript that cause myotonic dystrophy type 1 (DM1). A positive cross-selection was performed using a small library of 30 monomeric alkyne- and azide-containing ligands capable of producing more than 5000 possible di- and trimeric click products. The monomers were incubated with d(CTG)₁₆ or r(CUG)₁₆ under physiological conditions and both sequences showed selectivity in the proximity-accelerated azide-alkyne [3+2]-cycloaddition click reaction. The limited number of click products formed in both selections and the even smaller number of common products suggests that this method is a useful tool for the discovery of single-target and multi-target lead therapeutic agents.

Fragment-based drug discovery (FBDD)^{1–5} and target-guided synthesis (TGS)^{6–9} have emerged as important approaches to discover lead therapeutic agents. Both FBDD and TGS are broadly defined and involve a suite of different strategies that share conceptual and some operational similarities, including the formation of tight binding ligands from smaller fragments. Most relevant to this work is the *in situ* TGS approach wherein the target receptor serves as a template, binding two or more fragments whose complementary and reactive functional groups are in close proximity and capable of reaction. The resultant covalent coupling product typically has a much higher affinity. We are particularly interested in applying a TGS approach to myotonic dystrophy type 1 with the goal of discovering highly selective multivalent and multitargeting lead therapeutic agents.

*Corresponding Author: sczimmer@illinois.edu.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

[‡]These authors contributed equally.

Supporting Information includes details on synthesis and characterization of all ligands, HPLC analysis of template-mediated click reaction, and EMSA experiments. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Myotonic dystrophy type 1 (DM1) is an incurable neuromuscular disease for which there is currently no direct treatment.^{10–13} Figure 1a shows schematically the RNA gain-of-function model for the disease, where aberrantly expanded d(CTG-CAG) trinucleotide repeats in the 3'-UTR of the DMPK gene undergo bidirectional transcription to produce r(CUG) and r(CAG) transcripts. The former sequesters the MBNL family of alternative splicing regulators, including MBNL1, into nuclear foci¹⁴ and the loss of MBNL function leads to the mis-regulation of pre-mRNA splicing (>100 pre-mRNAs).¹⁵ An understanding of the complex pathobiology of the disease has led to several therapeutic strategies for mitigating disease phenotypes, including gene editing and transcription inhibition.^{16–24} However, the greatest effort arguably has focused on small molecule inhibitors of the r(CUG)^{exp}-MBNL interaction. Indeed, our group^{23,25,26} and others^{27,28} have reported potent dimeric and oligomeric ligands that recognize r(CUG) repeats. Additionally, we have reported several multitarget strategies for increased efficacy that include agents capable of simultaneously inhibiting r(CUG)^{exp}-MBNL foci formation, cleaving RNA, inhibiting transcription, and decreasing r(CUG) levels in cells and *in vivo*.^{23,24} In searching for a more effective approach to developing multitarget agents that inhibit both RNA synthesis and RNA-protein interactions (Figure 1b), we sought to adapt the TGS strategy pioneered by Sharpless²⁹ and developed by others^{30,31} for the discovery of multitarget and multivalent small molecules from a single library and assay (Figure 1c).

The development of click chemistry has significantly advanced TGS with a number of tight-binding ligands for proteins discovered in recent years.^{29,30} In contrast, there are few studies that use DNA or RNA targets for TGS.^{31–33} In 2003, Dervan and coworkers demonstrated the ability of DNA to template the alkyne-azide [3+2] cycloaddition of hairpin polyamides, giving a dimeric, triazole product with high binding affinity and specificity.³⁴ To the best of our knowledge there are no reports of screening a single mixed fragment library to discover multitarget agents for proteins or nucleic acids. Herein, we describe a novel library approach that uses a template-assisted azide-alkyne click reaction in a selection format with both d(CTG)₁₆ and r(CUG)₁₆ targets. The assay involves just three parallel runs against two targets. Hit dimeric compounds bound more tightly to d(CTG) and r(CUG) repeats than their monomeric counterparts, inhibited transcription and r(CUG)-MBNL interactions *in vitro*, and formed specific interactions with UU mismatches in RNA hairpin structures in solution. A limited number of products were formed on each target and even fewer formed in both selections, suggesting that this method is useful for the discovery of both single and multitarget lead therapeutic agents.

MATERIALS AND METHODS

Nucleic acid oligomers d(CTG)₂, r(CCUG)₈, r(CUG)₁₂, tRNA, and d(CTG)₁₆ were purchased from IDT technologies and used without further purification (sequences are below). Duplex DNA containing d(CTG-GAC)₁₆ and the consensus sequence for T7 polymerase was also purchased from IDT without further purification for *in vitro* transcription of r(CUG)₁₆. r(CUG)₁₆ was synthesized by *in vitro* transcription from d(CTG)₁₆ containing a T7 promoter and T7 RNA polymerase (see Supporting Information).

r(CCUG)₈: 5'—GG(CCUG)₈CC—3'

r(CUG)₁₂: 5'—GGG(CUG)₁₂CCC—3'

d(CTG)₁₆: 5'—GGG(CTG)₁₆CCC—3'

d(CTG)₁₆ with T7 promoter sequence:

5'—TTCTAATACGACTCACTATAGGG(CTG)₁₆CCC—3'

d(GAC)₁₆ with T7 promoter sequence:

3'—AAGATTATGCTGAGTGATATCCC(GAC)₁₆GGG—5'

r(CUG)₁₆: 5'—GGG(CUG)₁₆CCC—3'

Selection Assay Using DNA and RNA Templates and Azide- and Alkyne-containing Ligands.

