APPROACHES TO THE ASSEMBLY OF POTENT THERAPEUTIC AGENTS FOR THE TREATMENT OF MYOTONIC DYSTROPHY

BY

LONG M. LUU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Steven C. Zimmerman, Chair and Director of Research
Professor Paul J. Hergenrother
Professor M. Christina White
Assistant Professor Auinash Kalsotra

ABSTRACT

Myotonic dystrophy type 1 (DM1), the most common form of adult-onset muscular dystrophy, is an incurable neuromuscular disease. DM1 is caused by an expansion of the CTG repeat, whose RNA transcript sequesters the MBNL1 protein into nuclear foci, leading to the misregulation of various pre-mRNAs.

Based on a published x-ray structure of a $r(CUG)_6$ duplex, a benzamidinium ligand was rationally designed and found to selectively bind rCUG repeats and inhibit the MBNL1-$rCUG^{exp}$ interaction with a low micro-molar inhibition potency. Thus, this bisamidinium RNA groove binder represented one of the most promising lead compounds for DM1 treatment. The development of the bisamidinium analogs containing functional substituents will be discussed in Chapter 2.

Chapter 3 describes the development of a bisamidinium dimeric ligand as a possible compound for the treatment of DM1. This ligand reveals both the promise and challenge for the small molecule approach. The agent is a potent inhibitor of the MBNL1-$rCUG^{exp}$ both *in vitro* and in DM1 model cells, and is able to improve two distinct disease phenotypes in a DM1 *Drosophila* model. However, its relatively low maximum tolerated dose in mice and limited cell uptake provide insights into directions for future development.

In addition, a powerful selection method that uses the target DNA or RNA to select their own binders will be discussed in Chapter 4. Compounds discovered using this method are able to undergo the alkyne-azide cycloaddition on the target DNA and RNA. The selectivity of these compounds along with their click products to other DNA/RNA sequences can also be readily accessed. The click products formed have better efficacy and can be used as multi-target agents.

This work is dedicated to my family and my advisors

ACKNOWLEDGEMENTS

From an underdog chemistry student in high school to a defending PhD student today, I am very thankful for the many people who supported and guided me along the way. First, I would like to thank Prof. Steven Zimmerman for being an amazing research advisor. His establishment of the collaboration between UIUC and Hanoi University of Science, which has allowed myself and many other Vietnamese students to pursue doctoral studies in a world-leading institution, has impacted my life tremendously. I am grateful to Prof. Paul J. Hergenrother for being a great collaborator. I would like to recognize Prof. M. Christina White and Prof. Auinash Kalsotra for serving on my committee and giving me valuable advice.

My research could not have been completed without the help of my colleagues and our collaborators, especially Dr. Lien Nguyen, JuYeon Lee, Prof. Edwin Chan, and Dr. Chun-Ho Wong. I am also fortunate to have worked with young talented scientists, including Nacho Vergara and Emily Tonogai, who assisted me with the synthesis of various compounds. Beyond being great chemists, my lab mates in the Zimmerman group have truly become good friends, from whom I have learned a lot both scientifically and culturally. I am also grateful to have many good Vietnamese friends, who made my time in Champaign-Urbana enjoyable and exciting. I would like to thank Prof. Scheeline and Prof. Lisy for their guidance and support in both my academic career and personal life.

I cannot thank my parents enough for their unwavering support. My daughter has been the biggest distraction to my work, but at the end of the day, all the stress and worry disappear after seeing her smile. Finally, I would like to send my biggest thanks to my wife, who has traversed this journey with me through all ups and downs.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Overview of Myotonic Dystrophy

Myotonic dystrophy (Dystrophia Myotonia, DM) is an autosomal dominant neuromuscular disorder, which affects 2−14 in 100,000 people worldwide.[1] DM patients have a wide range of symptoms, including muscle weakness (myopathy), trouble relaxing certain muscles (myotonia), progressive muscle wasting (atrophy), cataracts, cardiac defects, and insulin-dependent diabetes.[2,3] Myotonic dystrophy exists in two major types: type 1 (DM1), the most common form with a global prevalence of 1 per 8,000 people, and type 2 (DM2) with a high prevalence in Europe.[4]

1.2 Pathogenic Mechanisms of Myotonic Dystrophy Type 1

DM1 is caused by a progressive expansion of the trinucleotide CTG repeat ($dCTG^{exp}$) in the 3'-untranslated region of the dystrophia myotonica-protein kinase (*DMPK)* gene on chromosome 19q13.3.[5,6] Healthy individuals have a low number of CTG repeats, fewer than 35, whereas the number of repeats for a DM1 patient can range from 50 to thousands.[6] The length of CTG repeats directly correlates with the symptom severity and the age of onset.[7] Previously, there were three possible pathogenic mechanisms considered for DM1:[8] 1) haploinsufficiency of *DMPK*,[9] 2) decreased expression of neighboring genes, including *SIX5* and *DMAHP*,[10,11] and 3) a gain-of-function for the expanded RNA transcript ($rCUG^{exp}$).[2] Recent findings have provided evidence against the first two hypotheses, leaving the RNA gain-of-function as the most probable mechanism for therapeutic intervention.[2,12]

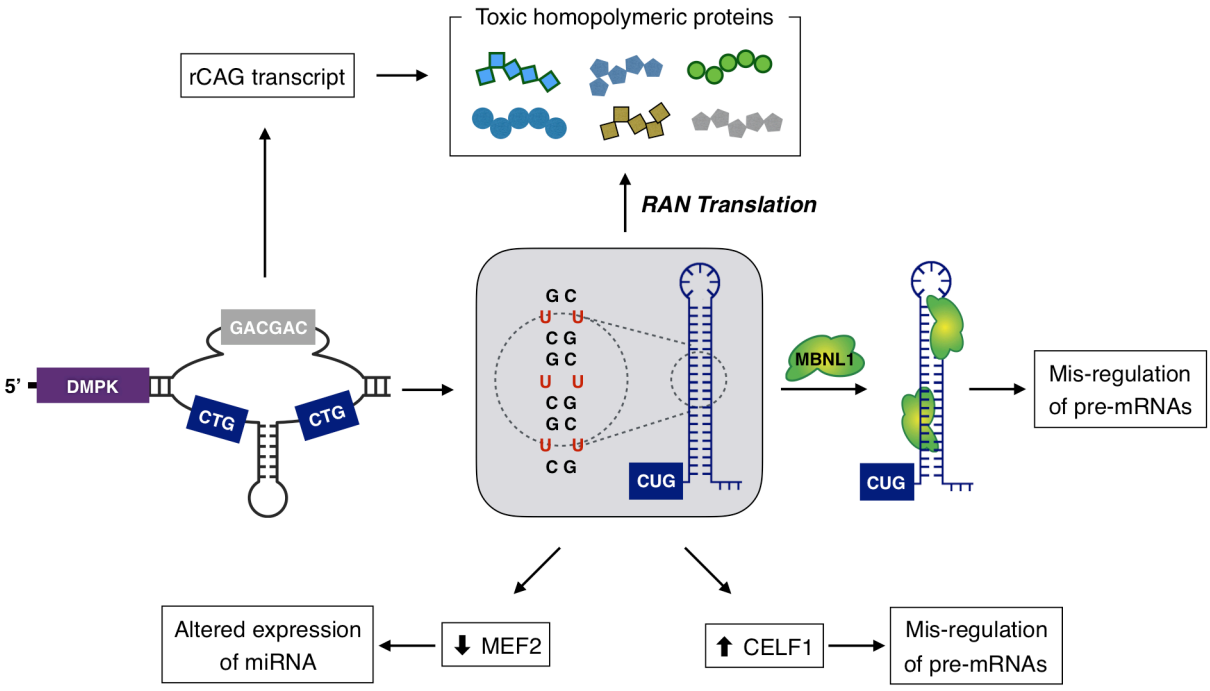


Figure 1. Pathogenic mechanisms of DM1.

The RNA gain-of-function model involves the sequestration of the muscleblind (MBNL) protein family of splicing regulators by expanded rCUG repeats ($rCUG^{exp}$) (Figure 1).[13,14] In DM1 cells, the co-localization of MBNL1 protein and rCUG repeats in nucleus results in the formation of nuclear foci.[15] It was found that more than 80% of aberrant alternative splicing events are directly related to the sequestration of MBNL1 by $rCUG^{exp}$ in a DM1 mouse model.[16] Many of the symptoms of DM1 result from these mis-splicing events.[17] For example, myotonia, insulin resistance, cardiac abnormalities, muscle weakness are caused by the aberrant splicing of chloride channel 1 (*ClC-1*),[18] insulin receptor (*IR*),[19] cardiac troponin T (*cTNT*),[8] and sarcoplasmic/endoplasmic reticulum Ca^{2+}-ATPase 1 (*SERCA1*),[20] respectively.

In contrast to MBNL1 loss-of-function, the expression of CUG-BP1 and ETR-3-like factor 1 (CELF1) is up-regulated in DM1 heart and skeletal muscle tissues and a DM1 mouse model.[21] The process of up-regulation is attributed to the protein kinase C-activated

hyperphosphorylation and stabilization of CUG-BP1.[22] Unlike MBNL1, CUG-BP1 binds short, single-stranded rCUG repeats, and does not co-localize with the toxic RNA to form nuclear foci.[15] In 2005, Cooper and coworkers found that transgenic mice overexpress 4−6 fold above the endogenous level of CUG-BP1 and have aberrant splicing of *cTNT*, *Mtmr1*, and *ClC-1* in heart and skeletal muscle tissues, suggesting the increase in CUG-BP1 level as one of the DM1 pathogenesis.[21]

Apart from these two disease mechanisms, Kalsotra *et al.* found that the significant reduction of MEF2 protein level results in altered miRNA and mRNA expression in heart tissue.[23] Recently, Ranum showed that both rCAG and rCUG repeats can undergo the repeat-associated non-ATG (RAN) translation, producing six toxic homopolymeric proteins.[24]

1.3 Therapeutic Approaches to the Treatment of Myotonic Dystrophy Type 1

Recent studies suggest that inhibiting the formation of the MBNL1-rCUG^{exp} complex could be a potential therapeutic approach for DM1.[3] In 2003, Swanson and coworkers showed that a muscleblind knockout mouse model (*Mbnl1*$^{\Delta 3/\Delta 3}$) exhibited phenotypic features associated with DM1, such as myotonia, cataract defect, and abnormal splicing of *ClC-1*, *cTNT*.[25] Conversely, the overexpression of MBNL1 in *HSA*^{LR} mice using adeno-associated virus mediated transduction was found to relieve myotonia and correct the missplicing of *ClC-1*, *TNNT3*, and *SERCA1* to nearly wild-type levels after 23 weeks.[26] Three possible approaches to the inhibition of MBNL1-rCUG^{exp} complex include (1) targeting RNA to displace MBNL1 protein, (2) targeting DNA to reduce rCUG transcript levels or suppress the repeat expansion, and (3) upregulating MBNL1 level.

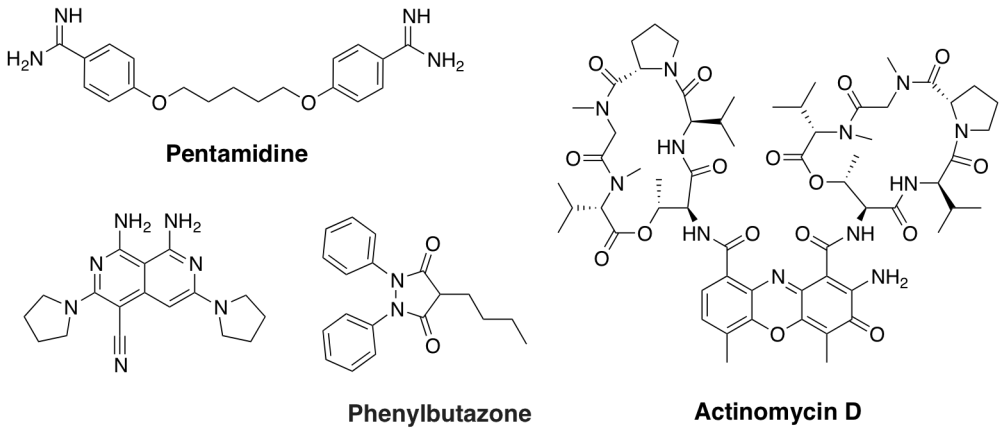


Figure 2. Possible small molecules for treatment of DM1.

Targeting RNA. RNA has become the central target in part because of the availability of high-resolution structures of rCUG repeats.[27,28] A number of RNA binding ligands, including antisense oligonucleotides and small molecules, have been developed over the past decade. Most of small molecule inhibitors of MBNL1 binding were discovered from screening methods, such as a pentamidine[29] from a library of 26 nucleic acid binding molecules, and a naphthyridine-containing compound[30] from the NIH Chemical Genomics Center (Figure 2). These compounds exhibit promising therapeutic properties, including dissolving the MBNL1-rCUG^{exp} nuclear foci and correcting splicing defects in DM1 model cells and mice. To improve the binding affinity of small molecules toward rCUG repeats and inhibition potency of MBNL1 complex formation, Disney and co-workers reported modularly assembled ligands by attaching several bisbenzimidazole Hoechst 33258 or kanamycin units on a peptoid backbone.[31,32] Moreover, upon photolysis, a rCUG binder conjugated with an *N*-hydroxypyridine-2(1*H*)-thione moiety is capable of producing hydroxyl radicals and cleaving rCUG repeats specifically.[33] In addition, a compound that binds and covalently conjugates to rCUG repeats was found to exhibit improved inhibition potency.[34] Although many small molecules have promising therapeutic properties, none have advanced into clinical trials. However, a 2'-O-methyl-phosphorothioate modified-

$(CAG)_7$ oligonucleotide developed by Thornton was shown to significantly reduce the toxic rCUG level in DM1 mouse models, and is currently being tested in phase 2 clinical trials.[35]

Transcription inhibition of dCTG repeats. Recently, this approach has emerged as an alternative therapeutic strategy for DM1, as it may lead to the reduction of toxic $rCUG^{exp}$ level.[36] In 2015, Berglund and coworkers reported that actinomycin D, an FDA-approved chemotherapeutic agent, is capable of reducing the rCUG transcript levels in DM1 model cells and mice at low nanomolar concentration.[37] In addition, a pyrrole-imidazole polyamide-chlorambucil conjugate was shown to induce the DNA alkylation of the target sequence and inhibit transcription of the CAG-CTG repeat.[38] However, the potential for expansion or contraction of the repeat regions upon binding to dCTG remains unknown.

Overexpression of MBNL1 protein. Another approach to the treatment of DM1 is overexpression of MBNL1. Recently, it was found that treatment with phenylbutazone, a nonsteroidal anti-inflammatory drug, resulted in the unregulated expression of MBNL1 myoblasts and DM1 mouse model, leading to the improvement of abnormal splicing events, muscle weakness, and muscle pathology.[39]

1.4 Development of Small Molecule Agents in the Zimmerman Group

Based on the available structures of rCUG repeats and using a rational design method, we have developed several small molecules that bind rCUG repeats and inhibit its interaction with MBNL1 protein (Figure 3, Table 1). It is known that rCUG repeats can form stable stem loop structures with U-U mismatches flanked by G-C and C-G base pairs.[27,28] Compounds **1−4** contain triaminotriazine rings that can serve as a Janus-wedge recognition unit capable of forming hydrogen bonds to the uracil-uracil (U-U) and thymine-thymine (T-T) mismatches.[40]

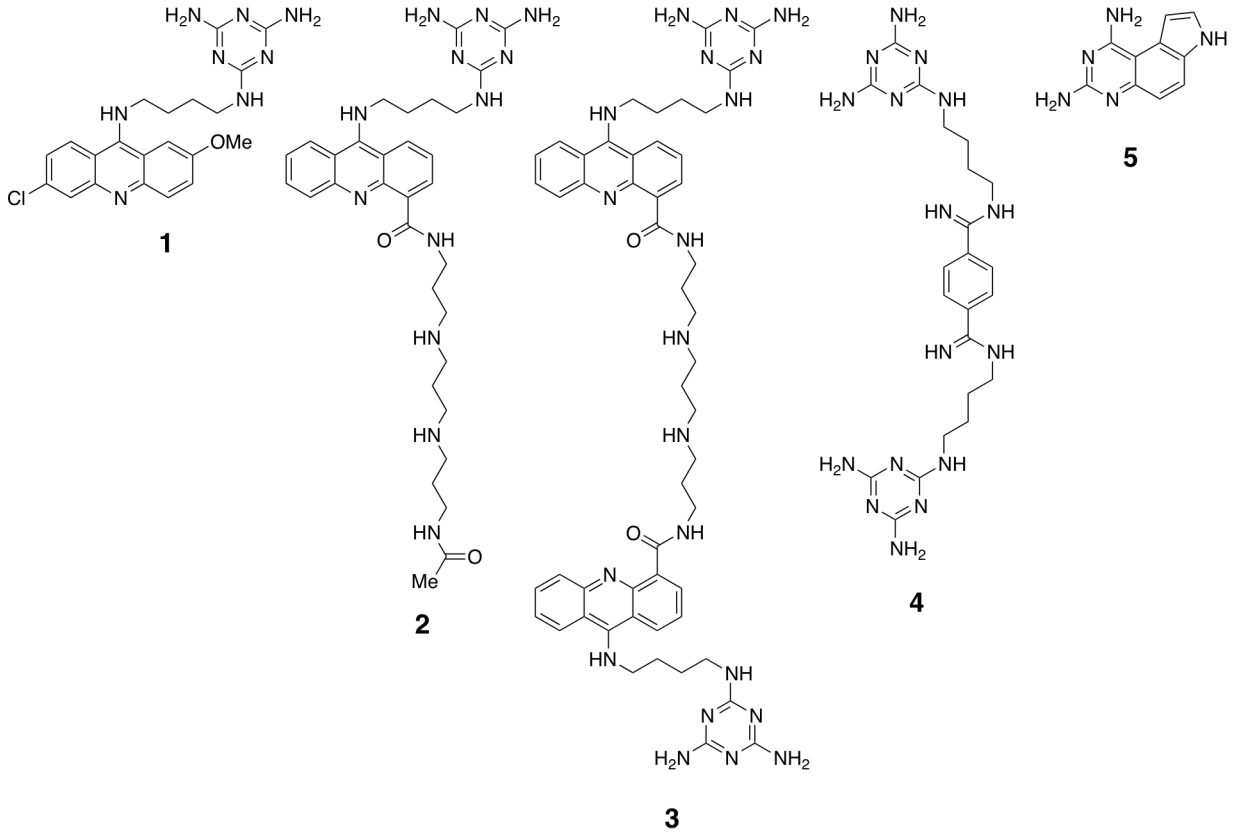


Figure 3. Five compounds developed in the Zimmerman group.

Ligand **1** is a melamine-acridine conjugate, in which the acridine intercalator helps drive binding by slipping between the melamine ring and a G-C base pair, while the triaminotriazine ring serves as the recognition unit.[41] Compound **1** is highly selective for T-T and U-U mismatches. It was suggested that the triaminotriazine ring in compound **1** recognizes the U-U mismatch in the RNA major groove, whereas it forms hydrogen bonds to the T-T mismatch in the DNA minor groove (Figure 4).[40] Compound **1** was one of the first small molecule inhibitors of the MNBL1-rCUG^{exp} complex to be reported in the literature. We previously discovered that the acridine-based ligands and MBNL1 can form a ternary complex with rCUG repeats, indicating that inhibition may be noncompetitive.[42] However, compound **1** suffers from the poor water solubility, an inability to penetrate the nuclear membrane, and high toxicity. Therefore, the efficacy of **1** in DM1 model cells was not studied.

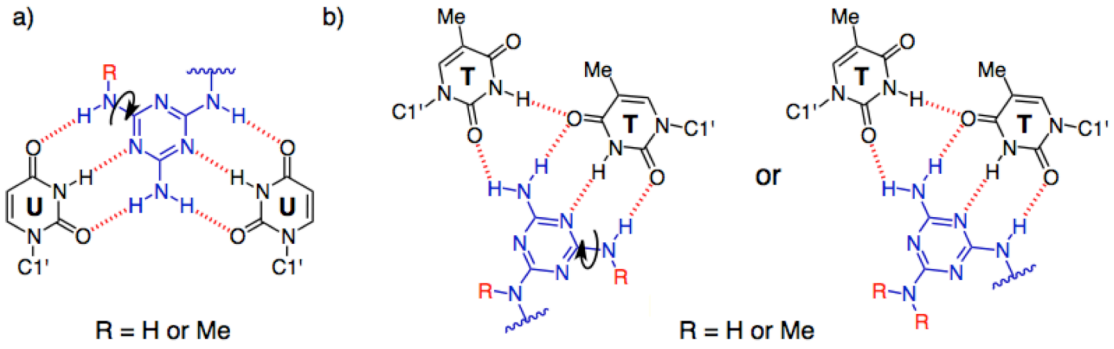


Figure 4. Proposed recognition of the triaminotriazine ring to (A) U-U mismatch in the RNA major groove and (B) T-T mismatch in the DNA minor groove. (Figure was adapted from Ref 40)

Table 1. Biological activities of ligands **1**–**5**. Dissociation constants (K_D) determined by (a) Isothermal Titration Calorimetry (ITC) and (b) reverse fluorescence titration of TAMRA-r(CUG)$_6$ to ligands. IC_{50} for the inhibition of MBNL1-r(CUG)$_{12}$ determined by (c) electrophoretic mobility shift assay (EMSA) and (d) surface plasmon resonance (SPR). K_I value was calculated using the equation: $K_I = IC_{50} \times K_D / [\text{protein}]_{total}$, where K_D is the dissociation constant of the MBNL1–RNA complex. Toxicity to (e) HeLa cells.

Comp.	Binding to CUG K_D, μM	Inhibition of MBNL1-CUG		Cytotoxicity[e] IC_{50}, μM	Foci dispersion	Splicing reversal
		IC_{50}, μM	K_I, μM			
1	0.43 ± 0.11[a], r(CUG)$_4$	46 ± 7[c]	7 ± 1	N/A	N/A	N/A
2	N/A	15 ± 2[d]	N/A	<10, 72 h	✓	✓
3	0.32 ± 0.02[b], r(CUG)$_6$	1.1 ± 0.1[d]	N/A	<50, 24 h	✓	N/A
4	8 ± 2[a], r(CUG)$_{12}$	115 ± 14[c]	8 ± 2	>100, 72 h	✓	✓
5	N/A	61 ± 15[c]	15 ± 3	~24, 72 h	N/A	N/A

To overcome these issues, Dr. Amin H. Jahromi prepared ligand **2** by attaching a polyamine chain to the acridine moiety.[43] In DM1 model cells, treatment of **2** at 75 μM for 48 h resulted in a ca. 86% reduction of cells containing MBNL1-r(CUG)$_{960}$ foci. It was also found that ligand **2** is capable of rescuing 40% of the *IR* splicing defect. In addition, compound **3**, discovered from a small library of bivalent ligands containing different polyamine or polyether linkers, shows significantly improved binding affinity to rCUG repeats, and inhibition

of MBNL1-rCUG^{exp} complex compared to ligand **1**.[44]

In addition to the triaminotriazine-acridine conjugate, Dr. Chun-Ho Wong rationally designed ligand **4** as a rCUG major groove binder.[45] Compound **4** contains two triaminotriazine recognition units, each linked to a bisamidinium benzene by a four methylene linker. Ligand **4** has a number of features that make it a promising lead candidate, including a low micromolar binding affinity (K_D = 8 ± 2 μM) to r(CUG)$_{12}$ and a modest inhibition of the MBNL1-r(CUG)$_{12}$ complex (K_I = 8 ± 2 μM). In addition, ligand **4** is water soluble, and non-toxic to HeLa cells up to 100 μM after 3-day incubation. In DM1 model cells, the nuclear foci formed by rCUG and MBNL1 was significantly reduced upon treatment of **4** at 100 μM for 48 h. Furthermore, the mis-splicing of *cTNT* and *IR* pre-mRNAs were partially restored. Ligand **4** was also found to reverse the rCUG-induced rough and glossy eye phenotypes in a DM1 *Drosophila* model.

Using fluorescence anisotropy, Eric (Yen Jun) Ho discovered compound **5** from the Diversity Set III of the National Cancer Institute compound library as a low micromolar inhibitor of the MBNL1-rCUG^{exp}. However, preliminary results showed that ligand **5** exhibits weak interactions with rCUG and MBNL1, and a low micromolar binding affinity to a CTG repeat.

CHAPTER 2

DEVELOPMENT OF THE BISAMIDINIUM ANALOGS[1]

2.1 Introduction to the Bisamidinium Ligand

Bisamidinium-based ligand **4** has many promising therapeutic properties and was considered as a key lead compound in this work. Ligand **4** exhibits modest binding affinity toward rCUG^{exp} (K_D = 8 ± 2 μM) and inhibition of MBNL1-rCUG^{exp} (K_I = 8 ± 2 μM).[45] With high water solubility, and low cytotoxicity in HeLa cells, the efficacy of **4** in DM1 model cells was studied, showing that **4** was able to dissolve nuclear foci and partially correct splicing defects. Therefore, for further development of **4**, its binding mode to rCUG repeats should be elucidated. Herein, I provide some preliminary binding data, and uncover a point of modification on **4** where substituents can be placed without affecting the binding affinity. This work provides a foundation for future developments of more efficacious melamine-bisamidinium ligands.

2.2 Binding of the Bisamidinium Ligand to rCUG Repeats

The rational design of ligand **4** was based on the reported NMR structure of the complex between a frame-shifting bisamidinium stimulator ligand DB213 and its HIV-1 frameshift site (FS) RNA stem-loop target (Figure 5a).[46] It was suggested that amidinium groups form hydrogen

[1] Some of the material in this chapter was adapted from the following publications:
Luu, L. M.; Lee, J.; Nguyen, L.; Vergara, J. I.; Zimmerman, S. C. A Fragment-Based, Multi-target Drug Discovery Approach for Both DNA and RNA Targets Using a Template-Assisted Selection. (Manuscript in preparation)
Luu, L. M.; Nguyen, L.; Peng, S.; Lee, J.; Lee, H. Y.; Wong, C.-H.; Hergenrother, P. J.; Chan, H. Y. E.; Zimmerman, S. C. *ChemMedChem* **2016**, *11*, 1428–1435.
Nguyen, L.; **Luu, L. M.**; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2015**, *137*, 14180–14189.
Wong, C.-H.; Nguyen, L.; Peh, J.; **Luu, L. M.**; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 6355–6361.

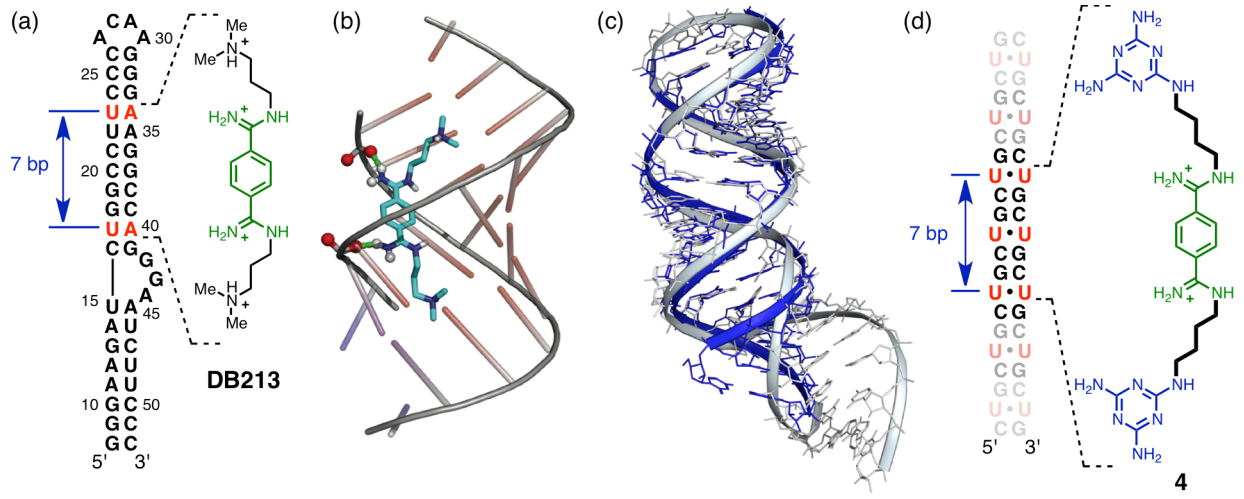


Figure 5. Rational design of ligand **4**. a) Schematic diagram and b) Illustration of the DB213-HIV-1 FS RNA complex. c) An overlay of the r(CUG)$_6$ x-ray structure (blue, PDB: 3GM7) and the NMR-determined HIV-1 FS RNA structure (grey, PDB: 2L94). (d) Proposed binding of compound **4** to rCUG repeats. (The image was adapted from Ref 45 and 46 with minor changes)

bonds with the phosphate oxygens on opposite strands of the HIV-1 FS RNA stemloop (Figure 5b).[46] An overlay of the r(CUG)$_{12}$ x-ray structure and the NMR-determined HIV-1 FS RNA structure showed striking similarities (Figure 5c), suggesting that the bisamidinium moiety could potentially serve as a rCUG major groove binder. Thus, the bisamidinium unit of **4** was designed to cover the central CUG with two triaminotriazine rings to recognize the two adjacent U-U mismatches (Figure 5d).[45] We anticipated that the 1,3 U-U mismatches are essential for the recognition of ligand **4**, whereas the internal U-U mismatch would have no impact on the binding affinity of **4** toward A and B.

Isothermal titration calorimetry (ITC) was employed to study the binding of **4** to different rCUG oligonucleotide sequences. Thermodynamic parameters and binding stoichiometry can be derived from ITC experiments, allowing the users to gain insight into the binding mode of the designed ligands.[47] Although a large quantity of material is needed for a single experiment, this

method does not require labeling of either the small molecule or RNA.

RNA constructs A and B, containing three and two U-U mismatches, respectively, were employed to investigate the binding mode of **4** with rCUG repeats (Figure 6). These sequences contain a UUCG tetraloop to promote the hairpin formation.[48] Before the ITC measurement was performed, the RNA sequences were refolded by heating the RNA solution at 95 °C for 5 min and cooling to room temperature for 2 h. After refolding, both sequences existed as a single conformation as indicated by native gel. The binding of **4** to rCUG repeats was studied at 25 °C at pH 7.0 (MOPS 20 mM, NaCl 150 mM).

It was found that no binding isotherm was detected for the interaction of **4** with construct B, whereas a low micromolar binding affinity towards construct A was observed (Figure 6). A two sequential binding model was employed to fit the binding isotherm of **4** with construct A. The strongest binding event with a K_D value of 1.4 μM is exothermic and entropy driven. The second interaction (K_D = 103 μM) is 74-fold weaker than the first binding event. Although, the interaction of **4** with construct B was also characterized by two sequential binding events with the K_D values of 36 and 372 μM, a negligible amount of heat was generated from the titration, signifying a weak interaction of **4** to construct B. Because the equivalence point was not clearly observed from the isotherm, the binding stoichiometry was not determined. Although increasing concentrations of both ligand and RNA could resolve this problem, the high cost of RNA makes it economically impractical.

This ITC study demonstrates the importance of the middle U-U mismatch in the interaction of **4** to CUG repeats. However, one may argue that two consecutive U-U mismatches provide a sufficient binding site for ligand **4**, given by the weak binding of **4** to construct B,

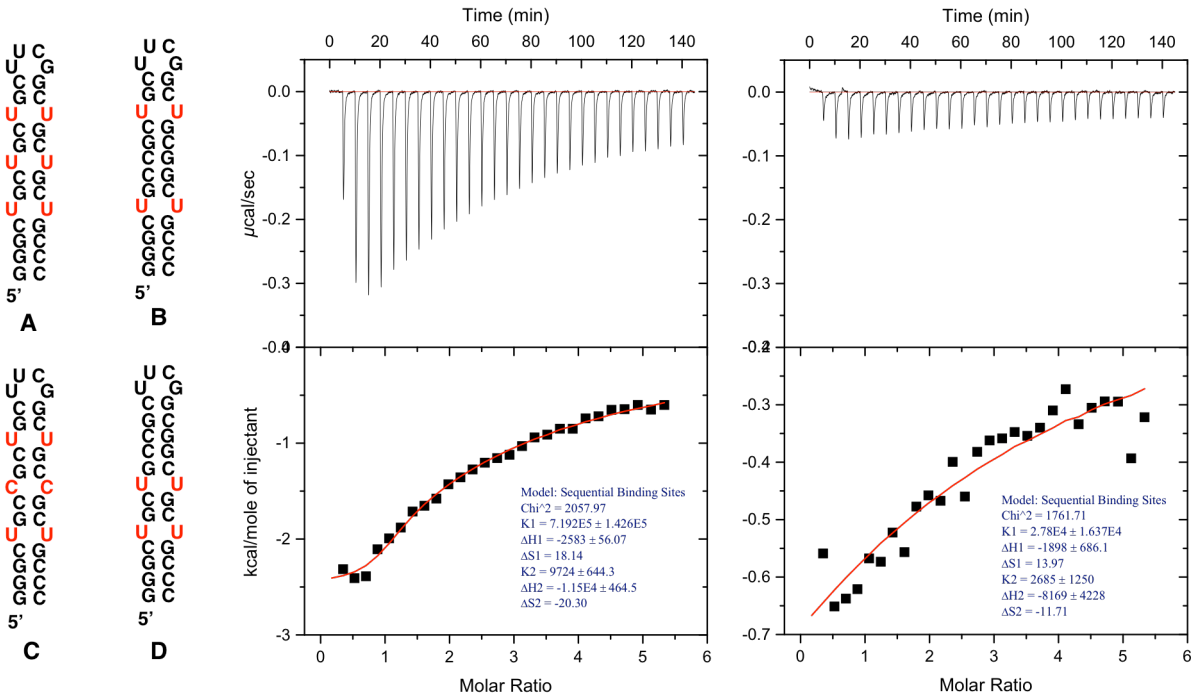


Figure 6. ITC binding curves represented for the interactions of ligand **4** with construct A (left) and construct B (right).

where the middle U-U mismatch is replaced by a G-C base. On the other hand, the internal mismatch may be more accessible for the bisamidinium moiety than the fully paired bases. Therefore, binding of **4** toward construct C and D would shed light on this question (Figure 6).

2.3 Synthesis and Binding Study of Bisamidinium Analogs

The two most important considerations in designing the bisamidinium analogs include their synthetic accessibility and their ability to span adjacent binding sites without disrupting the primary interactions with rCUG repeats. The substituted analogues of dicyanobenzene and 1,4-diaminobutane are quite limited in availability, while unsymmetrically substituted triaminotrazines are readily synthetically accessible. Thus, the logical functionalization point was the triaminotriazine amino group of **4**. The key question was whether *N*-substitution of one or both of the triaminotriazine units of **4** would affect its ability to bind rCUG^{exp}. Previously, we

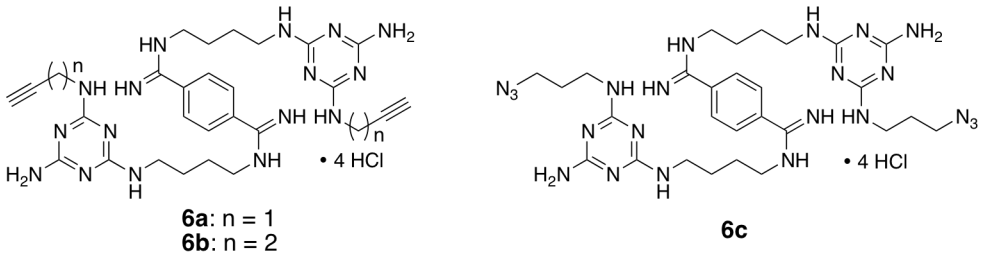
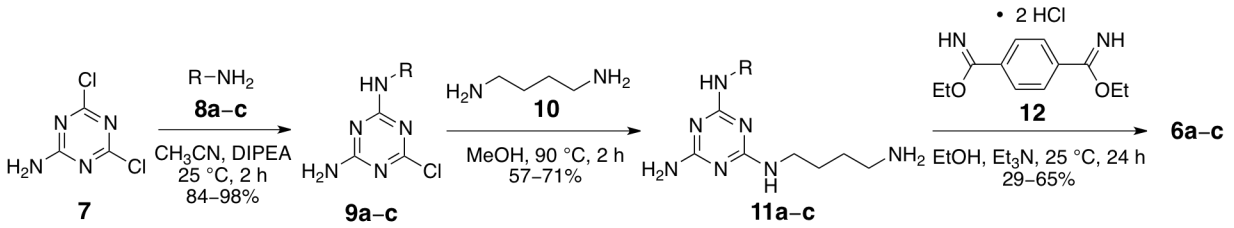


Figure 7. Structures of compounds **6a**−**b**.

showed that methylation of amino groups in the triaminotriazine-acridine ligand **1** may have a profound effect on both affinity and selectivity.[40] To address this issue in the bisamidinium ligands, compound **6a**−**c** were prepared (Figure 7 and Scheme 1), and their binding affinity to rCUG repeats was determined.

The preparation of **6a**–**c** containing two alkyne or azide groups was based on the originally proposed synthetic route of ligand **4** (Scheme 1).[45] The triazine **7** was treated with propargylamine **8a**, 3-azido-propylamine **8b**, or 3-azido-propylamine **8c** at room temperature, affording **9a**–**c** in good yields. The reaction of **9a**–**c** with an excessive amount of 1,4-diaminobutane was refluxed at 90 °C to give compounds **11a**–**c**, which was reacted with bis-imidate ester **12** in anhydrous ethanol to afford ligands **6a**–**c** in moderate to high yields. Compounds **6a**–**c** were purified by either silica gel or C18 column chromatography. The purification yielded tetrahydrochloride salts of **6a**–**c** with a purity of at least 95% as indicated by

Scheme 1. Synthesis of **6a**−**b**.



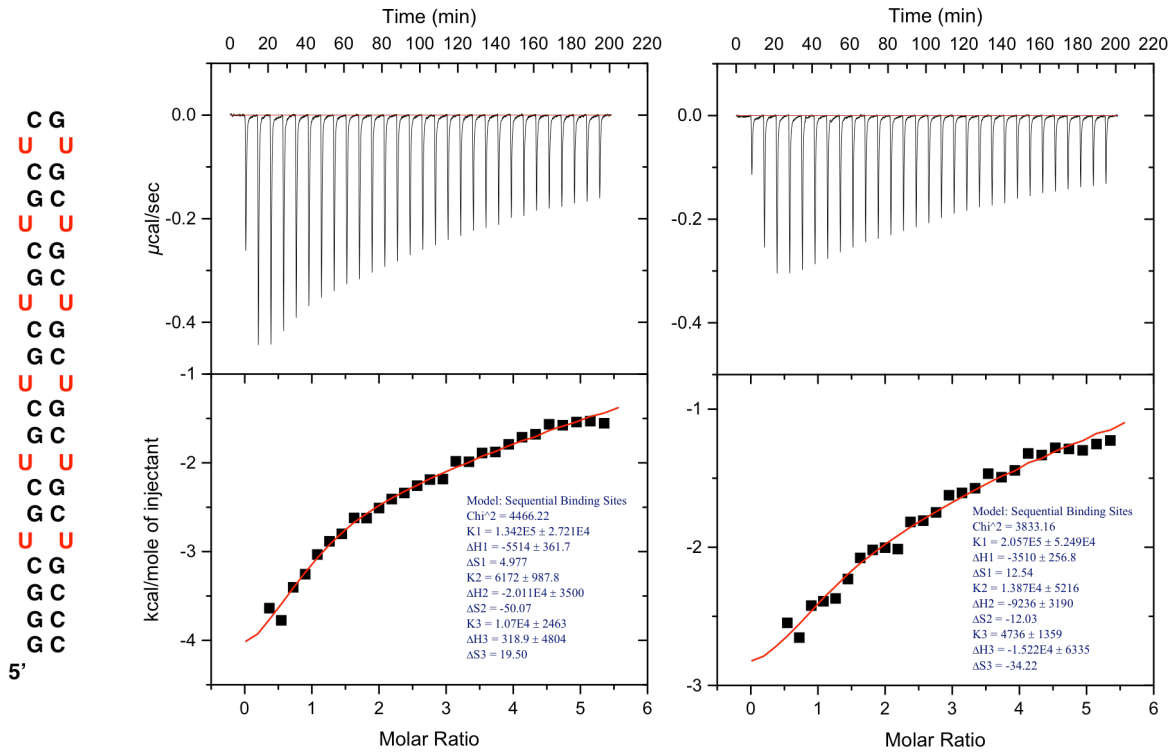


Figure 8. ITC binding isotherm to r(CUG)$_{12}$ of **6a** (left) and **6c** (right). The binding isotherms were fitted with the 3 sequential binding sites model. The heat generated at the end of the titration may result from the dilution of the RNA storage buffer in the RNA solution as the ligand solution was added.