The templated reactions (400 μ L, 25 μ M each compound, 12.5 μ M DNA/RNA, 2 mM Tris-HCl pH 7.0, 2 mM KCl, 2 mM MgCl₂, and 1 mM CaCl₂, 5% DMSO) were incubated at 37 °C for one week. After 1, 2, 3, 4, 5, 6 and 7 d, 10 μ L of the RNA-templated reaction was withdrawn and mixed with 10 μ L of aqueous urea (8 M) to denature the RNA hairpin and 1 μ L of RNase A (10 mg/mL) to degrade the RNA. Similarly, 10 μ L of the DNA-templated reaction was removed at 1, 2, 3, 4, 5, 6, and 7 d and mixed with 1 μ L of TURBO-DNase A 10X to degrade the DNA. The resulting samples were incubated at 37 °C for 20 min and analyzed by MALDI-TOF MS using DHB as the matrix. The reaction without DNA/RNA was conducted using a similar procedure. All experiments were performed in triplicate.

NMR Studies with r(CUG) Hairpin.

1D NMR, 2D NOESY and 2D HSQC data were collected as previously described.³⁵ Assignment of uridine imino ¹H and ¹⁵N chemical shifts in the CUG repeat regions were made by analysis of NOEs in 2D NOESY data and comparison to ¹H and ¹⁵N chemical shifts of previously assigned U-U wobble pairs in helical conformations.³⁶ The formation of U-U wobble pairs in the free hairpin conformation is supported by the NMR chemical shift data and X-ray crystallography of CUG repeats in RNA.³⁷

Inhibition Assays.—These experiments were conducted as previously described with minor changes.²⁴ For an *in vitro* transcription inhibition experiment, ligands at desired concentrations were incubated with 15 ng of linearized d(CTG·CAG)₇₄ plasmid in 1X T7 RNA polymerase buffer (80 mM Tris pH 8.3, 10 mM MgCl₂, 2 mM spermine, 0.1% Triton-S, 10 mM NaCl) containing 0.5 mM each of rATP, rGTP, rCTP, and rUTP at 37 °C. After 3.5 h, T7 polymerase (Biolab) was added to give a final activity of 0.5 U, and the transcription reaction was allowed to proceed for 1.5 h. The reaction was quenched by adding 7 μ L of urea (8 M), 1 μ L of RNA denaturing dye (95% formamide, 5 mM EDTA, 0.025% each xylene cyanol and bromophenol blue), and 2 μ L of HIV fs RNA (1 μ M) as a spike-in standard and heated to 95 °C for 5 min. The solution (15 μ L) was loaded on an 8% denaturing PAGE gel, and the RNA was detected by ethidium bromide post-staining. Bands were quantified using Fiji (ImageJ, NIH). The ratio of the intensity of r(CUG)₇₄ to HIV fs RNA from the ligand-treated sample was compared with that from the untreated sample,

providing a percentage of transcription for each ligand and concentration. % inhibition = 100 - % transcription. For details on EMSA assay, see Supporting Information.

RESULTS AND DISCUSSION

Template-Assisted Click Reaction and Design of Clickable Fragment Library.

To determine the approximate reaction time required for the template-assisted click reaction, a model study was performed using an acridine intercalator with either an azide or activated alkyne side-chain. Acridine intercalates into dsDNA non-specifically following the neighbor-exclusion model, i.e., binding to every other base pair site.³⁸ Computational modeling with a short DNA duplex was used to identify linkers of sufficient length for proximity-induced, template-assisted [3+2] cycloaddition (Figure 2a–b). To verify our computational results, alkyne **1** and azide **2** were incubated at 37 °C with d(CTG)₂ duplex. The formation of triazole product **3** was detected by MALDI-TOF MS and monitored by HPLC over 7 d. Dimeric product was detected at 1 d. Moreover, the concentration of alkyne **1** and azide **2** decreases by 76% and 70%, respectively, over 7 d when incubated with the DNA target (Figure 2c). Having established the conditions for the click selection, we chose a clickable fragment library that included small molecule ligands known to bind to nucleic acid secondary structures. To increase the likelihood of r(CUG) and d(CTG) selective products, we included fragments of known DM1 selective ligands (**4–6**) that bind strongly to these targets and corrected DM1 phenotypes in cells.^{22–25} These ligands were developed to selectively target the toxic r(CUG)^{exp} or d(CTG)^{exp} hairpin through recognition of the UU or TT mismatch by the triaminotriazine units coupled with a general duplex binder through either A-form RNA-bisamidinium groove complexation (**4**) or acridine intercalation (**5–6**). Acridineintercalating ligands (**7–13**) and bisamidinium-based ligands (**14–18**) were mono- or difunctionalized with azide and alkyne functional groups (Chart 1) using straightforward syntheses (see supporting information for full synthetic procedures and characterization). In addition to ligands **7–18** that are designed to complex d(CTG)^{exp} and r(CUG)^{exp}, mono- and diazides **19–36** were chosen to increase both the size and structural diversity of the library and build on known scaffolds of DNA- and RNA-targeting ligands.