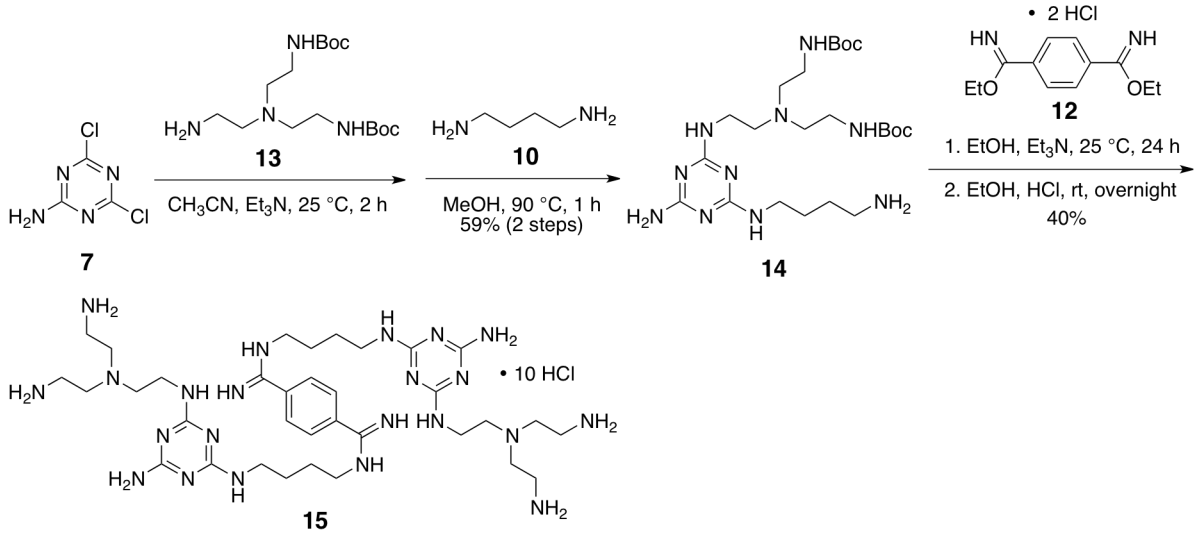
NMR and HPLC.

It was discovered from a single ITC experiment that the binding affinity toward r(CUG)$_{12}$ of compound **6a** (K_D = 7.7 μM) and **6c** (K_D = 4.8 μM) were similar to that of **4** (K_D = 8 ± 2 μM) (Figure 8).[45] They also exhibit similar binding isotherms. This finding supports the proposed binding mode of the triaminotriazine to the U-U mismatch where one of the remaining NH_2 groups can be mono-alkylated.[40] This finding enabled the rapid development of a wide range of functional analogs of ligand **4** using the [3+2] copper-assisted alkyne-azide cycloaddition of **6a** and **6c**.

This binding study also suggests that substitutions on the triaminotriazine rings would not

affect the binding affinity of the bisamidinium ligands. In fact, compound **15**, prepared as described in Scheme 2, showed comparable binding affinity to $r(CUG)_{12}$.[49] Remarkably, Dr. Lien Nguyen found that **15** is able to 1) bind dCTG repeats and inhibit its transcription, 2) bind rCUG repeats and inhibit MBNL1-$rCUG^{exp}$ complex, and 3) degrade rCUG repeats.[49] In chapter 3, development and evaluation of the potent bisamidinium bivalent ligands will be described.

Scheme 2. Synthesis of ligand **15**.



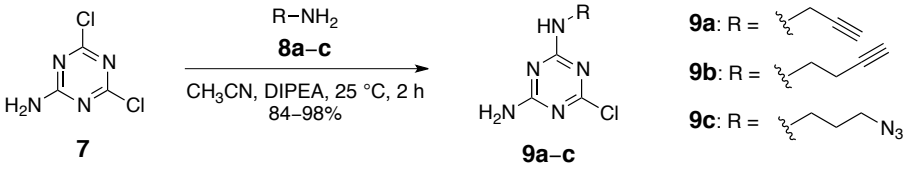
2.4 Conclusions and Outlook

In conclusion, the binding study of **4** by ITC suggests that three consecutive U-U mismatches provide a sufficient binding site for the bisamidinium-based ligand **4**. Although, more evidence could be obtained to further elucidate the binding mode of **4** toward rCUG repeats and advance the development of bisamidinium ligands, thus far we found that the triaminotriazine ring can be functionalized with different substituents with minimal structural and binding perturbation. This finding has allowed us to develop the bisamidinium ligands for different purposes, including multi-target agents and multivalent ligands.

2.5 Materials and Methods

2.5.1 Synthesis

All reactions were carried out under a dry N_2 atmosphere with oven-dried (115 °C) glassware. All solvents and reagents were of reagent quality, purchased commercially, and used without further purification. Reactions were monitored by thin layer chromatography using EMD pre-coated silica gel 60 F_{254} plates. Flash chromatography was performed using Silicycle SiliaFlash P60 (230–400 mesh) silica gel. ^{1}H and ^{13}C NMR spectra were recorded on a 500 MHz Varian Unity Inova spectrometer. Chemical shifts and coupling constants (*J*) were reported in ppm and Hertz, respectively. Electrospray ionization mass spectra (LR ESI-MS) were obtained by the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois at Urbana-Champaign. High performance liquid chromatography (HPLC) was performed using a Dynamax SD-200 system with a UV detector set at 260 nm and a DENALI C18 Vydac column (250 × 10 mm) with a dual solvent system of 0.1% TFA in H_2O (Solvent A) and 0.1% TFA in MeOH (Solvent B).



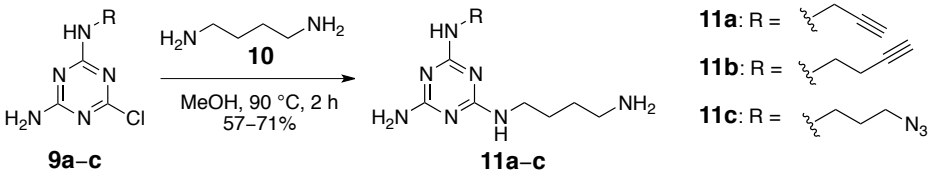
Compound **9a**: To a suspension of 1.5 g (9.1 mmol) of **7**, and 1.6 mL (9.2 mmol) of DIPEA in 20 mL of CH_3CN was added dropwise 0.6 mL (9.3 mmol) of propargylamine **8a** over 30 min at room temperature. The mixture was stirred at room temperature for 2 h. The solid was filtered, washed with water and dried under nitrogen overnight to give 1.4 g (85%) of **9a** as a white solid (R_f = 0.43, CH_2Cl_2/MeOH = 9:1). ^{1}H NMR (DMSO-d_6): δ 8.11−7.94 (m, 1H),

7.44−7.23 (m, 2H), 4.01−3.98 (m, 2H), 3.07 (t, J = 2.5, 1H). ^{13}C NMR (DMSO-d_6): δ 168.3, 166.6, 165.4, 80.7, 72.2, 29.3. LR-ESI-MS (m/z) calculated for $[M + H]^+$: 184.0; found 184.0.

Compound **9b**: Using the similar procedure as described above, 600 mg (3.6 mmol) of **7**, 0.95 mL (5.5 mmol) of Et_3N, and 0.35 mL (4.3 mmol) of 1-amino-3-butyne **8b** afforded 660 mg (98%) of **9b** as a white solid. ^{1}H NMR (DMSO-d_6): δ 7.84−7.66 (m, 1H), 7.35−7.14 (m, 2H), 3.36−3.28 (m, 2H), 2.84−2.82 (m, 1H), 2.39−2.33 (m, 2H).

Compound **9c**: Using the similar procedure as described above, 8.4 g (50.9 mmol) of **7**, 7.5 mL (53.8 mmol) of Et_3N, and 6.0 g (59.9 mmol) of 3-azido-propylamine **8c** afforded 9.8 g (84%) of **9c** as a white solid. ^{1}H NMR (DMSO-d_6): δ 7.83−7.65 (m, 1H), 7.33−7.10 (m, 2H), 3.40−3.36 (m, 2H), 3.28−3.23 (m, 2H), 1.75−1.68 (m, 2H).

Note: compound **7**, propargylamine **8a**, 3-azido-propylamine **8b**, 3-azido-propylamine **8c** are available from Sigma-Aldrich or can be easily prepared using reported procedures.

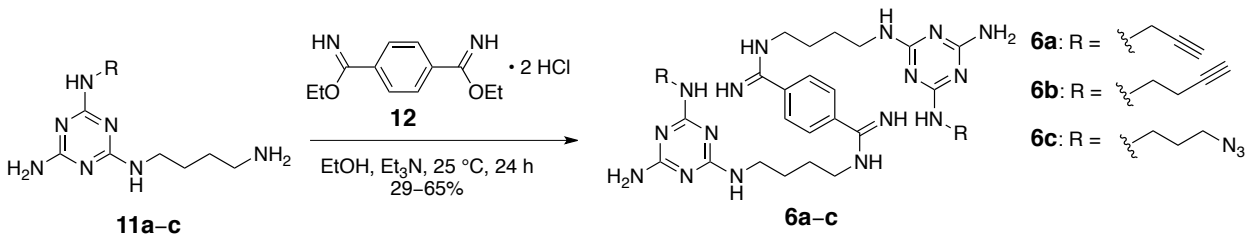


Compound **11a**: To a mixture of 22 mL (218.9 mmol) of 1,4-diaminobutane **10** in 50 mL of MeOH was added slowly 13.0 g (71.1 mmol) of **9a** over 30 min at 90 °C. The mixture was stirred at 90 °C for 1 h. The solvent and the excess amount of 1,4-diaminobutane were removed by distillation. The crude was purified by silica gel column chromatography with a gradient mixture of CH_2Cl_2/MeOH from 9/1 to 4/1 (v/v). Fractions containing product were combined and concentrated using a rotary evaporator to give 11.9 g (71%) of **11a** as a white solid. ^{1}H NMR

(DMSO-d_6): δ 6.83−6.52 (m, 2H), 6.26−5.98 (m, 2H), 4.68 (br s, 2H), 3.97−3.95 (m, 2H), 3.21−3.15 (m, 2H), 3.00 (t, J = 2.4, 1H), 2.62 (t, J = 6.9, 2H), 1.47−1.39 (m, 4H). ^{13}C NMR: δ 166.97, 166.73, 165.97, 165.77, 165.57, 82.72, 72.16, 41.31, 30.41, 29.22, 26.89. LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 236.2; found 236.1.

Compound **11b**: Using the similar procedure as described above, 0.7 g (3.5 mmol) of **9b** and 1.5 mL (14.9 mmol) of 1,4-diaminobutane **10** afforded 0.5 g (57%) of **11b** as a white solid. ^{1}H NMR (DMSO-d_6): δ 6.59−6.45 (m, 2H), 6.22−5.75 (m, 2H), 4.12 (br s, 2H), 3.34−3.29 (m, 4H), 2.81 (t, J = 2.7, 1H), 2.66 (t, J = 6.8, 2H), 2.33 (br s, 2H), 1.45 (br s, 4H).

Compound **11c**: Using the similar procedure as described above, 9.3 g (40.7 mmol) of **9c** and 13 mL (129.3 mmol) of 1,4-diaminobutane **10** afforded 7.5 g (66%) of **11c** as a white solid. ^{1}H NMR (DMSO-d_6): δ 8.07 (br s, 2H), 6.67−6.46 (m, 2H), 6.17−5.89 (m, 2H), 3.36 (t, J = 6.8, 2H), 3.23−3.15 (m, 4H), 2.74 (t, J = 7.5, 2H), 1.72−1.69 (m, 2H), 1.59−1.46 (m, 4H). LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 281.2; found 281.1.

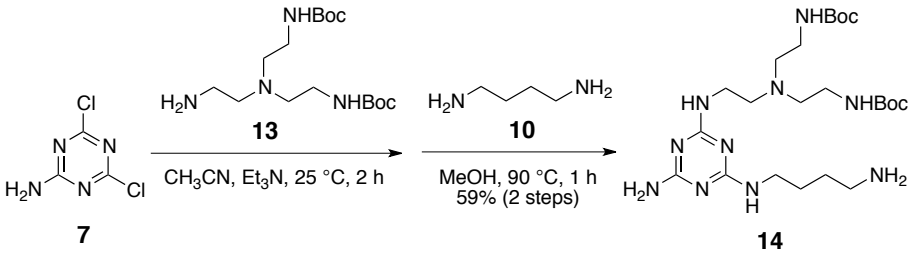


Compound **6a**: To a white suspension of 1.5 g (5.1 mmol) of diethyl terephthalimidate hydrochloride **12** in 30 mL of anhydrous EtOH was added 3.3 mL (23.7 mmol) of anhydrous Et_3N. To the resulting clear solution was added 13 mL (13.0 mmol) of ethanolic solution of **11a** 1 M. The reaction mixture was stirred at room temperature for 24 h. The solvent was removed using a rotary evaporator. The white solid was purified by silica gel column chromatography

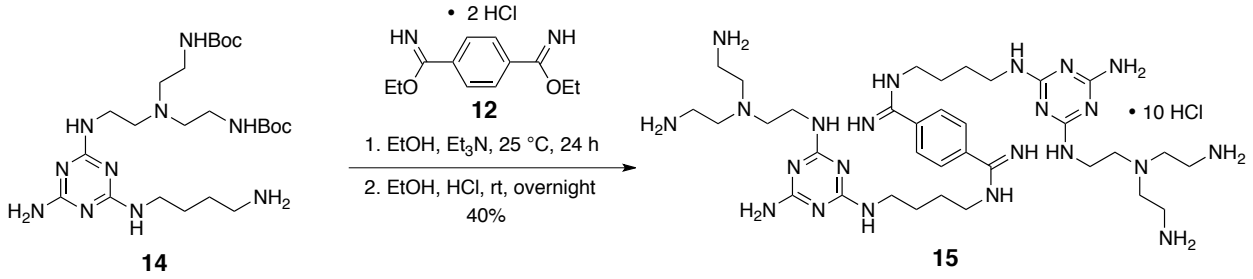
with a gradient mixture of CH_2Cl_2/MeOH from 9/1 to 1/9 (v/v). After using 1 L of CH_2Cl_2/MeOH (1:9, v/v), the eluent was acidified with a gradient of 0.1–0.2 mL of dioxane solution of HCl 4 M per liter of eluent. The product-containing fractions were combined, filtered, and concentrated using a rotary evaporator to give 1.1 g (29%) of **6a** as a white tetra-HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.29 (s, 2H), 9.85 (s, 2H), 9.50−9.47 (m, 2H), 8.70−8.65 (m, 4H), 8.28−7.99 (m, 8H), 4.10−4.07 (m, 4H), 3.51−3.49 (m, 4H), 3.40−3.33 (m, 4H), 3.24−3.21 (m, 4H), 1.71−1.62 (m, 8H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 599.3544; found 599.3536. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 5.1 min (98%).

Compound **6b**: Using the similar procedure as described above, 181 mg (0.61 mmol) of **12**, 0.25 mL (1.8 mmol) of anhydrous Et_3N, and 7 mL (1.4 mmol) of ethanolic solution of **11b** 0.2 M afforded 250 mg (52%) of **6b** as a white tetra-HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.22 (s, 2H), 9.78 (s, 2H), 9.40 (s, 2H), 8.21−7.73 (m, 14H), 3.48 (q, J = 6.6, 4H), 3.40−3.40 (m, 16H), 1.71−1.59 (m, 8H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 627.3857; found 627.3840. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 6.1 min (100%).

Compound **6c**: Using the similar procedure as described above, 261 mg (0.9 mmol) of **12**, 0.35 mL (2.5 mmol) of anhydrous Et_3N, and 545 mg (1.9 mmol) of **11c** afforded 484 mg (65%) of **6c** as a white tetra-HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.32−10.27 (m, 2H), 9.83 (s, 2H), 9.49−9.45 (m, 2H), 8.54−8.38 (m, 4H), 8.07−7.92 (m, 8H), 3.51−3.47 (m, 4H), 3.43−3.32 (m, 12H), 1.80−1.62 (m, 12H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 689.4198; found 689.4182. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 8.7 min (100%).



Compound **14**: To a 50 mL round-bottomed flask containing 690 mg (4.2 mmol) of compound **7** was added 10 mL of CH_3CN and 0.7 mL (5.0 mmol) of Et_3N. The suspension was stirred at room temperature for 5 min. To the reaction flask was added 1.7 g (4.9 mmol) of **13** at once, giving a clear solution. The reaction was stirred at room temperature for 2 h. Acetonitrile was removed using a rotary evaporator. The solid was dissolved in 30 mL of MeOH. To the resulting solution was added 2.0 mL (19.9 mmol) of 1,4-diaminobutane **10** at once. The reaction was heated to 90 °C, and stirred for 1 h. Methanol and the excess amount of diaminobutane were removed by a rotary evaporator. The crude was purified by silica gel column chromatography with a mixture of CH_2Cl_2/MeOH/NH_4OH (90/9/1, v/v/v). Fractions containing product were combined and concentrated to give 1.3 g (59%) of compound **14** as a white solid. ^{1}H NMR (DMSO-d_6): δ 6.75 (br s, 2H), 6.53–5.81 (m, 4H), 3.19–3.14 (m, 4H), 2.97 (br s, 4H), 2.56–2.53 (t, J = 6.9, 2H), 2.46–2.44 (t, J = 6.5, 6H,), 1.49–1.30 (m, 4H), 1.37 (br s, 18H). LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 527.4; found 527.2.



Compound **15**: To a 100 mL oven-dried round-bottomed flask was added 250 mg (0.9 mmol) of **12**. The white solid was dissolved in 15 mL of anhydrous EtOH. To the resulting suspension was added 0.3 mL (2.2 mmol) of anhydrous Et_3N followed by 1.0 g (1.9 mmol) of **14** at once. The resulting suspension was stirred at room temperature for 1 d. The solvent was removed using a rotary evaporator. The crude was dissolved in 20 mL of ethanolic solution of HCl 2 M. The reaction was stirred at room temperature overnight. Ethanol was removed using a rotary evaporator. The crude was purified by a Sephadex CM-25 column chromatography using aqueous solution of NH_4HCO_3 from 0.1 M to 1.0 M. Fractions containing products were combined and concentrated at 60 °C using a rotary evaporator. The solid was dissolved in 80 mL of aqueous solution of HCl 0.1 M. The resulting solution was concentrated using a rotary evaporator to give 390 mg (40%) of compound **15** as a white HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.10 (s, 2H), 9.70 (s, 2H), 9.26 (s, 2H), 8.42–7.79 (m, 26H), 3.45–3.32 (m, 12H), 2.90 (s, 8H), 2.70–2.61 (m, 12H), 1.69–1.60 (m, 8H). LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 781.6; found 781.6.

2.5.2 Isothermal Titration Calorimetry

RNA sequences were purchased from Integrated DNA Technology (Coralville, IA). The RNA solid was dissolved in the RNA storage buffer containing 1 mM sodium citrate pH 6.4 (Ambion® #AM7001).

The ligand stock solutions (10 mM in water), MOPS buffer (100 mM; pH 7.0 ± 0.2), and NaCl (5000 mM) were prepared. The ligand solution (1 mL) contained 0.5 mM of ligand, 20 mM of MOPS buffer, and 150 mM of NaCl. The RNA solution (2 mL) contained 0.02 mM of RNA, 20 mM of MOPS buffer, and 150 mM of NaCl. The resulting RNA solution was heated in

a water bath at 90 °C for 5 min, then allowed to cool slowly to room temperature for 2 h.

ITC measurements were performed at 25 °C on a MicroCal VP-ITC calorimeter (Northampton, MA) in Microanalysis Laboratory, University of Illinois at Urbana-Champaign. A typical experiment consisted of 30 injections of a ligand solution (500 μM) from a 298 μL syringe into an RNA solution (10 μM, 1.42 mL) stirred at 300 rpm (1 μL for the first injection and 10 μL for the remaining). The initial delay prior to the first injection was 60 s. The duration of each injection was 20.5 s and the delay between injections was 400 s. Data was analyzed using Origin 5.0 software (MicroCal). Binding parameters such as the dissociation constant (K_D), enthalpy change (ΔH), and entropy change (ΔS) were determined by fitting the experimental binding isotherms with an appropriate model.

CHAPTER 3

DEVELOPMENT OF THE BISAMIDINIUM DIMERIC LIGANDS[2]

3.1 Introduction

As described in Chapter 1, the causative agent of DM1 is the rCUG repeats whose length varies from 50 to thousands repeats. Although ligands **1**–**4** exhibit promising inhibitory activities, they likely span only 1-3 adjacent U-U sites and certainly no more than three U-U mismatches. We anticipate that a multivalent or modularly assembled ligand containing two or more RNA-binding molecules would exhibit improved inhibition of the MBNL1-rCUG^{exp} complex. In 2013, Zimmerman and coworkers evaluated a library of dimeric ligands, which are comprised of two melamine-acridine units connected by different oligoether and oligoamino linkers.[44] Compared with the monomeric ligand, the most potent compound **3**, exhibited a 206-fold greater binding affinity to r(CUG)$_6$ and a 266-fold greater inhibition of the MBNL1-r(CUG)$_{12}$ complex. Advantageously, compound **3** is also water soluble, able to enter the nuclei and fully disrupt nuclear foci at 50 μM.

In this chapter, the synthesis and biological activities of the bisamidinium-based bivalent ligands are described. Dimer **20a** was found to be a potent MBNL1-rCUG^{exp} complex inhibitor with a low nanomolar K_I. However, its efficacy in DM1 model cells was not what might be expected for such a strong inhibitor, which may be attributed to its low cellular permeability resulting from the high molecular weight of the dimer (M_r = 1166.4 Da). In addition, the

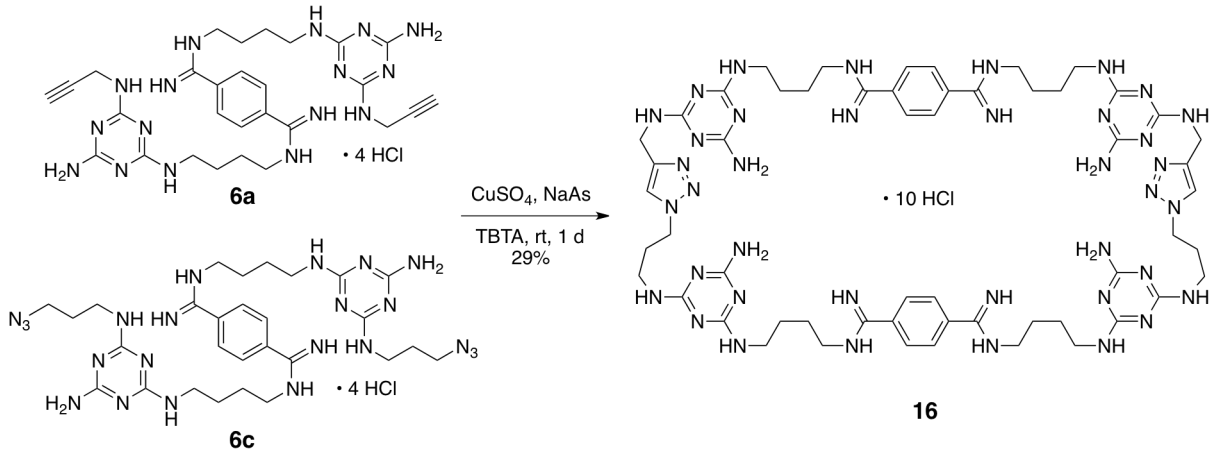
[2] Some of the material in this chapter was adapted from the following publication:
Luu, L. M.; Nguyen, L.; Peng, S.; Lee, J.; Lee, H. Y.; Wong, C.-H.; Hergenrother, P. J.; Chan, H. Y. E.; Zimmerman, S. C. *ChemMedChem* **2016**, *11*, 1428–1435.

compound has a significantly lower maximum tolerated dose (MTD) in mice than that of the monomeric ligand. These important observations suggest a need to access the potency of the dimer while still retaining some of the favorable attributes of the monomer.

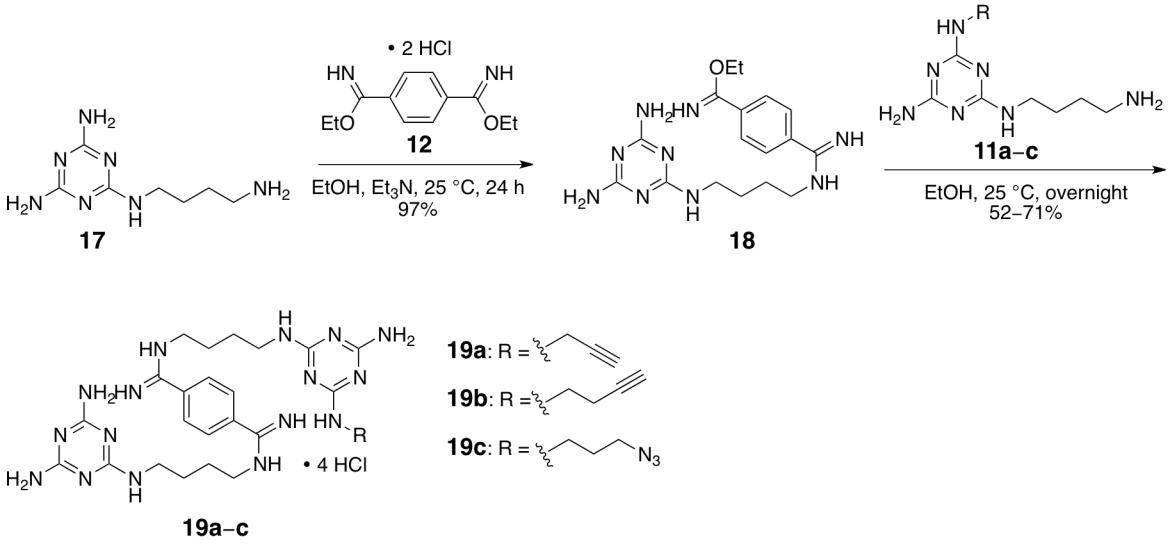
3.2 Design and Synthesis of the Dimeric Ligands

It was previously shown that compounds **6a**−**c** containing alkyne and azide groups exhibit comparable binding affinities toward rCUG repeats with compound **4**. Thus, the linker that connects the two modules of the bisamidinium dimer should be placed on the triaminotriazine unit. Initial attempt to prepare a bisamidinium dimer was to link ligands **6a** and **6c** using the copper-catalyzed click reaction (Scheme 3). The copper(I) was generated *in situ* from the reduction of copper(II) sulfate by sodium ascorbate.[50] Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA) was employed to prevent Cu^{I} from oxidizing to Cu^{II}.[51] Interestingly, the alkyne-azide cycloaddition of **6a** and **6b** mainly afforded cyclic dimer **16**, which was characterized by the loss of the acetylene proton and carbon in ^{1}H and ^{13}C NMR. The low yield was attributed to the loss of compound during the purification process.

Scheme 3. Synthesis of cyclic dimer **16**.



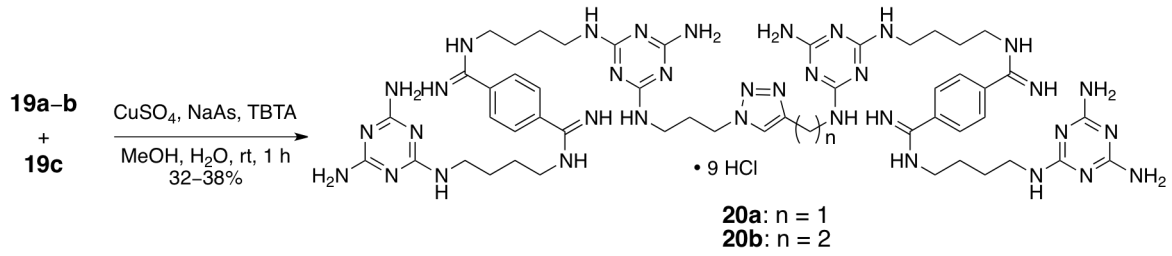
Scheme 4. Synthesis of ligands **20a**−**c**.



To prepare linear bisamidinium dimers, compounds containing one clickable moiety were synthesized (Scheme 4). The preparation of unsymmetrical ligands was more challenging, because the remaining imidate ester group in compound **18** is susceptible to hydrolysis and alcoholysis. Therefore, compound **17** was added slowly to the excess amount of imidate ester **12** to produce **18** with a small amount of **4** as indicated on TLC and ESI-MS. Reactions of **18** with the melamine **11a**−**c** afforded ligands **19a**−**c** in moderate yields.

Dimers **20a**−**b** were prepared from the click reaction of **19a**−**b** and **19c** (Scheme 5). Because the starting materials and the byproducts are composed of varying number of positive charges, dimers **20a**−**b** were purified by an ion exchange column chromatography (Sephadex

Scheme 5. Synthesis of dimers **20a**−**b**.



CM-25). Compounds **20a**−**b** in HCl salt form were obtained with a purity of greater than 95% pure as indicated by ^{1}H NMR spectroscopy, and analytical HPLC.

3.3 Inhibition of the MBNL1-r(CUG)$_{16}$ Complex by Dimeric Ligands

Inhibition of the MBNL1-r(CUG)$_{16}$ complex was evaluated by electrophoretic mobility shift assay (EMSA). Because of the high sensitivity of the radioisotope-labeled RNA, EMSA requires small amounts of both RNA and protein, making it widely used to study the binding of protein and RNA as well as the inhibition of protein-RNA interactions.[52] Because monomer **4** was designed to span three U-U mismatches, the r(CUG)$_{16}$ containing a maximum of eight U-U mismatches would provide at least one binding site for dimers **20a**−**b**.[45] In this study, MBNL1N, a truncated MBNL1 containing four zinc finger motifs of MBNL1 and a hexahistidine tag (C-terminus) was employed. The MBNL1N construct and the full-length MBNL1 have similar affinities to rCUG^{exp}.[53] The MBNL1-r(CUG)$_{16}$ complex and the free RNA were separated on a 6% polyacrylamide gel at 4 °C, pH 8.2−8.3, and visualized by the autoradiography of the ^{32}P-labeled RNA. The ability of ligands to displace MBNL1 from its complex with rCUG repeats was judged by the ratio of bound RNA to the total amount of RNA.

Firstly, a head-to-head comparison between monomers **4**, **6a**, **6c**, and dimers **16**, **20a**, **20b** was conducted (Figure 9). The concentrations of MBNL1 and r(CUG)$_{16}$ were selected to have ca. 46% of r(CUG)$_{16}$ complexed to MBNL1. Under the treatment of ligands, more free RNA was observed. As can be seen in Figure 9, ligands **4** and **6a** exhibited similar inhibition potency at 10 and 100 μM, whereas a small improvement was observed for ligand **6c**. It is noteworthy that all three monomers were inactive at 10 μM, whereas most of the MBNL1-r(CUG)$_{16}$ complex was disrupted by the dimers. When the concentration of the dimers was reduced to 1 μM, ca. 25−40%

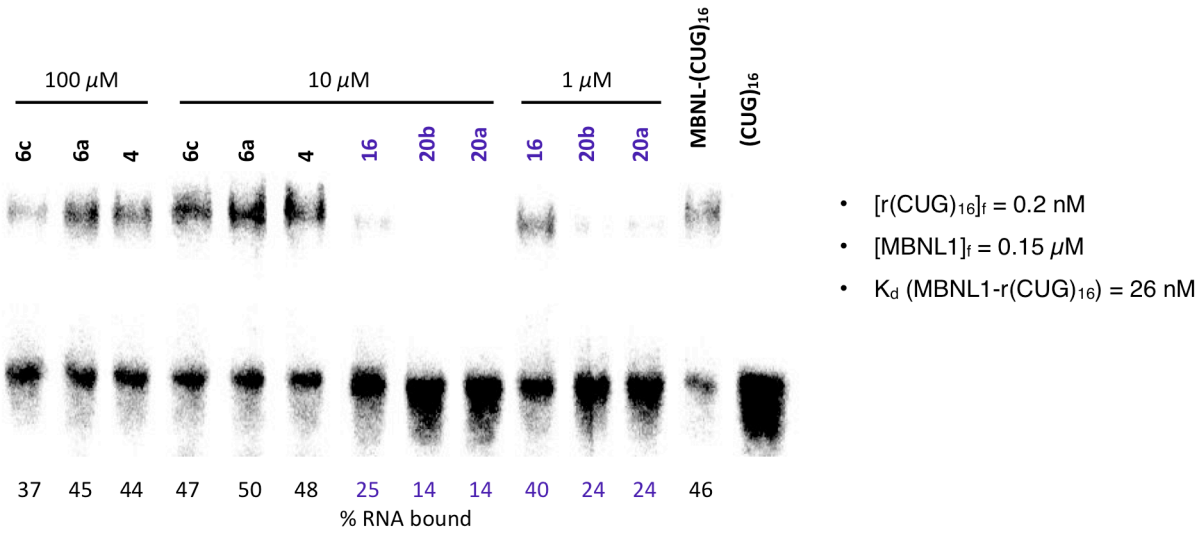


Figure 9. Head-to-head comparison of inhibition of the MBNL1-r(CUG)$_{16}$ complex by monomers **4**, **6a**, **6c** and dimers **16**, **20a**, **20b** using EMSA. (This experiment was conducted by Dr. Lien Nguyen)

of the MBNL1-r(CUG)$_{16}$ complex remained. This single experiment suggested that linear dimers exhibited highest inhibition potency.

The *in vitro* inhibition constants (K_I) of monomer **4** and dimers **20a**−**b** were quantitatively determined by EMSA (Figure 10a). In this study, an apparent binding constant (K_D) of MBNL1-r(CUG)$_{16}$ was 5.1 ± 1.6 nM, which is comparable to reported K_D values by EMSA and other techniques.[43,45] To determine the inhibition constant, the MBNL1-r(CUG)$_{16}$ complex, in which ca. 50% of the total RNA was free, was incubated with ligands at different concentrations. The concentration of dimers were not raised above 5 μM because of the their aggregation at higher concentrations. As the ligand concentration increased, more RNA was released from the complex with MBNL1. The IC$_{50}$ value was determined by fitting the plot of bound RNA fractions versus ligand concentrations with an appropriate equation (Figure 10b).

Compound **4**, dimers **20a**−**b** exhibited IC$_{50}$ values of 188 ± 32 μM, 290 ± 20 nM, and 430

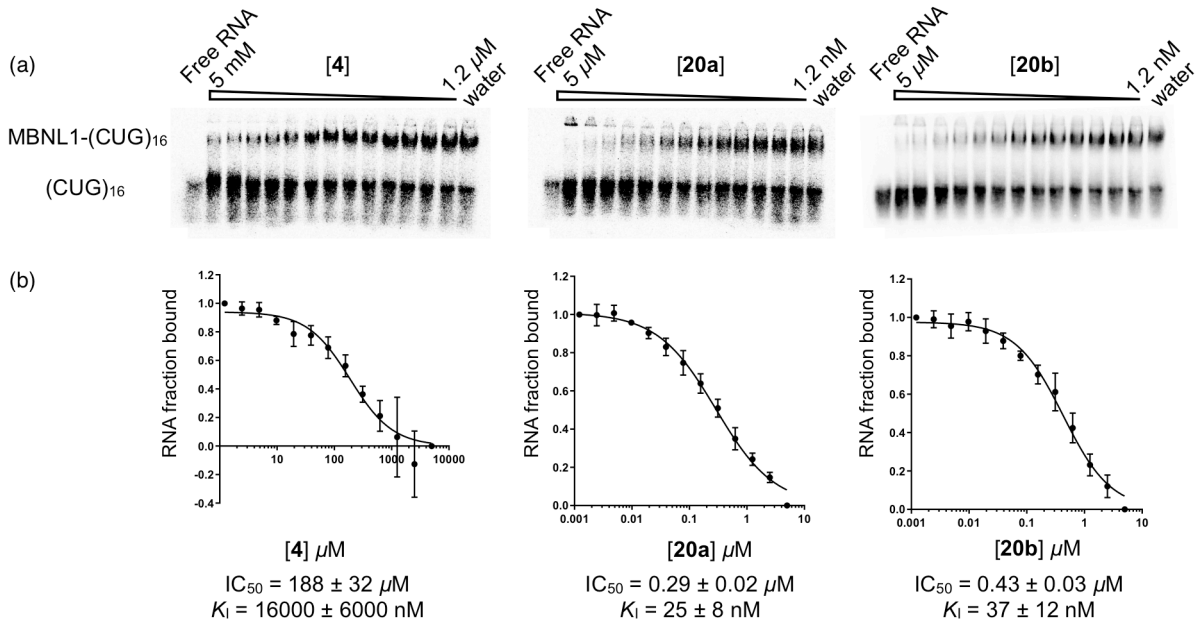


Figure 10. Determination of inhibition constants, K_I, by EMSA. (a) Representative gels show the dissociation of r(CUG)$_{16}$ from the complex with MBNL1 at different concentrations of **4** and **20a**−**b**. (b) The IC_{50} value was obtained by plotting the normalized fraction RNA bound versus the ligand concentration. Error bars represent the standard deviation from three independent experiments.

± 30 nM, and calculated K_I values of 16 ± 6 μM, 25 ± 8 nM, and 37 ± 12 μM, respectively. The significantly improved inhibitory potency (ca. 1000-fold) is likely attributed to bivalent binding to r(CUG)$_{16}$ of dimers. A binding study by ITC showed that the binding of **20a** to r(CUG)$_{16}$ produced significantly more heat than does the monomers (Figure 11). Although, the binding isotherm indicates a positively cooperative system, additional experiments are needed to fully elucidate the binding mode and to quantitatively determine the affinity of the dimeric ligands. Beyond their greater RNA affinity, dimeric ligands cover a greater surface area of the rCUG^{exp} target than does monomer **4**.

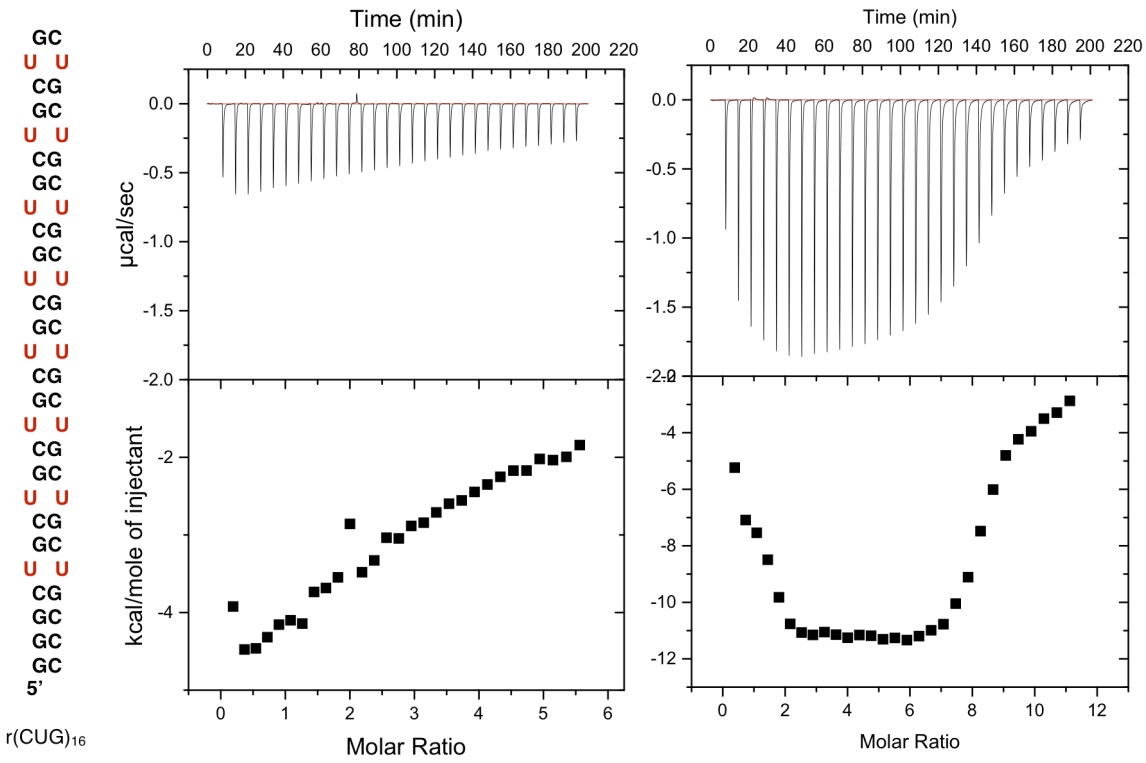


Figure 11. ITC binding isotherm obtained from the titration of **4** (left) and **20a** (right) to r(CUG)$_{16}$ at 25 °C, pH 7.0 (MOPS 20 mM, NaCl 150 mM).

3.4 Biological Activities of the Dimers in DM1 Model Cells

3.4.1 Foci Disruption

Because of the slightly lower K_I value and the lower cost of starting materials, dimer **20a** was selected for the cellular and animal assays. Before the efficacy of **20a** was evaluated in cells, its cytotoxicity to HeLa cells was studied in a sulforhodamine B (SRB) colorimetric assay. It was found that the dimer exhibited negligible cellular toxicity up to 100 μM after a 3-day treatment (Figure 12). At 100 μM, only ~10% HeLa cell death was observed, demonstrating the advantage of the bisamidinium-based dimer over the acridine-containing dimers.[44]

Disruption of the MBNL1-rCUG^{exp} foci in DM1 model cells was conducted by Dr. Lien

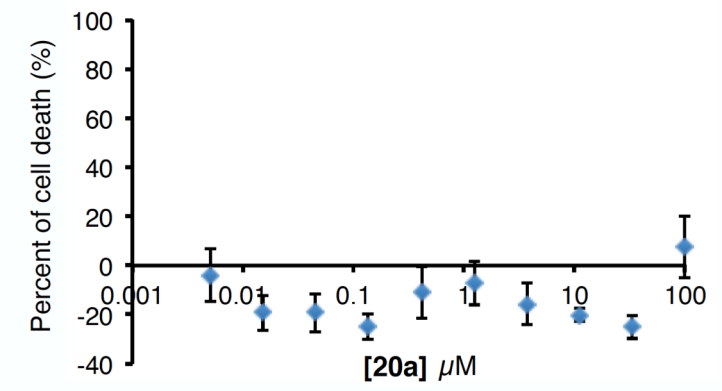


Figure 12. Cytotoxicity of **20a** to HeLa cells. (This experiment was conducted by Dr. Lien Nguyen)

Nguyen using a confocal microscopy experiment with immunofluorescence staining and fluorescent *in situ* hybridization (Figure 13a).[49] DM1 model cells were generated by transfecting HeLa cells with GFP-DT960 plasmid containing $(CTG)_{960}$ in exon 15 of a truncated *DMPK* gene, while cells with GFP-DT0 served as normal cells. The endogenous MBNL1 protein and rCUG transcript were detected by Alexa Fluor 647 dye-labeled antibody (green spots in column 1), and $(CAG)_{10}$ conjugated with a Cy3 dye (red spots in column 2). The blue areas are the nucleus stained by Hoechst 33342 (column 3).