DNA/RNA-Templated Click Selection.

Charts 1 and 2 contain four alkyne-, six amine-, and 23 azide-containing compounds. Considering that the alkyne-azide cycloaddition may produce 1,4- or 1,5-adducts and that the propiolamide-based ligands can undergo Michael addition with amines, there are 196 possible AB-type dimeric adducts and 5,219 ABA/ABC-type trimeric products. Although oligomeric or even polymeric products might form from the difunctionalized molecules, the length of the DNA and RNA templates, d(CTG)₁₆ and r(CUG)₁₆ was chosen to favor dimeric and trimeric products. For example, ligand **4** was designed to complex three consecutive UU mismatch sites whereas ligands **5** and **6** only require one UU or TT mismatch site.

In considering methods for monitoring the template-assisted click reaction, we chose MALDI-TOF mass spectrometry. The overall approach to screening is shown schematically in Figure 3. Equimolar amounts of all compounds were mixed together with the d(CTG)₁₆ or

r(CUG)₁₆ template, incubated for up to 7 d, and analyzed by MALDI-TOF MS. Negative screens were performed in the same manner but without template or with the following nucleic acid sequences: a standard, random 24-mer DNA duplex, a DNA duplex containing d(CTG-CAG)₁₆, r(CCUG)₆, and tRNA (Figure S3). The interest in the r(CCUG) tetranucleotide repeat arises from its role in myotonic dystrophy type 2 and that it contains mismatched pyrimidines. Furthermore, both acridine and bisamidinium ligands structurally related but with different recognition motifs to those studied here recognize r(CCUG)^{exp}.^{39,40}

As shown in Figure 4a and Figure S3, no significant peaks corresponding to click products appeared after 7 d of incubation in the absence of RNA/DNA or in the presence of r(CCUG)₈ or tRNA, respectively. In the presence of a fully paired 24-mer DNA duplex or a paired DNA duplex containing d(CTG-CAG)₁₆, small amounts of adducts were detected after 1 d (Figures S3a and S3b). In contrast, multiple intense peaks in the spectra were observed after 1 d in selections with d(CTG)₁₆ and r(CUG)₁₆ (Figure 4b and c, Table S1) and these peaks grew in intensity over the 7 d incubation. MALDI MS peaks at *m/z* 1027.4, 1084.4, 1215.6 were identified as the click products **D(7+8)**, **D(7+13)**, **D(7+17)**, respectively (Figure 4d). Some potential combinations of ligands gave very similar calculated *m/z* values (Table S1). To firmly establish the structure, a templated selection was performed without alkyne **7**. This selection yielded no click products even after a 7 d incubation (Figure S4). The combined results suggest that the products at *m/z* 1041.4, 1099.5, 1141.5 are **D(7+9)**, **D(7+19)**, and **D(7+10)**, respectively (Figure 4d). The results of the selection support the notion that the products are the result of a template-assisted click reaction, a form of target-guided synthesis. Thus, no products are observed in the absence of the DNA or RNA template. There are clearly other factors that help determine the product outcome. Most clearly, the reactivity of the alkyne is critical with no triazole products observed from unactivated alkynes such as **14** and **15**. Although the activated propiolamide unit is required, it is not sufficient because alkyne **16** was not observed to undergo template-assisted click (Figure S4). The fact that the acridine-melamine series is mostly found in the hit compounds does suggest that stronger binders may out compete weaker binders. Finally, there is a limited dataset that allowed the effect of chain-length to be examined. Somewhat surprisingly, it appears that chain-length is a less important factor than might be expected with azides **8**, **9**, and **10** all showing similar intensity peaks in the MALDI-MS. Without better quantification of the relative amounts of these products and examining the 1,4- to 1-5-triazole ratio, it isn't possible to speculate further on the origin of this effect.

Dimeric Ligand D(7+9) Binds Tightly to Expanded r(CUG) and d(CTG) Making Specific Contacts with Base Mismatches.

With several template-assisted selection hit compounds in hand, we chose dimer **D(7+9)** resulting from the reaction between alkyne **7** and azide **9** for further study as a potential multitarget binding and lead therapeutic agent. This pair was selected for additional testing because it was observed as one of the most prominent peaks in all three of the MALDI-TOF MS selections using d(CTG)₁₆ and r(CUG)₁₆. Additionally, in the HPLC analysis of the template-assisted click reaction with alkyne **7** and azide **9** with r(CUG)₁₆ template, both ligands were almost completely consumed, and a peak for **D(7+9)** was observed after 1 d

incubation compared to negligible change in the peak area with no template present even after 7 d (Figure S5).

Dimeric ligand **D(7+9)** was independently synthesized by copper-catalyzed azide-alkyne click (CuAAC). Note that the copper-catalyzed click reaction gives the 1,4-triazole exclusively, whereas the template-assisted click can produce the 1,4- or 1,5-isomer. Because the MALDI MS method does not distinguish the isomers, the click reaction of **7** and **9** was monitored by HPLC which showed the major isomeric product to be the 1,4-triazole (Figure S6). The relative binding constants for complexation between ligands **7** and **D(7+9)** to d(CTG)₁₆ and r(CUG)₁₆ was determined using a reverse fluorescence titration (Figure S9 and Figure S10). Dimer **D(7+9)** binds to d(CTG)₁₆ with $K_{D,app} = 23 \pm 1$ nM and r(CUG)₁₆ with $K_{D,app} = 230 \pm 10$ nM. Thus, the dimer binds its DNA and RNA targets ca. 21- and 15-fold more tightly than does monomer **7** that binds to d(CTG)₁₆ with $K_{D,app} = 350 \pm 27$ nM and r(CUG)₁₆ with $K_{D,app} = 4820 \pm 310$ nM. The dimer formed on the DNA or RNA template appears to be a stronger binder than the monomers alone but does not bind as tightly as one might expect for a multivalent interaction.²⁵ This could reflect a less than ideal linkage between the two monomers.

We next sought to investigate the specific interactions that would result in selective template-assisted reactions. Docking experiments with alkyne **7** and azide **9** bound to r(CUG)₁₂ duplex (PDB: 3GM7) were performed using a potential mode of binding where the triaminotriazine unit serves to recognize UU mismatches and the acridine unit acts as an intercalator (Figure 5a). The triaminotriazine units were posed to interact with the UU mismatch through hydrogen bonds to one U of the mismatch whereas the other U is weakly paired or flipped out. The acridine moiety was docked to intercalate into the hairpin and stack with a CG base pair below. Placing **7** and **9** in consecutive CUG sites brought the reaction partners into proximity (ca. 3.2 Å) with some flexibility in the linker, expected to allow them to undergo cyclization to give the triazole product **D(7+9)**.

The ligand complex with RNA was examined by NMR using an RNA hairpin containing two consecutive UU mismatches. We were specifically interested in investigating the changes in UU mismatches upon ligand binding and global changes by intercalation. There are several published structures of r(CUG) repeats that can serve as a reference point for this study to better understand structural changes upon ligand binding.^{37,41,42} A 1D titration experiment was performed following the imino peak region of a 1H NMR (e.g., 10–15 ppm) in D₂O (Figure 5b). Addition of **D(7+9)** disrupts the base-pairing of the UU mismatches as shown by loss of the peaks ~10.25 ppm. A 3:1 ratio of ligand to RNA is needed to fully disrupt the UU mismatch, which may be the result of poor solubility of the compound. Indeed, higher concentrations of compound showed observable precipitation from solution even when using 5% DMSO. Nonetheless, disruption of UU mismatches supports the model where the triaminotriazine recognition unit pairs with U bases in CUG sites. Further, 2D HSQC with ¹³C- and ¹⁵N-labelled RNA hairpin show clear bound and free states for UU mismatches. The new peak for the bound state is consistent with hydrogen bonding from a U base to a triaminotriazine recognition unit (Figure 5c). The combined solution NMR results support that **D(7+9)** binds to consecutive CUG repeats and makes specific interactions with the UU mismatches of the hairpin. The model offers new evidence for multivalent targeting

of trinucleotide repeat RNA and DNA through interaction with particular bases in a sequence-selective manner.

Ligand **D(7+9)** is a Dual Targeting Agent Functioning as a Potent Inhibitor of Both d(CTG)^{exp} Transcription and MBNL-r(CUG)^{exp} Sequestration.

The inhibition of d(CTG)^{exp} transcription has emerged as potential therapeutic strategy to treat DM1. Thus, we^{22–24} and Berglund^{21,43,44} reported that the formation of r(CUG)^{exp} can be suppressed by agents that bind d(CTG)^{exp} *in vitro*, in cells, and *in vivo*. To determine whether the DNA template-assisted selection product **D(7+9)** could function in this manner, a linearized T7-promoter-containing d(CTG·CAG)₇₄ plasmid was incubated with ligands **7**, **9**, and **D(7+9)** followed by *in vitro* transcription with T7 RNA polymerase.^{24,45} Ligand **D(7+9)** led to a maximum of 80–90% transcription inhibition at 1–5 μ M indicating a significant suppression of r(CUG)₇₄ levels (Figure 6 and S11). In contrast, treatment with monomeric ligands **7** and **9** up to 25 μ M showed no significant effect on the production of r(CUG)₇₄ relative to the untreated sample (Figure 6b). This result suggests that the dimeric click product formed *in situ* on the DNA target is a more potent transcription inhibitor than its monomeric counterpart. Interestingly, for **D(7+9)**, transcription inhibition was observed over a narrow concentration window. One reasonable explanation is a cooperative binding of the polymerase to d(CTG·CAG) and the compound to the d(CTG) hairpin.