MBNL1 was distributed throughout the nucleus in normal cells, whereas in DM1 model cells, it was concentrated in nuclear foci with rCUG repeats. A head-to-head comparison between **4** and **20a** was performed to evaluate the improvement of the dimeric ligand. Under the treatment of **20a** at 1 μM for 48 h, the number of foci decreased to ca. 72%. A similar effect was achieved with **4** at 100 μM over the same period of time (Figure 13b). Although it is encouraging that dimeric ligand **20a** is as active at 1 μM as monomeric ligand **4** is at 100 μM, it is surprising that full foci dispersion is not observed at higher concentrations. Incubation of **20a** at 10 μM decreased the foci area to ~36%, but a similar level was also observed for treatment at 100 μM. The lack of full dose dependence and its inability to fully disrupt nuclear foci suggest a cellular

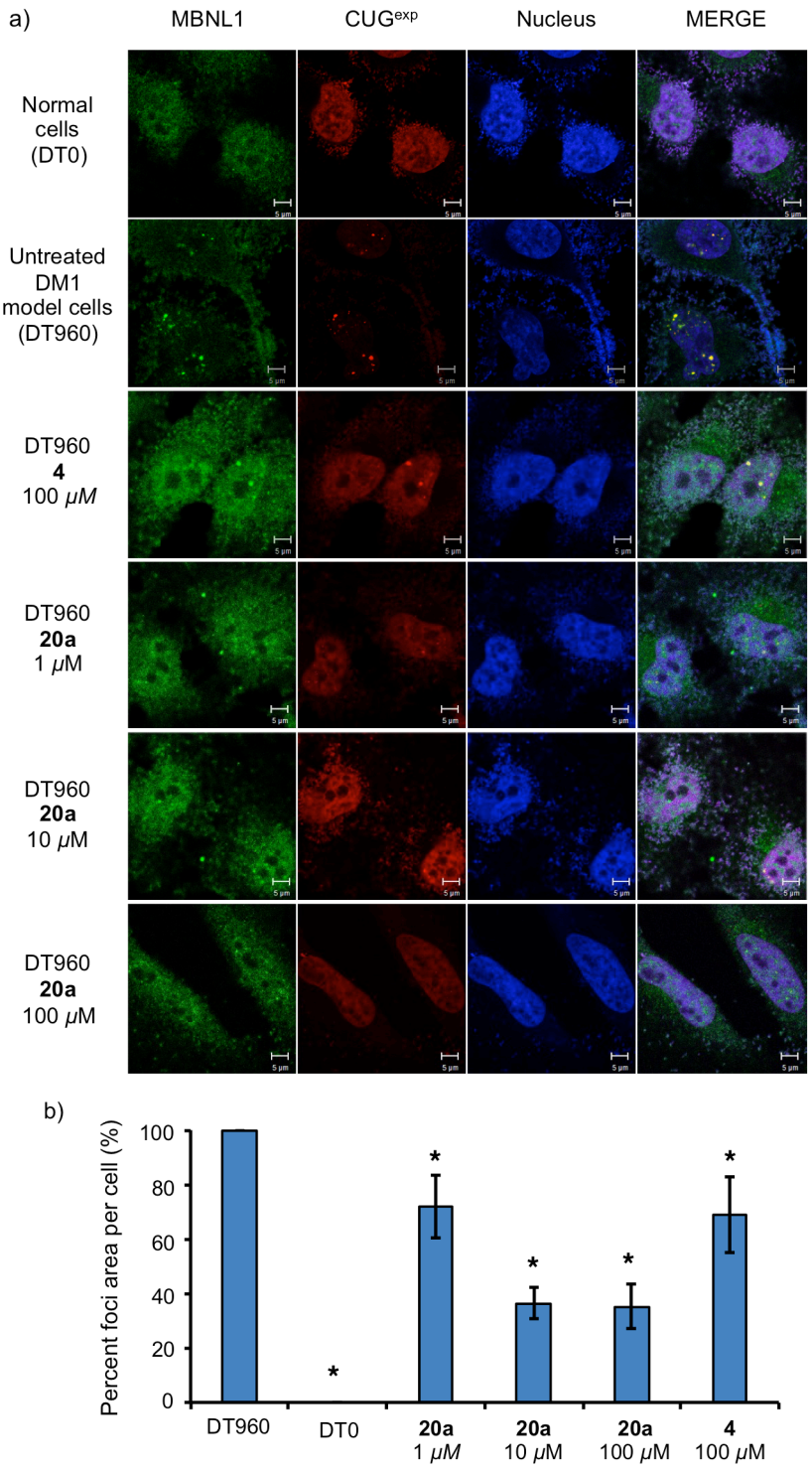


Figure 13. Foci disruption in DM1 model cells by ligands **4** and **20a**. (a) Representative confocal microscope images. (b) Percent of foci area per cell after treatment of ligands for 48 h. Error bars represent standard errors of mean from at least four independent experiments, two-tailed *t*-test * $P<0.005$. (This experiment was conducted by Dr. Lien Nguyen)

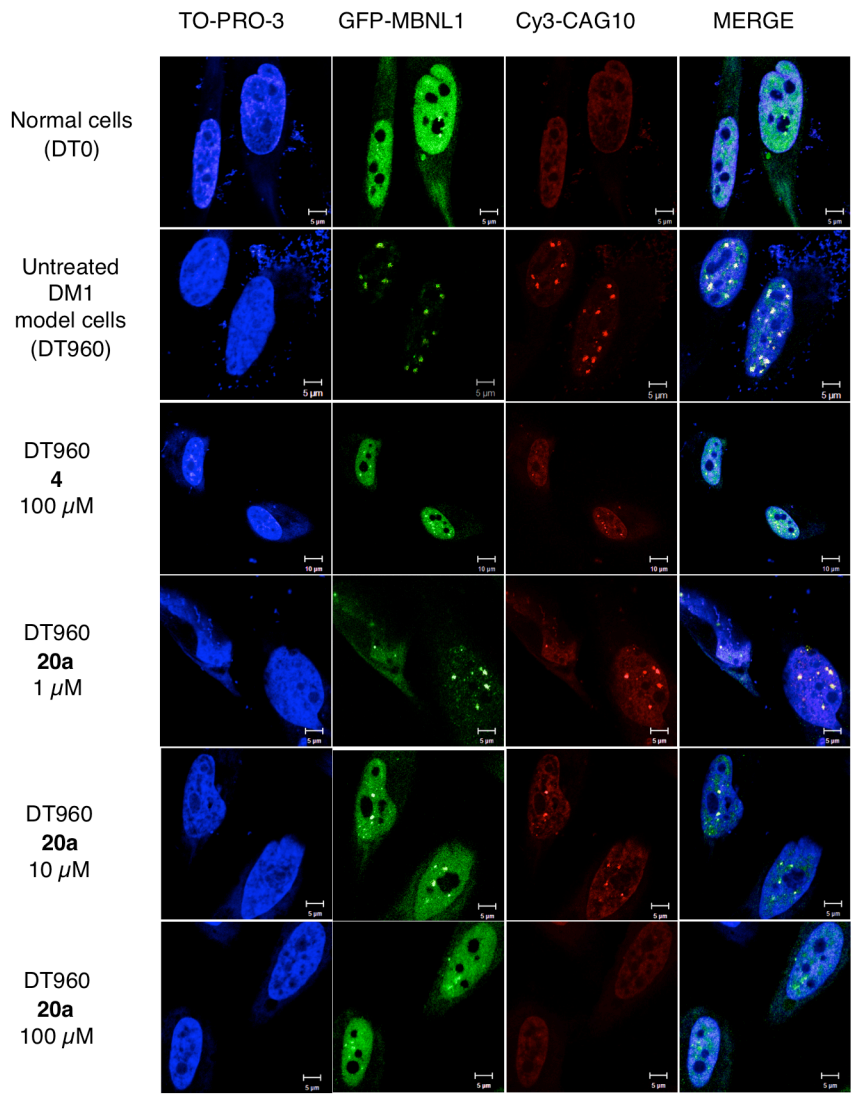


Figure 14. Representative confocal microscopy images. r(CUG)$_{960}$ was visualized by Cy3-(CAG)$_{10}$, and appeared as red dots in the second column. The nuclei was stained by TO-PRO-3, and appeared as blue in the third column. Images in columns 1−3 were merged in column 4. (This experiment was conducted by Dr. Lien Nguyen)

penetration issue for **20a**.

The foci dispersion was also carried out by Dr. Lien Nguyen using exogenous GFP-MBNL1 with similar results (Figure 14). For the normal cells, HeLa cells were transfected with a plasmid containing GFP-MBNL1. DM1 model cells were generated by transfecting HeLa cells

with 2 plasmids, GFP-MBNL1 and a DMPK plasmid containing $(CTG)_{960}$. The MBNL1 were labeled in green by GFP, whereas $r(CUG)_{960}$ was visualized by the FISH probe Cy3-$(CAG)_{10}$ and was represented by the red color. In normal cells, MBNL1 was distributed in the entire nucleus. Because of the absence of rCUG repeats, no red signal of Cy3 was observed. In DM1 model cells, MBNL1 protein was concentrated in bright green spots, which were co-localized with the red spots of rCUG repeats in the nucleus. It was found that the number of nuclear foci was decreased to 70%, 30%, and 12% after treatment with **20a** for 48 h at 1, 10, 100 μM, respectively. On the other hand, treatment of ligand **4** at 100 μM for 48 h incubation resulted in a reduction of foci to ca. 60%.

3.4.2 Splicing Correction

Previous studies provided evidence for the inhibition of MBNL1-$rCUG^{exp}$ by **20a** both *in vitro* and in DM1 model cells. Therefore, its capacity to correct the splicing defects in DM1

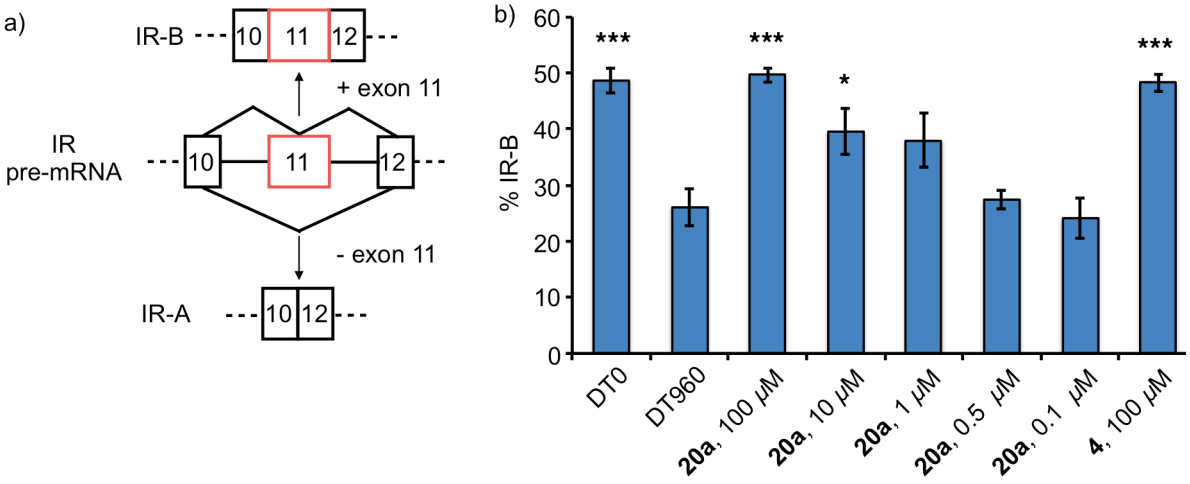


Figure 15. a) Schematic representation of *IR* alternative splicing. b) Correction of *IR* splicing by **20a** and **4** determined by conventional RT-PCR. Error bars represent standard error of the mean from at least three independent experiments, two-tailed *t*-test * $P < 0.05$, *** $P < 0.001$. (This experiment was conducted by JuYeon Lee and Dr. Lien Nguyen)

model cells was addressed in this part using an *IR* minigene (Figure 15). During the alternative splicing of *IR*, two products are generated: isoforms IR-A and IR-B with the exclusion and inclusion of exon 11, respectively (Figure 15a). In normal skeletal muscle, IR-B is the dominant isoform, whereas IR-A is more abundant in DM1 skeletal muscle. [19] We found that normal HeLa cells and the DM1 model cells contained ~49 and 26% isoform B, respectively (Figure 15b). The treatment of **20a** in DM1 model cells showed a dose-dependent reversal of *IR* splicing with a full recovery at 100 μM. Although **20a** exhibited a low nanomolar inhibition constant for the MBNL1-rCUG^{exp} complex *in vitro*, a full recovery of *IR* splicing defect was not achieved at a lower concentration, demonstrating the limited cellular uptake of **20a**.

3.5 Efficacy of Dimers in a DM1 *Drosophila* Model

Determining drug efficacy in a living organism is a critical step in drug development. The *Drosophila* line with tissue-specific expression of 480 interrupted CTG repeats, $i(CTG)_{480}$, shows many key characteristics of DM1, such as nuclear accumulation of rCUG^{exp} and MBNL proteins, CUG-induced eye phenotypes, muscle degeneration, and abnormal alternative splicing of muscle genes.[54,55] In contrast, the fly line expressing 60 uninterrupted CTG repeats, $(CTG)_{60}$, does not exhibit severe pathologies. Several small molecules were reported to improve the CUG-induced phenotypes in the DM1 *Drosophila* model,[56] including our previous studies of **4**.[45] The comparative effectiveness of **4** and **20a** in suppressing the external eye degeneration phenotype and improving larva mobility was examined in the DM1 *Drosophila* model.

Similar to the reported DM1 *Drosophila* model,[54] our transgenic fly model with the *gmr-GAL4* driven expression of $i(CUG)_{480}$ repeats (*gmr-GAL4>UAS-*$i(CTG)_{480}$) showed external eye morphological phenotypes, including disorganized ommatidia and mechanosensory bristles,

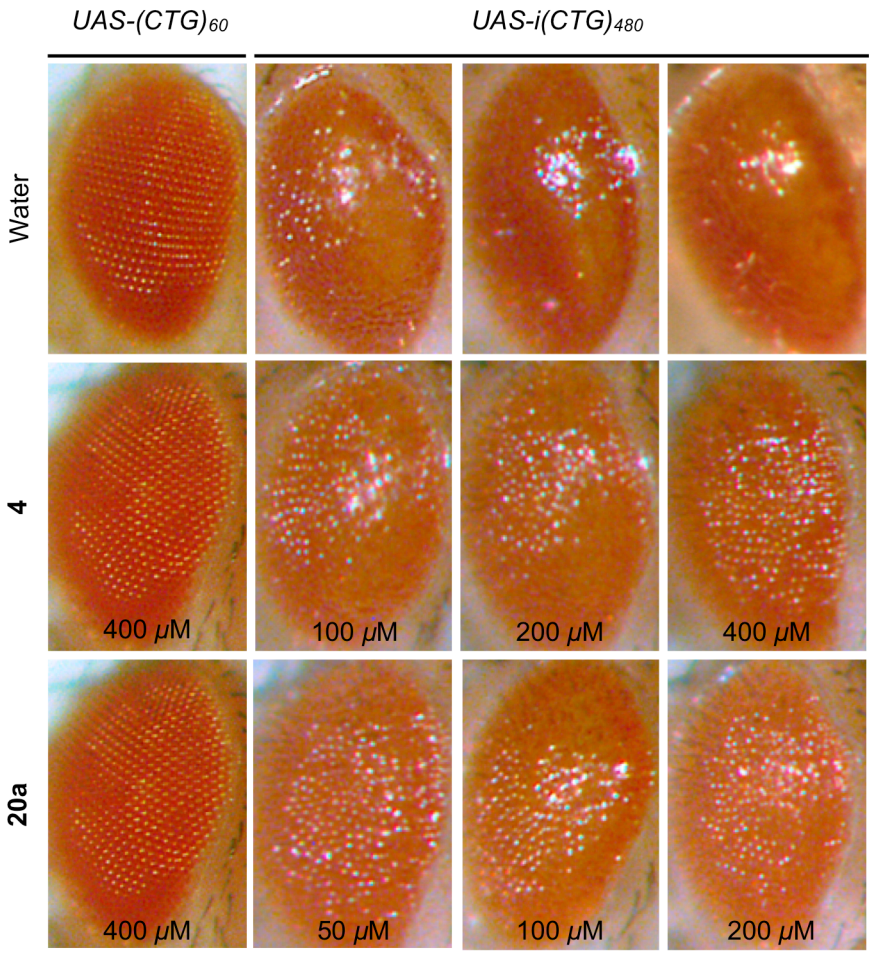


Figure 16. Improvement in adult external eye degeneration phenotype in DM1 *Drosophila* fly model. Each experiment was conducted in triplicate. (This experiment was conducted by Shaohong Peng and Prof. H. Y. Edwin Chan, The Chinese University of Hong Kong)

resulting in a glossy eye (Figure 16). The microscopic images showed that treatment with **20a** at 50 μM led to partial reversal of the degenerative phenotype, whereas the same effect was achieved by monomer **4** at 400 μM (Figure 16). On the other hand, treatments of *(CTG)*$_{60}$ *Drosophila* flies treated with either **4** or **20a** at 400 μM did not change the eye structure.

In addition, **4** and **20a** significantly improved the locomotor ability of *i(CTG)*$_{480}$ DM1 *Drosophila* larvae in a larval crawling assay.[57] In this assay, an individual larva generated from transgenic flies was placed on a Petri dish over graph paper, and allowed to crawl (Figure 17a).

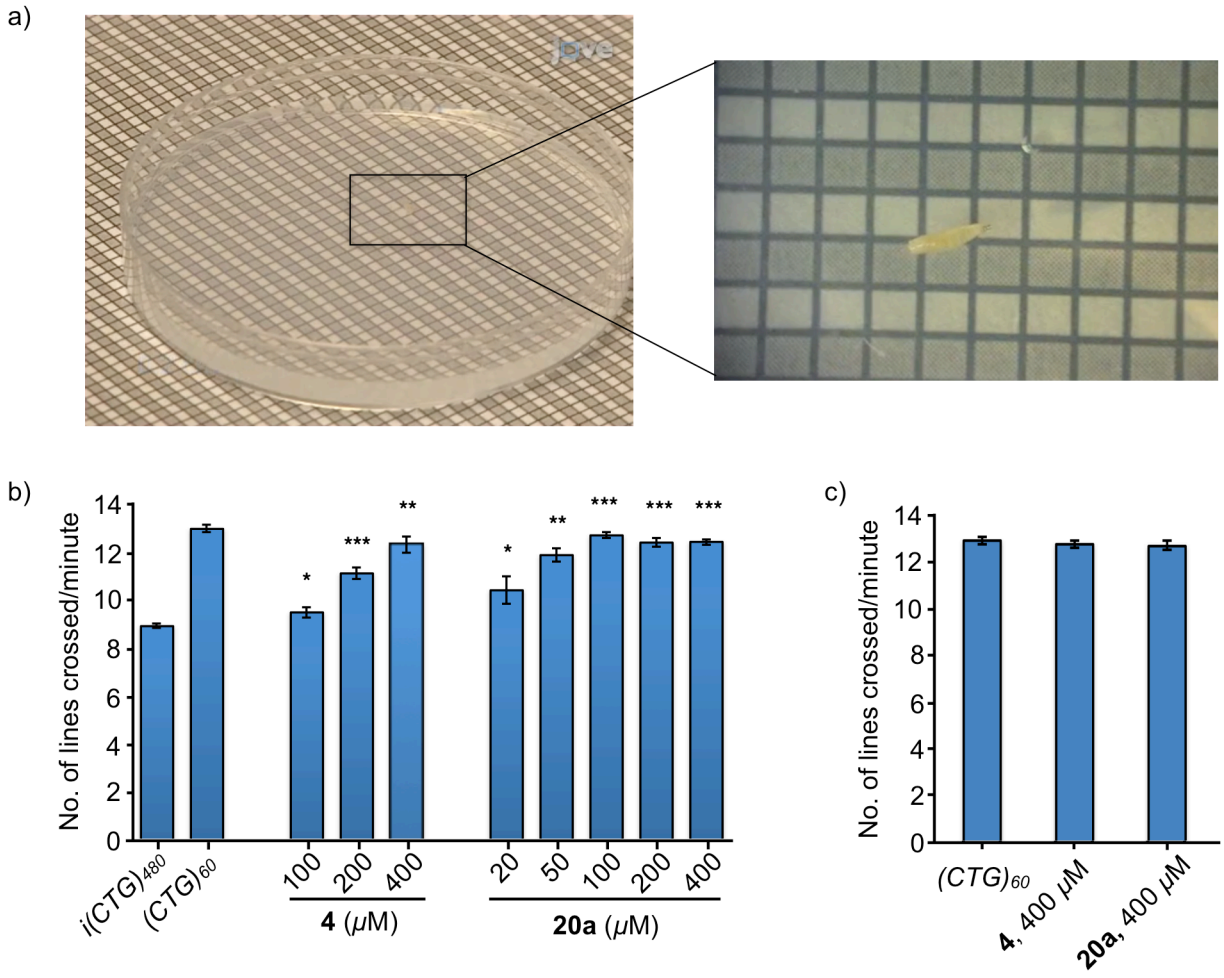


Figure 17. Effect of **4** and **20a** in a larval crawling assay. a) Illustration of the larvae crawling on the Petri dish. Images were adapted from the video in Ref 57. b) Dose-dependent effect of ligands on the DM1 larvae (*24B-GAL4 > UAS-i(CTG)$_{480}$*). (c) Effect of ligands on the control *Drosophila* larvae (*24B-GAL4 > UAS-(CTG)$_{60}$*). Each experiment was conducted in triplicate, 10 individual larvae for each trial. Error bars represent standard deviation, two tailed *t*-test * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. (This experiment was conducted by Shaohong Peng and Prof. H. Y. Edwin Chan, The Chinese University of Hong Kong)

The *(CTG)$_{60}$* control larvae were able to crawl ~13 lines per minute, whereas the *i(CTG)$_{480}$* DM1 larvae could cross only nine lines (Figure 17b), demonstrating their locomotor defect. When the control larvae were raised with an aqueous solution of fly food containing either **4** or **20a** at 400 μM, the mobility was not affected (Figure 17c), suggesting low toxicity of the bisamidinium

ligands. Treatment of the *i(CTG)*$_{480}$ DM1 larvae with **4** at 100, 200, 400 μM rescued ~14, 54, and 84% of the crawling defect, respectively, demonstrating dose-dependent response (Figure 17b). Compared with **4**, dimer **20a** showed greater efficacy with a nearly full recovery at 100 μM. Even at concentrations as low as 20 μM, **20a** induced a 37% improvement in the DM1 larva's mobility. To our knowledge, this is the first example of a small molecule showing a substantial improvement in locomotor behavior or muscle performance defect of the DM1 *Drosophila* larvae.