The ability of monomer **7** and **D(7+9)** to inhibit MBNL sequestration of r(CUG)^{exp} was evaluated using an electrophoretic mobility shift assay (EMSA). A truncated MBNL1 protein containing four zinc finger motifs of MBNL1 and a hexa-His tag was employed.¹⁴ This truncated MBNL1 and radiolabeled-r(CUG)₁₂ form a tight complex with a K_D of 9 ± 1 nM (Figure S7). Using EMSA and ligand **7**, an IC₅₀ value of $26,000 \pm 2,000$ nM was determined with a calculated K_I of $3,300 \pm 500$ nM, whereas **D(7+9)** was ca. 144-fold more potent, with IC₅₀ and K_I values of 180 ± 5 nM and 23 ± 3 nM, respectively (Figure 7 and S8). Collectively the biophysical and biochemical results suggest the dimeric click products have potential as tight-binding, sequence-selective, multitarget therapeutic agents.

CONCLUSION

We have described an accelerated strategy for finding agents that target the DNA and RNA that causes myotonic dystrophy type 1 (DM1). The approach combines elements of FBDD and TGS with a positive cross screen to discover multitargeting lead agents for DM1. Both template-assisted selections (with d(CTG) and r(CUG) repeats) produced only a small number of the potential products that could form. Importantly, selectivity was observed with product formation in the presence of the target DNA or RNA but not with other oligonucleotides such as tRNA or a random dsDNA. The selection appears to favor tight-binding ligands and those with more reactive clickable functional groups and was non-discriminating with respect to the length of the linker between the ligand and the functional group. The success of the templated-click selection using the toxic RNA and DNA in DM1 supports future applications to finding multitarget agents for other diseases caused by unstable trinucleotide repeats, such as Huntington's disease, fragile X syndrome, and Friedrich's ataxia.

As alluded to above, using a single mixture of fragments in the screen will likely lead to the strongest binders being selected. This could be advantageous from the perspective of discovering the tightest binding dimers and trimers, but it neglects other considerations such as toxicity. For example, the acridine series of ligands (**7–14**) are tighter binding but the bisamidinium series (**14–18**) are generally less toxic. Another limitation arises from the use of fragments that, with the exception of these two designed ligand types, are weak and nonspecific binders. For this reason, although the approach described here demonstrates the potential of the fragment-based screen, it did not lead to a new class of ligands or one that outperforms previous dimers prepared by rational design. To move beyond this proof of principle demonstration, a broader discovery effort might use new RNA fragment libraries that offer a higher level of diversity.^{46,47}

Beyond the general approach described here for the discovery of more potent *in vitro* inhibitors and multitarget agents, this same general protocol could be applied to DM1 cell culture. A cell-based approach adds the challenge of cellular and nuclear permeability, a critical aspect for drug development. More importantly, it opens the prospect of a self-assembling therapeutic strategy using smaller molecules whose improved pharmacokinetic properties allow them to reach the target that facilitates the synthesis of the active agent. Such an approach will require a new method of product analysis and a faster and more biocompatible click reaction. Efforts to accomplish these goals are currently underway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding Sources

Financial support was provided by the National Institutes of Health (R01AR069645 to S.C.Z.) and the National Science Foundation through the Graduate Research Fellow Program under Grant No. DGE - 1746047 (S.E.B. at University of Illinois). L.D.H. is a member of the NIH Chemistry-Biology Interface Training Grant (NRSA 1-T-32-GM070421). M.T. was supported by NIH grant P41GM103399. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH grant P41GM103399.

REFERENCES

- (1). Scott DE, Coyne AG, Hudson SA, and Abell C (2012) Fragment-Based Approaches in Drug Discovery and Chemical Biology. *Biochemistry* 51, 4990–5003. [PubMed: 22697260]
- (2). Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, and Jhoti H (2016) Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discov* 15, 605–619. [PubMed: 27417849]
- (3). Mello J. da F. R. e., Gomes RA, Vital-Fujii DG, Ferreira GM, and Trossini GHG (2017) Fragment-Based Drug Discovery as Alternative Strategy to the Drug Development for Neglected Diseases. *Chem. Biol. Drug Des* 90, 1067–1078. [PubMed: 28547936]
- (4). Erlanson DA, Davis BJ, and Jahnke W Fragment-Based Drug Discovery: Advancing Fragments in the Absence of Crystal Structures. *Cell Chem. Biol* 26, 9–15. [PubMed: 30482678]
- (5). Warner KD, Homan P, Weeks KM, Smith AG, Abell C, and Ferré-D'Amaré AR (2014) Validating Fragment-Based Drug Discovery for Biological RNAs: Lead Fragments Bind and Remodel the TPP Riboswitch Specifically. *Chem. Biol* 21, 591–595. [PubMed: 24768306]
- (6). Rideout D, Calogeropoulou T, Jaworski J, and McCarthy M (1990) Synergism through Direct Covalent Bonding between Agents: A Strategy for Rational Design of Chemotherapeutic Combinations. *Biopolymers* 29, 247–262. [PubMed: 2328289]

- (7). Mamidyala SK and Finn MG (2010) In Situ Click Chemistry: Probing the Binding Landscapes of Biological Molecules. *Chem. Soc. Rev* 39, 1252–1261. [PubMed: 20309485]
- (8). Ramström O and Lehn JM (2002) Drug Discovery by Dynamic Combinatorial Libraries. *Nat. Rev. Drug Discov* 1, 26–36. [PubMed: 12119606]
- (9). Bosc D, Camberlein V, Gealageas R, Castillo-Aguilera O, Deprez B, and Deprez-Poulain R (2020) Kinetic Target-Guided Synthesis: Reaching the Age of Maturity. *J. Med. Chem* 63, 3817–3833. [PubMed: 31820982]
- (10). Reddy K, Jenquin JR, Cleary JD, and Berglund JA (2019) Mitigating RNA Toxicity in Myotonic Dystrophy Using Small Molecules. *Int. J. Mol. Sci* 20, 4017–4037.
- (11). Konieczny P, Selma-Soriano E, Rapisarda AS, Fernandez-Costa JM, Perez-Alonso M, and Artero R (2017) Myotonic Dystrophy: Candidate Small Molecule Therapeutics. *Drug Discov. Today* 22, 1740–1748. [PubMed: 28780071]
- (12). Thornton CA, Wang E, and Carrell EM (2017) Myotonic Dystrophy approach to Therapy. *Curr. Opin. Genet. Dev* 44, 135–140. [PubMed: 28376341]
- (13). Day JW and Ranum LPW (2005) RNA Pathogenesis of the Myotonic Dystrophies. *Neuromuscul. Disord* 15, 5–16. [PubMed: 15639115]
- (14). Yuan Y, Compton SA, Sobczak K, Stenberg MG, Thornton CA, Griffith JD, and Swanson MS (2007) Muscleblind-like 1 Interacts with RNA Hairpins in Splicing Target and Pathogenic RNAs. *Nucleic Acids Res.* 35, 5474–5486. [PubMed: 17702765]
- (15). Singh RK and Cooper TA (2012) Pre-mRNA Splicing in Disease and Therapeutics. *Trends Mol. Med* 18, 472–482. [PubMed: 22819011]
- (16). Raaijmakers RHL, Ripken L, Ausems CRM, and Wansink DG (2019) CRISPR / Cas Applications in Myotonic Dystrophy : Expanding Opportunities. *Int. J. Mol. Sci* 20, 3689–3706.
- (17). van Agtmaal EL, André LM, Willemse M, Cumming SA, van Kessel IDG, van den Broek WJAA, Gourdon G, Furling D, Mouly V, Monckton DG, and Wansink DG (2017) CRISPR/Cas9-Induced (CTG-CAG)_n Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing. *Mol. Ther* 25, 24–43. [PubMed: 28129118]
- (18). Wang Y, Hao L, Wang H, Santostefano K, Thapa A, Cleary J, Li H, Guo X, Terada N, Ashizawa T, Xia G (2018) Therapeutic Genome Editing for Myotonic Dystrophy Type 1 Using CRISPR/ Cas9. *Mol. Ther* 26, 2617–2630. [PubMed: 30274788]
- (19). Provenzano C, Cappella M, Valaperta R, Cardani R, Meola G, Martelli F, Cardinali B and Falcone G (2017) CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. *Mol. Ther. - Nucleic Acids* 9, 337–348. [PubMed: 29246312]
- (20). Dastidar S, Ardui S, Singh K, Majumdar D, Nair N, Fu Y, Reyon D, Samara E, Gerli MFM, Klein AF, and De Schrijver W (2018) Efficient CRISPR/Cas9-Mediated Editing of Trinucleotide Repeat Expansion in Myotonic Dystrophy Patient-Derived IPS and Myogenic Cells. *Nucleic Acids Res.* 46, 8275–8298. [PubMed: 29947794]
- (21). Siboni RB, Nakamori M, Wagner SD, Struck AJ, Coonrod LA, Harriott SA, Cass DM, Tanner MK, and Berglund JA (2015) Actinomycin D Specifically Reduces Expanded CUG Repeat RNA in Myotonic Dystrophy Models. *Cell Rep.* 13, 2386–2394. [PubMed: 26686629]
- (22). Serrano JF, Lee J, Daniel Curet L, Hagler LD, Bonson SE, Schuster EJ, and Zimmerman SC (2019) Development of Novel Macrocyclic Small Molecules That Target CTG Trinucleotide Repeats. *Bioorganic Med. Chem* 27, 2978–2984.
- (23). Lee JY, Bai Y, Chembazhi UV, Peng S, Yum K, Luu LM, Hagler LD, Serrano JF, Edwin Chan HY, Kalsotra A, and Zimmerman SC (2019) Intrinsically Cell-Penetrating Multivalent and Multitargeting Ligands for Myotonic Dystrophy Type 1. *Proc. Natl. Acad. Sci. U. S. A* 116, 8709–8714. [PubMed: 30975744]
- (24). Nguyen L, Luu LM, Peng S, Serrano JF, Chan HYE, and Zimmerman SC (2015) Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc* 137, 14180–14189. [PubMed: 26473464]
- (25). Jahromi AH, Fu Y, Miller KA, Nguyen L, Luu LM, Baranger AM, and Zimmerman SC (2013) Developing Bivalent Ligands to Target CUG Triplet Repeats, the Causative Agent of Myotonic Dystrophy Type 1. *J. Med. Chem* 56, 9471–9481. [PubMed: 24188018]