However, the maximum tolerated dose (MTD) of **20a** in mice was a relatively low 3 mg/kg as determined by daily intraperitoneal injections.

3.6 Conclusions and Outlook

In summary, dimeric ligand **20a**, prepared by copper-catalyzed click reaction, exhibits an excellent *in vitro* inhibition of the MBNL1-rCUG^{exp} complex, but its desired efficacy in DM1 model cells was not observed. The potential issues with cell permeability may be attributed to its high molecular weight and increased number of hydrogen bond donors and acceptors.[58] However, the clear dose-dependent phenotypic improvement in the *Drosophila* models supports the notion that the bis-amidinium dimer may be useful if delivery and toxicity issues can be addressed. Two clear solutions involve the use of a delivery vehicle or *in vivo* self-assembly of the clickable monomers. The latter approach is particularly appealing, because monomer **4** is nontoxic to mice at 50−100 mg/kg.[45] Efforts along these lines will be reported in due course.

The use of copper catalyst to assemble the dimer and trimer is an alternative solution. Recently, Zimmerman and coworkers developed a copper nanoparticle that was found to catalyze the alkyne-azide click reaction at a ppm concentration, making this catalyst relatively

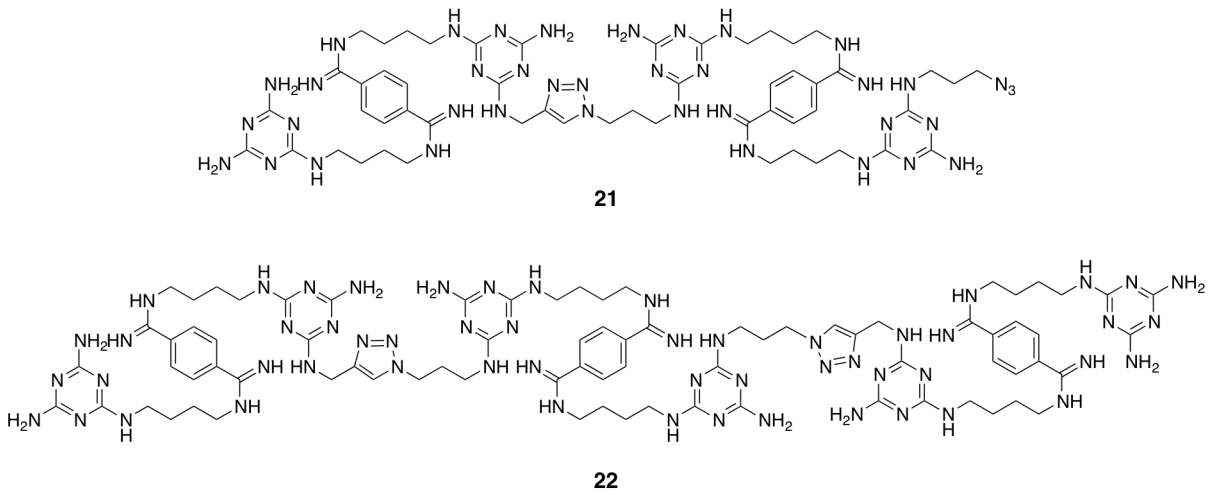
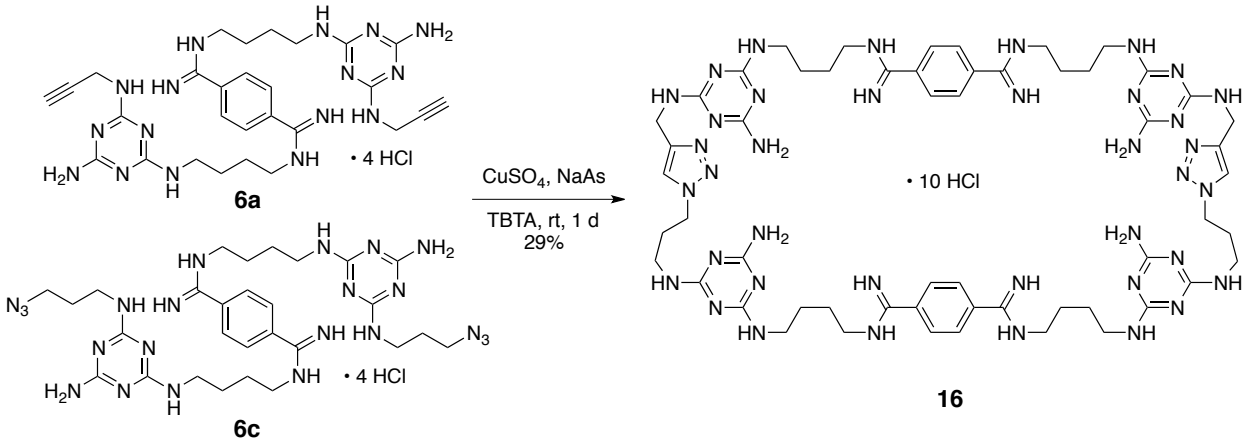


Figure 18. Structures of dimer **21** and trimer **22**.

nontoxic to cells at the working concentration.[59] The evaluation of this catalyst in the cycloaddition of diazide **6c** and alkyne **19a**, which potentially produces dimer **21** and trimer **22** (Figure 18), is currently ongoing.

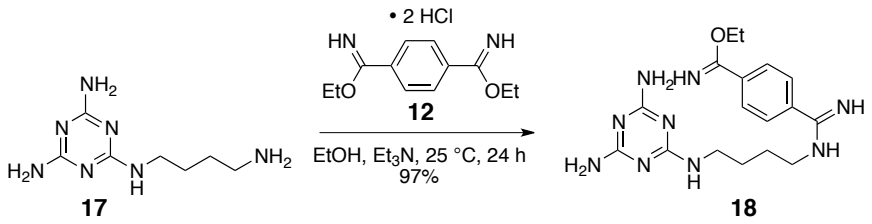
3.7 Materials and Methods

3.7.1 Synthesis



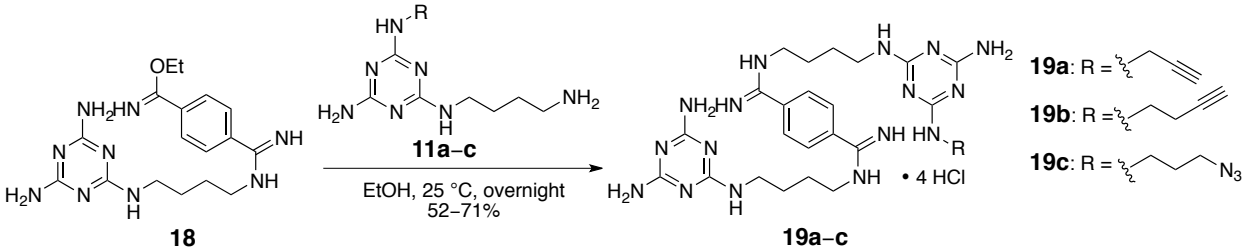
Compound **16**. To a solution containing 33 mg (0.042 mmol) of **6a** and 49 mg (0.043

mmol) of **6c** in 10 mL of MeOH was added 127 μL of DMSO solution of TBTA 0.1 M, 849 μL of aqueous solution of sodium ascorbate 0.1 M, and 212 μL of aqueous solution of $Cu_2SO_4.5H_2O$ 0.1 M. The mixture was stirred at room temperature for 24 h. Solvents were removed using a rotary evaporator. The resulting solid was suspended in 10 mL water. The suspension was filtered using a syringe filter (28mm diameter, 0.45 μm pore SFCA membrane, Corning® #431220) to obtain a clear solution. Concentrated aqueous solution of HCl was added to adjust to pH of the solution to 5−6. The resulting solution was loaded onto the Sephadex CM-25 column, and purified using aqueous solution of NH_4HCO_3 from 0.5 to 0.9 M. Fractions containing the product was collected, concentrated using a rotary evaporator at 60 °C. The solid was dissolved in 5 mL of aqueous solution of HCl 1 M. The resulting solution was frozen in an dry-ice bath, and dried in a lyophilizer for 2 days, gave 20 mg (29%) of **16** as a white solid. ^{1}H NMR (DMSO-d_6): δ 10.33 (s, 4H), 9.84 (s, 4H), 9.49 (s, 4H), 8.81−7.93 (m, 24H), 4.55−4.50 (m, 4H), 4.41−4.38 (m, 4H), 3.51−3.49 (m, 8H), 3.34−3.27 (m, 12H), 2.05 (q, J = 6.5, 4H), 1.68−1.60 (m, 16H). ^{13}C NMR (DMSO-d_6): δ 161.50, 155.84, 155.74, 132.75, 128.70, 128.66, 123.25, 47.16, 42.44, 37.66, 35.98, 29.52, 25.91, 24.65.



Compound 18. To a white suspension of 5.0 g (17.1 mmol) of **12** in 30 mL of anhydrous EtOH was added 7.5 mL (53.8 mmol) of anhydrous Et_3N, resulting in a colorless solution. A suspension of 1.1 g (5.7 mmol) of **17**[41] in 10 mL of anhydrous EtOH was added slowly to the reaction mixture. The reaction was stirred at room temperature for 24 h. The white solid was

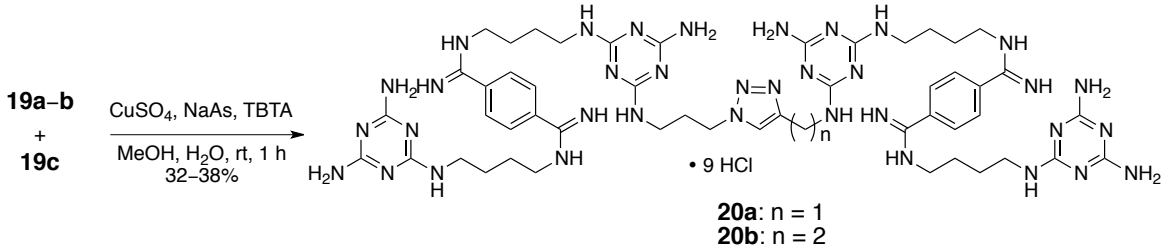
filtered, washed extensively with CH_2Cl_2, dried under nitrogen overnight to give 2.1 g (97 %) of compound **18** as a white solid. Compound **18** was used for the next step without further purification because the imidate ester is water sensitive. LR-ESI-MS (*m*/*z*) calculated for $[M + H]^+$: 372.2, found: 372.2.



Compound **19a**. A suspension of 4.0 g (10.8 mmol) of **18** and 2.8 g (11.9 mmol) of **11a** in 50 mL of anhydrous EtOH was stirred at room temperature overnight. The solvent was removed using a rotary evaporator. The resulting white solid was purified by silica gel column chromatography with a gradient mixture of CH_2Cl_2/ MeOH from 9/1 to 1/9 (v/v). After using 1 L of CH_2Cl_2/MeOH (1:9, v/v), the eluent was acidified with a gradient of 0.10−0.2 mL of dioxane solution of HCl 4 M per liter of eluent. The product-containing fractions were combined, filtered, and concentrated using a rotary evaporator to give 5.4 g (71%) of **19a** as a white tetra-HCl salt. ^{1}H NMR: δ 10.11 (s, 2H), 9.73 (s, 2H), 9.30 (s, 2H), 7.95 (s, 4H), 6.87−6.54 (m, 3H), 6.10−5.95 (m, 6H), 3.97 (s, 2H), 3.45 (t, J = 7.1, 4H), 3.25−3.21 (m, 4H), 3.01 (t, J = 2.4, 1H), 1.70−1.53 (m, 8H); LR-ESI-MS (*m*/*z*) calculated for $[M + H]^+$: 561.3, found: 561.4.

Compound **19b**. Using the procedure described above, 0.3 g (0.8 mmol) of **18** and 5 mL (1.0 mmol) of 0.2 M ethanolic solution of **11b** afforded 0.3 g (52%) of **19b** as a white tetra-HCl salt. LR-ESI-MS (*m*/*z*) calculated for $[M + H]^+$: 575.4, found: 575.5.

Compound **19c**. Using the procedure described above, 3.5 g (9.4 mmol) of **18** and 2.9 g (10.4 mmol) of **11c** afforded 4.3 g (61%) of **19c** as a white tetra-HCl salt. ^{1}H NMR: δ 10.13 (s, 2H), 9.74 (s, 2H), 9.32 (s, 2H), 7.95 (s, 4H), 6.63−6.48 (m, 3H), 6.11−5.95 (m, 6H), 3.45 (t, J = 7.2, 4H), 3.38−3.54 (m, 4H), 3.24−3.21 (m, 4H), 1.74−1.53 (m, 10H); LR-ESI-MS (m/z) calculated for $[M + H]^+$: 606.4, found: 606.2.



Compound **20a**. To a solution containing 34 mg (0.049 mmol) of **19a** and 37 mg (0.049 mmol) of **19c** in 11 mL of MeOH was added 148 μL (0.015 mmol) of DMSO solution of TBTA 0.1 M, 246 μL (0.025 mmol) of aqueous solution of $CuSO_4.5H_2O$ 0.1 M. To the resulting solution was added 985 μL (0.099 mmol) of aqueous solution of sodium ascorbate 0.1 M. The mixture was stirred at room temperature for 24 h. Solvents were removed using a rotary evaporator. The resulting solid was suspended in 10 mL water. The suspension was filtered using a micro-filter to obtain a clear solution. Concentrated aqueous solution of HCl was added to adjust the solution to pH 5−6. The resulting solution was loaded onto the Sephadex CM-25 column, and purified using an aqueous solution of NH_4HCO_3 from 0.1 to 0.9 M. Fractions containing the product was collected, concentrated using a rotary evaporator at 60 °C. The solid was dissolved in 5 mL of aqueous solution of HCl 1 M. The resulting solution was frozen in an dry-ice bath, and dried in a lyophilizer for two days, giving 27 mg (38%) of **20a** as a white solid. ^{1}H NMR: δ 10.30 (s, 4H), 9.83 (s, 4H), 9.50−9.46 (m, 4H), 8.58−7.86 (m, 26H), 4.57−4.52 (m, 2H), 4.42−4.39 (m, 2H), 3.51−3.47 (m, 8H), 3.38−3.27 (m, 10H), 2.08−2.06 (m, 2H), 1.72−1.59

(m, 16H); ^{13}C NMR: δ 161.61, 132.84, 128.70, 128.68, 47.08, 42.46, 29.52, 25.98, 24.61; LR-ESI-MS (*m*/*z*) calculated for $[M + H]^+$: 1166.7, found: 584.0 (35%) $[M + 2H]^{2+}$, 389.7 (60%) $[M + 3H]^{3+}$, 292.6 (100%) $[M + 4H]^{4+}$; HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 3.9 min (100%).

Compound **20b**. Using the general procedure described above, 39 mg (0.054 mmol) of **19b** and 40 mg (0.054 mmol) of **19c** afforded 26 mg (32%) of **20b** as a white solid: ^{1}H NMR: δ 10.30 (s, 4H), 9.83 (s, 4H), 9.50−9.47 (s, 4H), 8.51−7.87 (m, 26H), 4.39 (t, J = 7.2, 2H), 3.56−3.47 (m, 10H), 3.36−3.26 (m, 10H), 2.90−2.86 (m, 2H), 2.07−2.04 (m, 2H), 1.71−1.59 (m, 16H); ^{13}C NMR: δ 161.56, 155.70, 132.79, 128.63, 122.90, 47.01, 42.43, 25.93, 24.56; LR-ESI-MS (*m*/*z*) calculated for $[M + H]^+$: 1180.7, found: 591.1 (15%) $[M + 2H]^{2+}$, 394.4 (35%) $[M + 3H]^{3+}$, 296.1 (45%) $[M + 4H]^{4+}$, 237.1 (60%) $[M + 5H]^{5+}$, 197.7 (100%) $[M + 6H]^{6+}$; HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 3.8 min (97%).

3.7.2 Biological Assays

Protein expression and purification. An expression vector for MBNL1N was obtained from Maurice S. Swanson (University of Florida, College of Medicine, Gainesville, FL, USA). MBNL1N is comprised of the four zinc finger motifs of MBNL1 and contains a His_6 tag. The protein was expressed and purified as described previously.[45] The molecular weight of MBNL1N was determined by MALDI-MS, and the concentration was determined by an amino acid analysis.

Determination of dissociation constant (K_D) of the MBNL1-r(CUG)$_{16}$ complex by EMSA. r(CUG)$_{16}$ was purchased from Integrated DNA Technology (Coralville, IA), dissolved in RNA storage buffer containing 1 mM sodium citrate pH 6.4 (Ambion® #AM7001). r(CUG)$_{16}$

RNA was labeled with [γ-^{32}P]-ATP using T4 poly-nucleotide kinase (New England Biolabs). Labeled RNA was purified by phenol extraction and ethanol precipitation. Labeled RNA was heated at 95 °C for 5 min, then placed on ice for 10 min and diluted in protein binding buffer (175 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-HCl (pH 8), 1.25 mM 2-mercapto-ethanol (BME), 12.5% glycerol, 2mg/ml bovine serum albumin (BSA), 0.1 mg/ml heparin, 0.05% or 0.1% Triton X). For MBNL1-r$(CUG)_{16}$ binding study, MBNL1 was serially diluted by a factor of 3 in protein binding buffer. To each solution was added labeled RNA, so that the final concentration of RNA was 0.2 nM. The resulting mixtures were incubated at room temperature for 10 min and loaded onto a 6% polyacrylamide gel (80:1) at 4 °C. The gel was run for 2 h at 180 V in Tris-Borate 0.5X buffer (pH 8.2–8.3). Gels were dried at 80 °C under high vacuum for 2 h, developed overnight in phosphor exposure cassette, and visualized on a Molecular Dynamics Storm PhosphorImager. The K_D (1:1 stoichiometry assumption) was obtained by fitting the normalized fraction RNA bound versus the concentration of protein using the equation:

$$\text{Fraction RNA bound} = B_{max} \times [\text{MBNL1}]_{total}^{h}/(K_D^{h} + \text{MBNL1}]_{total}^{h})$$

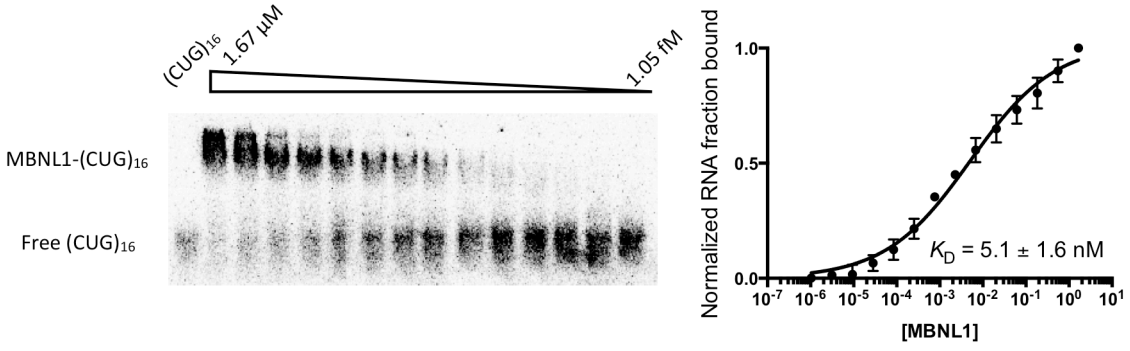


Figure 19. Dissociation constant of the MBNL1-r$(CUG)_{16}$ interaction determined by EMSA. The concentration of the protein was serially diluted by the factor of 3 from 1.67 μM to 1.05 fM. The concentration of r$(CUG)_{16}$ was 0.2 nM. Error bars represent standard deviation from three independent experiments.

where B_{max} is maximum fraction RNA bound, h is a Hill slope. Protein concentration in reaction mixture was larger 10-fold excess over RNA concentration.