- (26). Luu LM, Nguyen L, Peng S, Lee J, Lee HY, Wong C-H, Hergenrother PJ, Chan HYE, and Zimmerman SC (2016) A Potent Inhibitor of Protein Sequestration by Expanded Triplet (CUG) Repeats That Shows Phenotypic Improvements in a *Drosophila* Model of Myotonic Dystrophy. *ChemMedChem* 11, 1428–1435. [PubMed: 27245480]
- (27). Gareiss PC, Sobczak K, McNaughton BR, Palde PB, Thornton CA, and Miller BL (2008) Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA-MBNL1 Interaction in Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1). *J. Am. Chem. Soc* 130, 16254–16261. [PubMed: 18998634]
- (28). Pushechnikov A, Lee MM, Childs-Disney JL, Sobczak K, French JM, Thornton CA, and Disney MD (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. [PubMed: 19552411]
- (29). Grimster NP, Stump B, Fotsing JR, Weide T, Talley TT, Yamauchi JG, Nemezc Á, Kim C, Ho K-Y, Sharpless KB, and Taylor P (2012) Generation of Candidate Ligands for Nicotinic Acetylcholine Receptors via in Situ Click Chemistry with a Soluble Acetylcholine Binding Protein Template. *J. Am. Chem. Soc* 134 (15), 6732–6740. [PubMed: 22394239]
- (30). Bhardwaj A, Kaur J, Wuest M, and Wuest F (2017) In Situ Click Chemistry Generation of Cyclooxygenase-2 Inhibitors. *Nat. Commun* 8, 1–13. [PubMed: 28232747]
- (31). Umuhire Juru A, Cai Z, Jan A, and Hargrove AE (2020) Template-Guided Selection of RNA Ligands Using Imine-Based Dynamic Combinatorial Chemistry. *Chem. Commun* 56, 3555–3558.
- (32). McNaughton BR, Gareiss PC, and Miller BL (2007) Identification of a Selective Small-Molecule Ligand for HIV-1 Frameshift-Inducing Stem-Loop RNA from an 11,325 Member Resin Bound Dynamic Combinatorial Library. *J. Am. Chem. Soc* 129, 11306–11307. [PubMed: 17722919]
- (33). Rzuczek SG, Park H, and Disney MD (2014) A Toxic RNA Catalyzes the in Cellulo Synthesis of Its Own Inhibitor. *Angew. Chemie - Int. Ed* 53, 10956–10959.
- (34). Poulin-Kerstein AT, and Dervan PB (2003) DNA-Templated Dimerization of Hairpin Polyamides. *J. Am. Chem. Soc* 125, 15811–15821. [PubMed: 14677972]
- (35). Cornilescu G, Didychuk AL, Rodgers ML, Michael LA, Burke JE, Montemayor EJ, Hoskins AA, and Butcher SE (2016) Structural Analysis of Multi-Helical RNAs by NMR-SAXS/WAXS: Application to the U4/U6 Di-SnRNA. *J. Mol. Biol* 428, 777–789. [PubMed: 26655855]
- (36). Sashital DG, Venditti V, Angers CG, Cornilescu G, and Butcher SE (2007) Structure and Thermodynamics of a Conserved U2 SnRNA Domain from Yeast and Human. *Rna* 13, 328–338. [PubMed: 17242306]
- (37). Montemayor EJ, Virta JM, Hagler LD, Zimmerman SC, and Butcher SE (2019) Structure of an RNA Helix with Pyrimidine Mismatches and Cross-Strand Stacking. *Acta Crystallogr. Sect. F Struct. Biol. Commun* 75, 652–656. [PubMed: 31584014]
- (38). Crenshaw JM, Graves DE, and Denny WA (1995) Interactions of Acridine Antitumor Agents with DNA: Binding Energies and Groove Preferences. *Biochemistry* 34, 13682–13687. [PubMed: 7577959]
- (39). Wong CH, Fu Y, Ramisetty SR, Baranger AM, and Zimmerman SC (2011) Selective Inhibition of MBNL1-CCUG Interaction by Small Molecules toward Potential Therapeutic Agents for Myotonic Dystrophy Type 2 (DM2). *Nucleic Acids Res.* 39, 8881–8890. [PubMed: 21768123]
- (40). Nguyen L, Lee JY, Wong CH, and Zimmerman SC (2014) Small Molecules That Target the Toxic RNA in Myotonic Dystrophy Type 2. *ChemMedChem* 9, 2455–2462. [PubMed: 24938413]
- (41). Kiliszek A, Kierzek R, Krzyzosiak WJ, and Rypniewski W (2009) Structural Insights into CUG Repeats Containing the ‘Stretched U–U Wobble’: Implications for Myotonic Dystrophy. *Nucleic Acids Res.* 37, 4149–4156. [PubMed: 19433512]
- (42). Mooers BHM, Logue JS, and Berglund JA (2005) The Structural Basis of Myotonic Dystrophy from the Crystal Structure of CUG Repeats. *Proc. Natl. Acad. Sci* 102, 16626–16631. [PubMed: 16269545]
- (43). Coonrod LA, Nakamori M, Wang W, Carrell S, Hilton CL, Bodner MJ, Siboni RB, Docter AG, Haley MM, Thornton CA, and Berglund JA (2013) Reducing Levels of Toxic RNA with Small Molecules. *ACS Chem. Biol* 8, 2528–2537. [PubMed: 24028068]