Inhibition of MBNL1-r(CUG)$_{16}$ complex by small molecules. Labeled RNA was heated at 95 °C for 5 min, then placed on ice for 10 min. The MBNL1 and labeled r(CUG)$_{16}$ were diluted in protein binding buffer (175 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-HCl (pH 8), 1.25 mM 2-mercapto-ethanol (BME), 12.5% glycerol, 2mg/ml bovine serum albumin (BSA), 0.1 mg/ml heparin, 0.05% or 0.1% Triton X). The complex was incubated at room temperature for 10 min. For the inhibition assay, compound was serially diluted by a factor of 2 in water. To each solution was added the MBNL1-r(CUG)$_{16}$ complex, so that the final concentrations of RNA and protein were 0.2 and 60 nM, respectively. The resulting mixtures were incubated at room temperature for 10 min, and loaded onto a 6% polyacrylamide gel (80:1) at 4 °C. The gel was run for 1.5 h at 180 V in Tris-Borate 0.5X buffer (pH 8.2−8.3). Gels were dried at 80 °C under high vacuum for 2 h, developed overnight in phosphor exposure cassette, and visualized on a Molecular Dynamics Storm PhosphorImager. The IC_{50} value was obtained by fitting the normalized fraction RNA bound versus the concentration of compound using the equation in Prism:

$$y = ax^b/(IC_{50}{}^b + x^b)$$

where y is the fraction of RNA bound, x is the concentration of compound, $a = y_{max} - y_{min}$, b is a Hill slope. The average IC_{50} and standard deviation were obtained from three independent experiments.

The apparent inhibition constant (K_I) was determined using the equation:

$K_I = IC_{50} \times K_D / [\text{protein}]_{total}$, where $[\text{protein}]_{total} = 0.06\ \mu M$, K_D is the dissociation constant of the MBNL1-RNA complex and [protein] is at least 7-fold greater than the K_D.

Binding study using ITC. The ITC experiment was performed as described in Chapter 2. The concentrations of dimer **20a** in the syringe and r(CUG)$_{16}$ in the cell are 500 μM and 10 μM, respectively. [MOPS] = 20 mM, [NaCl] = 150 mM.

Foci dispersion in DM1 models cells using GFP-MBNL1. The foci disruption experiments using GFP-MBNL1 were conducted by Dr. Lien Nguyen as described previously.[43] Normal cells were produced by transfecting HeLa cells with a plasmid containing GFP-MBNL1 and a *DMPK* plasmid containing no CTG repeats. To generate DM1 model cells, HeLa cells were transfected with a GFP-MBNL1 plasmid and a *DMPK* plasmid containing (CTG)$_{960}$. The DM1 model cells were treated with ligands **4** and **20a** for 48 h. The Cy3-(CAG)$_{10}$ was employed to detect rCUG repeats, while the nuclei were stained with Hoechst 33342. The cells were visualized using the confocal microscope LSM 700 (Zeiss). GFP tag, Cy3 and Hoechst 33342 dyes were excited at 488, 564, and 405 nm, respectively. Each treatment was repeated at least three times.

Splicing assays. Splicing assays were carried out essentially by Dr. Lien Nguyen and JuYeon Lee as previously described.[45] Briefly, cells were seeded in each well in a 6-well plate the day before transfection to afford 70−80% of confluence on the following day. The cells were co-transfected with 500 ng of *IR* minigene and either 750 ng of DT960 or 500 ng of DT0 minigene with Lipofectamine 2000 (Invitrogen, Life Technologies) according to the manufacturer's protocol. After 4h, the media was changed to the complete growth media, and cells were treated with the compounds and harvested after 72h. For dose dependence experiments, cells were treated with **20a** at concentrations of at 10, 1, 0.5, 0.1 μM. And for the comparison of **20a** and **4**, 100 μM of each compound were treated to the cells. To harvest the cells, cells were detached using trypsin with 0.05% EDTA (Fisher Mediatech) and collected by

centrifugation at 500 g for 5 min. Out of the collected cells, was RNA immediately isolated using Total RNA kit I (Omega Bio-Tek). cDNA was synthesized from 1 μg of isolated RNA with iScript cDNA synthesis kit (Bio-Rad). The reverse transcription reaction was purified using QIAquick PCR purification kit (Quiagen). Then, 70 ng of cDNA was used for PCR amplification of IR isoforms A and B using specific primers. PCR cycle was optimized to 31–35 for each assay. The forward primer was 5'- GTA CCA GCT TGA ATG CTG CTC CT, and the reverse primer was 5'-CTC GAG CGT GGG CAC GCT. The PCR products were separated by running 8% nondenaturing polyacrylamide gel with 1X TBE (National Diagnostics) at 150 V for 30 min. The gel was post stained with ethidium bromide and subsequently imaged using Gel Doc XR+ system (Bio-Rad). The bands were quantified using ImageJ (NIH).

Drug treatment in *Drosophila*. Fly lines bearing *UAS-(CTG)$_{60}$ and UAS-i(CTG)$_{480}$*[54] were kind gifts of Prof. Rubén Artero Allepuz (Universitat de València, Estudi General, Spain). The external eye and larval crawling assays were performed by Shaohong Peng and Prof. H. Y. Edwin Chan, The Chinese University of Hong Kong as previously described.[49]

Determination of murine maximum tolerated dose (MTD). All experimental procedures were reviewed and approved by the University of Illinois Institutional Animal Care and Use Committee (Protocol #15232). Compound was formulated in sterile water and administered to C57BL/6 female mice (12-14 weeks old, Charles River) intraperitoneally for five consecutive days. All the mice were monitored over the course for signs of toxicity and weight loss. This experiment was performed by Hyang Yeon Lee, Prof. Paul J. Hergenrother, UIUC.

CHAPTER 4

IN VITRO SELECTION OF LIGANDS CAPABLE FORMING MORE POTENT AGENTS ON BOTH DNA/RNA TARGETS[3]

4.1 Introduction

Fragment-based drug discovery (FBDD)[60] and target-guided synthesis (TGS)[60-63] have emerged as promising approaches to discover new lead therapeutic agents. Both FBDD and TGS involve a suite of different strategies that share both conceptual and operational similarities, including the formation of tight binding ligands from smaller fragments. The most relevant to this work is the *in situ* TGS approach, wherein the target receptor serves as a template, binding two or more fragments in such a way that their reactive functional groups are in close proximity. This leads to an accelerated coupling reaction that produces a tighter binding ligand.

The utilization of click chemistry has significantly accelerated the development of TGS. A number of tight-binding ligands for targeting proteins have been discovered in recent years using this technique.[64,65] In contrast, the use of DNA and RNA targets with TGS has progressed more slowly. In 2003, Dervan and coworkers demonstrated the ability of DNA to catalyze the alkyne-azide [3+2] cycloaddition of hairpin polyamides, giving a heterodimeric product with a greater binding affinity and specificity.[66] Click and native chemical ligation reactions between small molecules formed on DNA or RNA templates have been also reported by other groups,[67-69] however, the number of examples are few in comparison to protein targets.

We are particularly interested in applying a TGS approach to both the DNA and RNA

[3] Some of the mateial in this chapter was adapted from the following publication:
Luu, L. M.; Lee, J.; Nguyen, L.; Vergara, J. I.; Zimmerman, S. C. A Fragment-Based, Multi-target Drug Discovery Approach for Both DNA and RNA Targets Using a Template-Assisted Selection. (Manuscript in preparation)

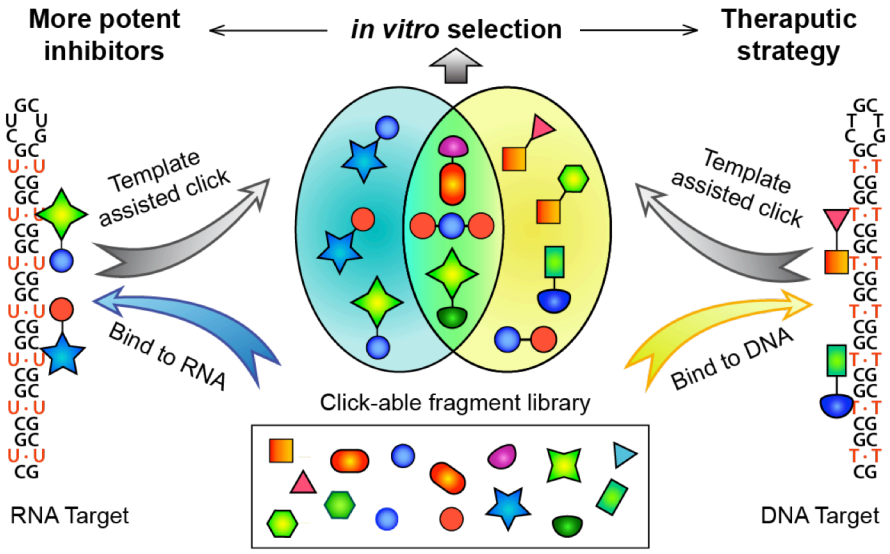


Figure 20. Illustration of the selection of multi-target agents using the click reaction templated by dCTG and rCUG repeats. (This image was obtained from Dr. Yugang Bai)

targets causing DM1. The repetitive nature of $dCTG^{exp}$ and $rCUG^{exp}$ makes them well suited to FBDD. Specifically, we are intrigued by the discovery of dimeric ligands for DM1 reported by Miller and coworkers in 2008.[70] They developed a bead-segregated, dynamic combinatorial library containing as many as 11,325 compounds equilibrated through a disulfide exchange reaction of 150 monomers. Among those dimeric ligands, the tightest binder was selected by a fluorescently labeled $r(CUG)_{10}$. While this is a promising approach to discover new RNA binders *in vitro*, the high abundance of thiols in a cellular system makes the disulfide-containing compounds potentially unstable.

Herein, we describe an approach that uses the template-assisted, irreversible alkyne-azide click reaction in a selection format with both $d(CTG)_{16}$ and $r(CUG)_{16}$ targets (Figure 20). A limited number of products were formed in each case and even fewer formed in both selections, suggesting this method can be useful for discovering both selective single and multi-target lead therapeutic agents.

4.2. Design and Synthesis of the Clickable Compound Library

To enable both FBDD and TGS to discover multi-target agents, initial focus was placed on ligands **1** and **4**, which were rationally designed to selectively target dCTG^{exp} or rCUG^{exp} hairpin structures. The binding mechanisms of these ligands involve 1) specific recognition of the T-T and U-U mismatch by the triaminotriazine unit and 2) acridine intercalation (**1**) or 3) RNA binding by A-form RNA-bisamidinium groove complexation (**4**). Indeed, ligand **1** was reported to bind d(CTG)$_4$ and r(CUG)$_4$ with K_D = 0.5 ± 0.2 μM and K_D = 0.43 ± 0.11 μM, respectively.[41] Ligand **4** shows strong affinity for r(CUG)$_{12}$ (K_D = 8 ± 2 μM) but no binding to dCTG repeats.[45] As discussed in Chapter 2, an aminoalkyl substituted analog of **4**, where the substituent is attached to the triaminotriazine moiety, binds both dCTG and rCUG repeats.[49]

Over the past two decades, the alkyne-azide cycloaddition has been widely employed in bioconjugation, because of its excellent bioorthogonality.[71,72] The small size of alkyne and azide groups makes them suitable for conjugating to both small- and macro-molecules. As discussed in chapter 2, a clickable moiety can be attached to the triaminotriazine ring of the bisamidinium containing ligands with minimal effect on binding to rCUG^{exp}. Therefore, compounds **6a**−**c** were included in the library, serving as clickable rCUG binders (Figure 21). In addition, the activated alkyne **24** was chosen with the expectation that it would undergo an accelerated click reaction.[66,73] In addition to azides, amine **23** was employed to examine the potential of a Michael addition to the activated alkyne. For the acridine compounds, clickable functionality was conjugated to the acridine ring using an amide coupling reaction. Thus, ligands **25**−**32** containing side-chains of varying lengths and terminated with either an azide, amine, or alkyne group were prepared. In addition, it was found that replacing one triaminotriazine hydrogen in **1** with a methyl group weakens, but does not abolish its binding to a rCUG site.[40] Thus, compound **31**

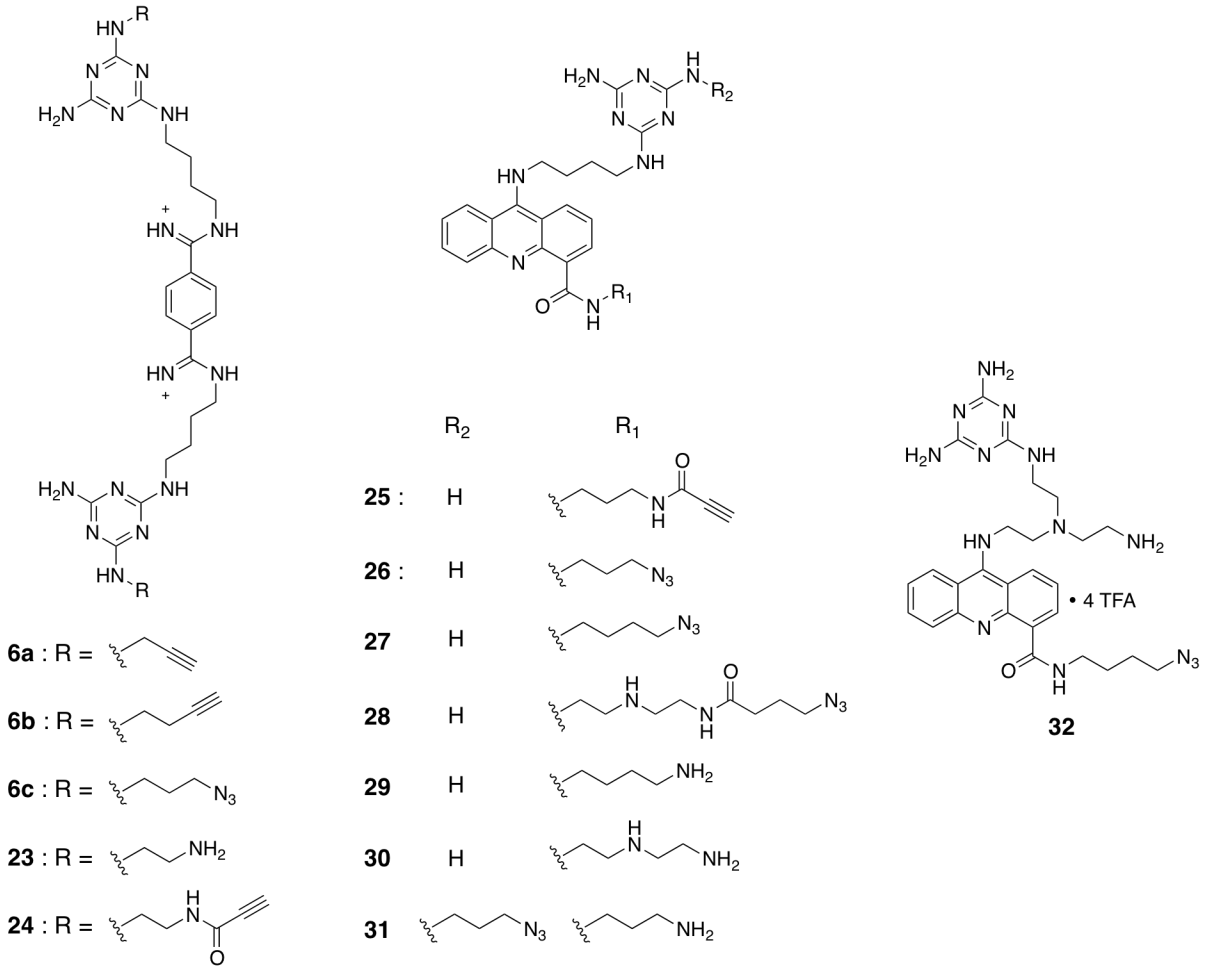
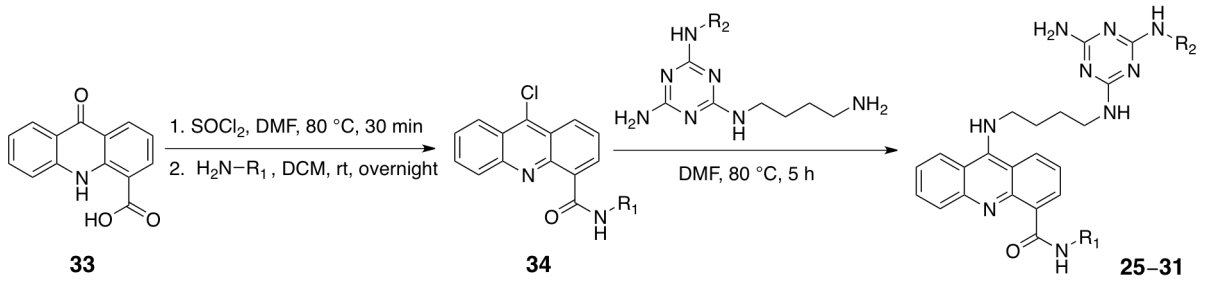


Figure 21. Structures of dCTG and rCUG binding ligands containing clickable groups.

containing a water-solubilizing aminopropyl group on the acridine ring and an *N*-azidopropyl group on the triaminotriazine recognition unit was prepared. The 5-atom linker in **32** was chosen based on our previous findings that compound **1** and an acridine-melamine conjugate with a pentyl linker have comparable binding affinities to rCUG repeats. Moreover, compound **32**, which contains two amino groups, is water-soluble and potentially serves as a Michael donor for the activated alkyne ligands **24** and **25**. It is note worthy that a derivative of **28** with an activated alkyne was found to undergo intramolecular Michael addition with a secondary amine.

Generally, the clickable bisamidinium compounds **6a**−**c**, and **23** were prepared as

Scheme 6. General synthesis of acridines **25**−**31**.



described in the previous chapters. The *N,N′*-Dicyclohexylcarbodiimide (DCC)-promoted coupling reaction of amine **23** with propiolic acid afforded an electron-deficient alkyne **24**. The general synthetic approach to acridine compounds **25**−**31** is outlined in Scheme 6. Thus, acridone **33**[43] was treated with thionyl chloride, followed by the amide formation with corresponding amines to give acridine **34** in 54−88% yield. Substitution reaction of **34** and an amine-containing triaminotriazine in DMF gave desired acridine ligands in 21−73% yield. Additional steps were needed for the preparation of several other compounds (see the Synthesis section for details). Compounds **6a**−**c**, **23**−**32** were isolated as either HCl or TFA salts, and were at least 95% pure as indicated by ^{1}H NMR and analytical HPLC, except for **31** and **24** (93% and 86% pure, respectively).