- (44). Siboni RB, Bodner MJ, Khalifa MM, Docter AG, Choi JY, Nakamori M, Haley MM, and Berglund JA (2015) Biological Efficacy and Toxicity of Diamidines in Myotonic Dystrophy Type 1 Models. *J. Med. Chem* 58, 5770–5780. [PubMed: 26103061]
- (45). Miller JW (2000) Recruitment of Human Muscleblind Proteins to (CUG)_n Expansions Associated with Myotonic Dystrophy. *EMBO J.* 19, 4439–4448. [PubMed: 10970838]
- (46). Morgan BS, Sanaba BG, Donlic A, Karloff DB, Forte JE, Zhang Y, and Hargrove AE (2019). R-BIND: an interactive database for exploring and developing RNA-targeted chemical probes. *ACS Chem. Biol* 14, 2691–2700. [PubMed: 31589399]
- (47). Abulwerdi FA, and Schneekloth JS Jr (2016). Microarray-based technologies for the discovery of selective, RNA-binding molecules. *Methods*, 103, 188–195. [PubMed: 27109057]

Figure 1. Schematic illustration of DM1 pathobiology and approaches for small molecule inhibition (adapted from ref. 23). (a) Expanded d(CTG·CAG) leads to r(CAG)^{exp} and r(CUG)^{exp} transcripts, the latter sequestering MBNL1. (b) Multitarget, small molecule inhibition of (left) transcription by binding d(CTG·CAG) or d(CTG)^{exp} hairpin and (right) MBNL1 sequestration by complexation with r(CUG) hairpins results in stalled transcription yielding no RNA products and free MBNL protein. (c) TGS can be used to find multivalent inhibitors that are single or multitarget agents. Small molecules from a TGS library bind to the target of interest and undergo a template-assisted covalent coupling reaction to yield more potent inhibitors.

Figure 2.

Model Template-Assisted Click Reaction. (a and b) Acridine intercalators **1** and **2** were docked onto DNA to find a minimum linker length for proximity-induced [3+2] cyclization to form triazole product **3**. (c) HPLC-monitoring of reaction progress by depletion of azide **1** and alkyne **2**. Area under the curve was calculated by Shimadzu Lab Solutions software. The concentration calculated from a standard curve for each ligand. Two independent runs gave values within 2–20%.

Figure 4.

Representative MALDI-TOF MS of click selections with (a) no DNA/RNA (b) r(CUG)₁₆, (c) d(CTG)₁₆. (d) MALDI-TOF MS peaks map to r(CUG)₁₆-templated click dimers. Concentration of each compound = 25 μ M and [DNA or RNA template] = 12 μ M. Molecular weight of monomers range from 200–700 Da, whereas those of conjugate products are between 800 and 1500 Da. The selections were performed at pH 7.0 and 37 $^{\circ}$ C in triplicate.

Figure 5.

(a) Docking experiments show the close proximity of alkyne **7** and azide **9** that would afford triazole product **D(7+9)** by a template-assisted click reaction on the $r(\text{CUG})_{12}$ target. (b and c) NMR solution experiments with **D(7+9)** and a $r(\text{CUG})_2$ -containing hairpin show specific and selective binding of the triaminotriazine recognition unit to UU mismatches. (b) UU mismatches are disrupted upon titration of **D(7+9)** to RNA and (c) new peaks emerge for hydrogen bonding between UU mismatches and triaminotriazine.

Figure 6.

In vitro transcription inhibition of d(CTG·CAG)₇₄ by ligands **7**, **9**, and **D(7+9)**. (a) Representative gels post-stained with EtBr and HIV fs used as an external standard. (b) Plot of the normalized level of r(CUG)₇₄ transcription inhibition versus concentration of ligands. Error bars represent the standard error of the mean for three independent experiments.

Figure 7.

Inhibition of MBNL1-r(CUG)₁₂ protein-RNA complex by **7** and **D(7+9)** by EMSA. (a) EMSA gel of compound **D(7+9)**, dosed from 5 μM to 1.2 nM. (b) IC₅₀ value was determined by fitting the plot of the normalized fraction of RNA bound vs. the ligand concentration with the equation: $y = ax^b/(IC_{50}^b + x^b)$, where y is the fraction of RNA bound and x is the concentration of the compound, $a = y_{max} - y_{min}$, b is a Hill slope. The inhibition constant K_I was calculated using the equation: $K_I = IC_{50} \times K_D / [protein]_{total}$, where $[protein]_{total} = 70$ nM. Error bars represent the standard error of the mean for three independent experiments.

Chart 1.
Bisamidinium and Acridine-Based Click Fragments

Chart 2.
Additional Fragments for Click Library