To increase both size and structural diversity, general DNA/RNA binders **35**−**59** were added to the library (Figure 22). Ligands **35** and **36** containing the bisamidinium core of **4** are potential major groove binder for A-form RNA,[46] whereas acridines **38** an **39** are intercalators that do not contain the triaminotriazine recognition unit. Alkyne **37** contains a naphthrydine moiety, which was part of molecules targeting G-A and G-G mismatch DNA.[74] Compound **40**, a derivative of **5** with an azide chain, was obtained from Julio Serrano. Although binding modes of **5** and **40** to dCTG and rCUG repeats remain to be elucidated, their tricyclic ring may insert

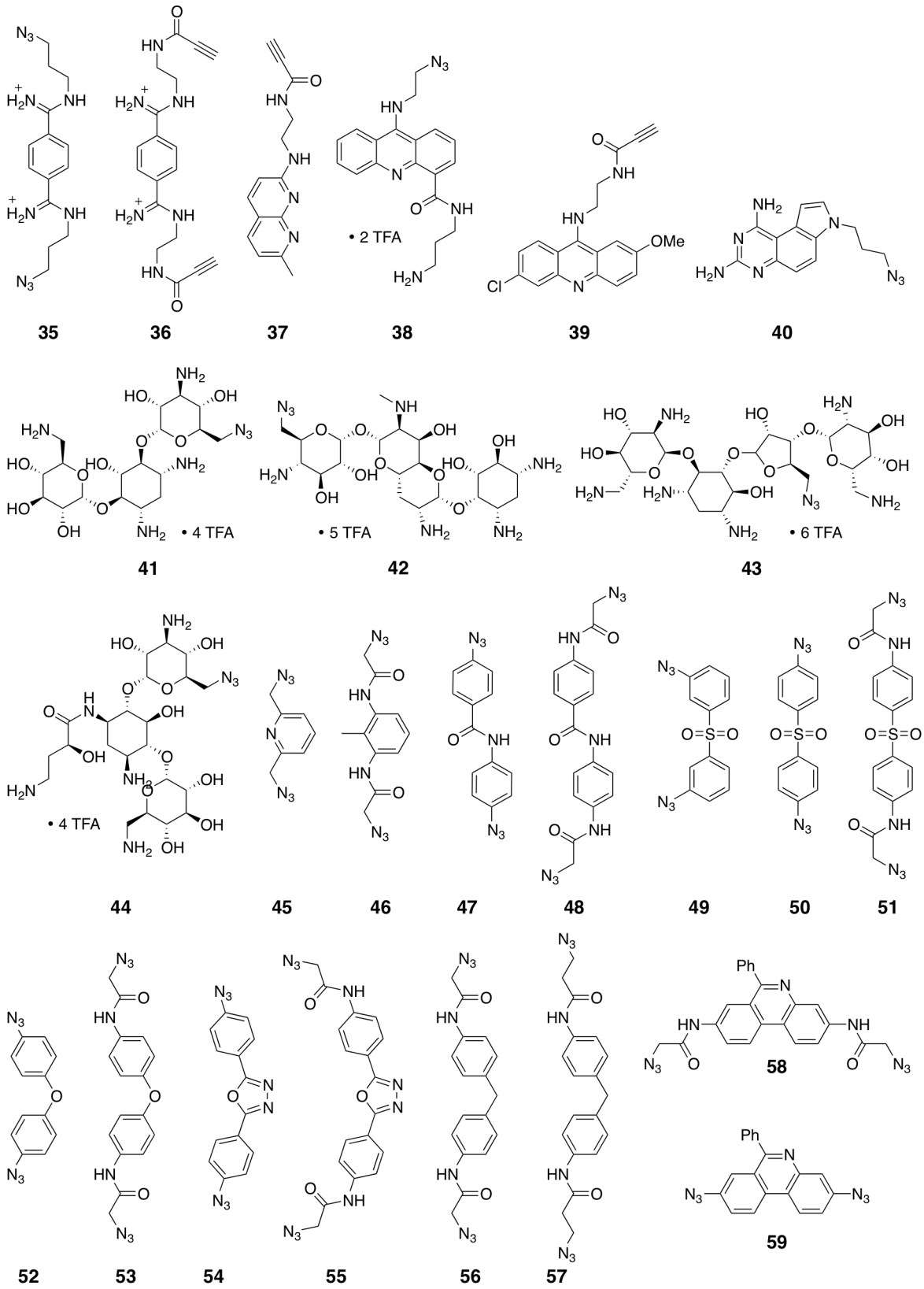


Figure 22. Structures of general DNA/RNA binders **35**−**59** containing clickable groups.

between DNA/RNA base pairs, whereas the amino groups can interact with T-T or U-U mismatches through hydrogen bonds. Aminoglycosides **41**−**44**, which are general RNA binders through electrostatic interactions, were prepared as previously reported.[75] Generally, the Boc-protected aminoglycosides underwent tosylation to convert the primary alcohol to a tosyl group, which was subsequently replaced by an azide. It was also found that neomycin B, kanamycin A/B, and amikacin bind various stem-loop mRNAs containing C-C, C-U, and A-A mismatch with nano to low micromolar K_D.[76] Interestingly, azide-containing kanamycin **41** can undergo the click reaction with an alkyne-bearing kanamycin in the presence of rCCUG repeats,[67] whereas neomycin B **43** was found to inhibit the MBNL1-rCUG^{exp} complex *in vitro*.[29] Diazides **45**−**59** were employed by the Hergenrother group as building blocks for the synthesis of deoxystreptamine dimers that selectively bound to RNA tetra- and octaloops with high affinity.[76] Although none of the diazides are known RNA binders, their aromatic structures resemble a number of known DNA and RNA intercalators or groove binders.[77] Especially, **58** and **59** are derivatives of ethidium bromide, an intercalating agent.

4.3 Design of the Selection Experiment

The overall approach is shown schematically in Figure 23. Because the molecular weight of click products are higher than those of monomers, their presence can be easily detected using mass spectroscopy. In addition, most of the click products have distinct molecular weights, allowing MS peaks to be assigned to the corresponding clickable agents. When considering possible methods to monitor the templated click reactions, we chose matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). In comparison to ESI-MS, MALDI generates less multiply charged ions and leads to fewer peaks in the spectrum of heterogeneous samples, making the data easier to analyze.[78] Although MALDI-MS does not give a quantitative

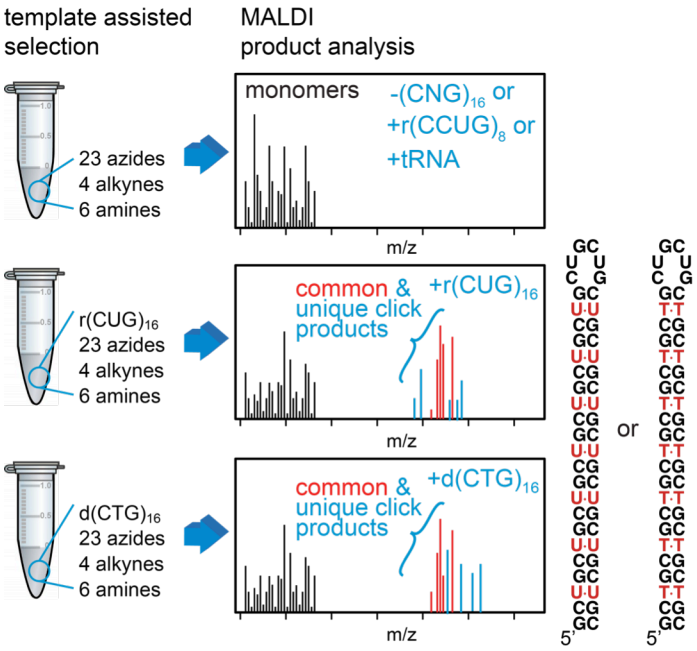


Figure 23. Illustration of the MALDI-MS detection of DNA/ RNA template-assisted selection. Negative selection (top) indicates click products formed off-target, whereas cross-selection shows possible multi-target leads (red peaks) and single target agents (blue).

assessment of the relative amount of each product formed, it has proven to be an excellent initial screen for assessing reaction conditions and degree of selectivity.

In the selection experiment, equal molar amounts of all the compounds (25 μM) were incubated with the template d(CTG)$_{16}$ or r(CUG)$_{16}$ (12.5 μM) at 37 °C and pH 7.0 for various amounts of time. Before MALDI-MS analysis, DNA and RNA were degraded by DNase I and RNase A, respectively, to release tight binders. Parallel control experiments were performed under identical conditions but without template or with other nucleic acid sequences, including a standard 24-mer DNA duplex, a DNA duplex containing d(CTG)$_{16}$-d(CAG)$_{16}$, r(CCUG)$_8$, r(CCCCGG)$_3$, HIV-1 FS RNA, and tRNA (Figure 24). We anticipated that the d(CTG)$_{16}$-d(CAG)$_{16}$ mainly exists as a fully paired DNA duplex. It was found that only 2% of the DNA containing 17 CTG-CAG repeats formed slipped-strand structures.[79] The interest in the rCCUG

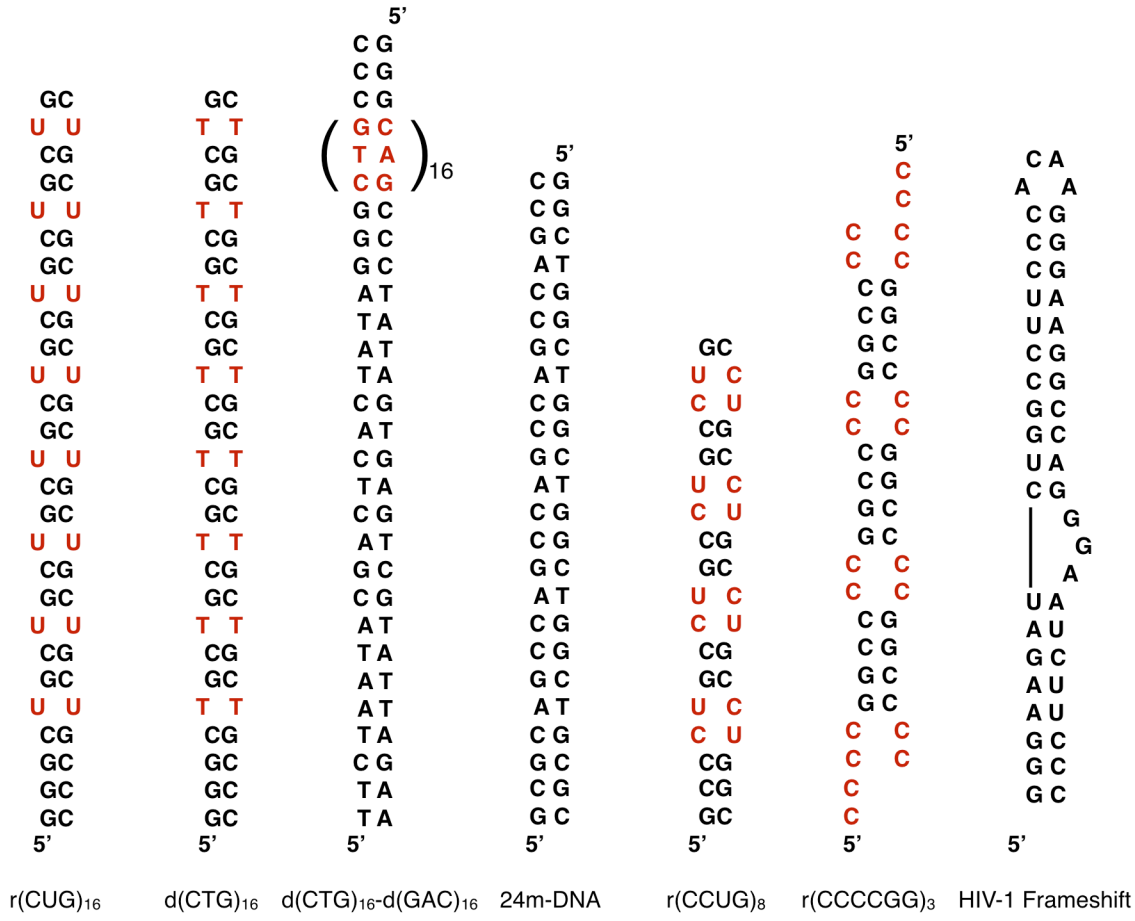


Figure 24. Structures of DNA/RNA sequences.

tetranucleotide repeat arises from its involvement in myotonic dystrophy type 2. Furthermore, both acridine and bisamidinium ligands structurally analogous to those studied here recognize rCCUG repeats selectively.[80,81] In addition, a hexanucleotide repeat r(CCCCGG)$_3$, an antisense RNA involved in c9FTD/ALS disease,[82] is a potential target of the acridine-triaminotriazine compounds, based on the finding that compound **1** binds C-C mismatch with a K_D value of 14 ± 2 μM. On the other hand, the bisamidinium moiety in compound **6a**−**c**, **23**, **24**, **35**, and **36** was found to interact with HIV-1 FS RNA by hydrogen bonds. By using different DNA/RNA sequences, the selectivity of the clickable agents as well as the click products and the ability of this selection method to discover new lead agents for DNA/RNA-mediated diseases can be

quickly accessed.

4.4 Analysis of DNA/RNA Templated Click Selections by MALDI-MS

The initial selection experiment involved 30 compounds, including **6a**−**c**, **23**−**32**, and **35**, **38**, **45**−**59** (Figure 21 and 22). Considering that these compounds may undergo Michael addition or the alkyne-azide cycloaddition that may produce 1,4- and 1,5-adducts, there are 196 possible AB-type adducts and 5,219 ABA/ABC-type coupled products. The templated reactions of dialkyne-, diazide-, and diamine-containing bisamidinium ligands **6a**−**c, 23, 24** by $d(CTG)_{16}$ and $r(CUG)_{16}$, which can form 8 T-T and U-U mismatches, were expected to favor dimeric and trimeric products.

In practice, the idealized results in Figure 23 were observed. Molecular weights of monomers are less than 713 Da, whereas those of dimers are between 716 and 1500 Da. As seen in Figure 25, no significant peaks corresponding to click products appeared after a 7 d incubation in the absence of RNA/DNA or in the presence of $r(CCUG)_8$ or tRNA. In the presence of a fully paired 24-mer DNA duplex or a paired DNA duplex containing $d(CTG)_{16}$-$d(GAC)_{16}$, small amounts of adducts were detected after 1 d.

In contrast, multiple intense peaks in the MALDI-MS were observed after just 1 d when the ligands were incubated with $d(CTG)_{16}$ or $r(CUG)_{16}$ and these peaks grew in intensity with additional time. Due to the observation of 6 peaks in both the DNA and RNA incubations, their corresponding structures were further characterized. Peaks at m/z 1027.4, 1084.4, 1215.6 were identified as the click product of **25**+**26**, **25**+**31**, and **25**+**6c**, respectively (Figure 26). Some click products could not be positively identified because the MALDI-MS *m*/*z* was consistent with more than one product. Thus, the peak at m/z 1041.4 could arise from three different click

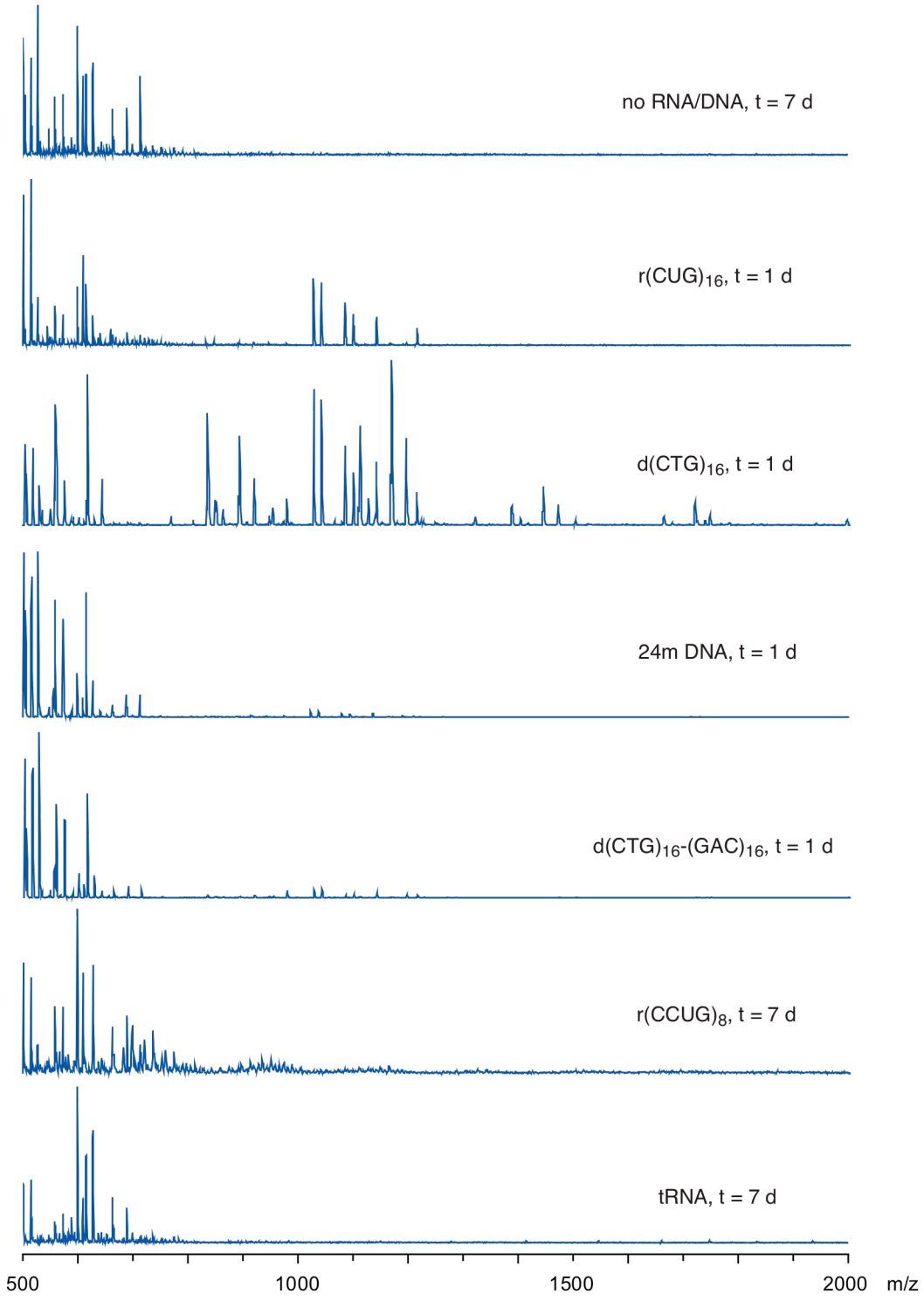


Figure 25. Representative MALDI-MS of templated click reactions of **6a**−**c**, **23**−**32**, and **35**, **38**, **45**−**59**. Analyzed samples were crystallized in 2,5-dihydroxybenzoic acid (DHB) matrix. MALDI-MS were obtained from the Mass Spectrometry Laboratory, UIUC. The selections were performed in triplicate. The spectra were formatted using Adobe Illustrator CS6.

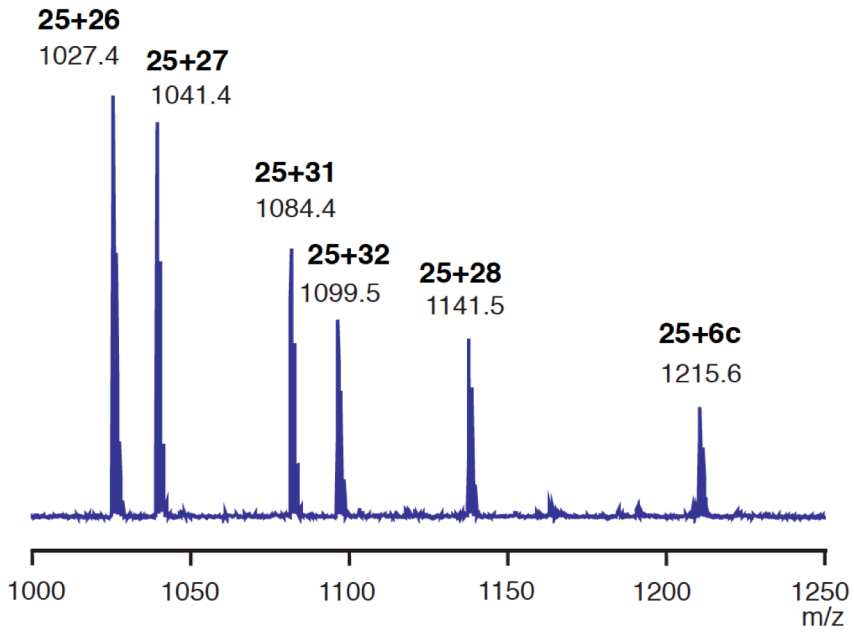


Figure 26. Matching the MALDI-MS peaks to the r(CUG)$_{16}$-templated click products.

couplings: **25**+**27**, **6b**+**35**, or **24**+**50**. Likewise, the peak at *m*/*z* 1099.5 could correspond to **25**+**32** or **6a**+**26** and the peak at *m*/*z* 1141.5 to **25**+**28** or **6b**+**26**.

Template selection assays under the same conditions as above with **6a**, **6b**, **6c**, **26**, **27**, **35**, **50**, but omitting **25** did not produce any click products after 7 d (Figure 27). This result suggests that the formed adducts at *m*/*z* 1041.4, 1099.5, 1141.5 are from the following pairings: **25**+**27**, **25**+**32**, and **25**+**28**, respectively (Figure 26). Not only does the template-assisted coupling appear quite selective, with only six major products formed, all six contain the same acridine-based alkyne partner **25**, presumably reflecting the higher reactivity of the activated alkyne. The selection also shows a considerable degree of selectivity with respect to the azide partner as five of the six participating azides (**26**–**28**, **31**, and **32**) are acridine-triaminotriazine conjugates. Indeed, only the compound at *m*/*z* 1215.6 was the product of acridine **25** and bisamidinium ligand **6c**.

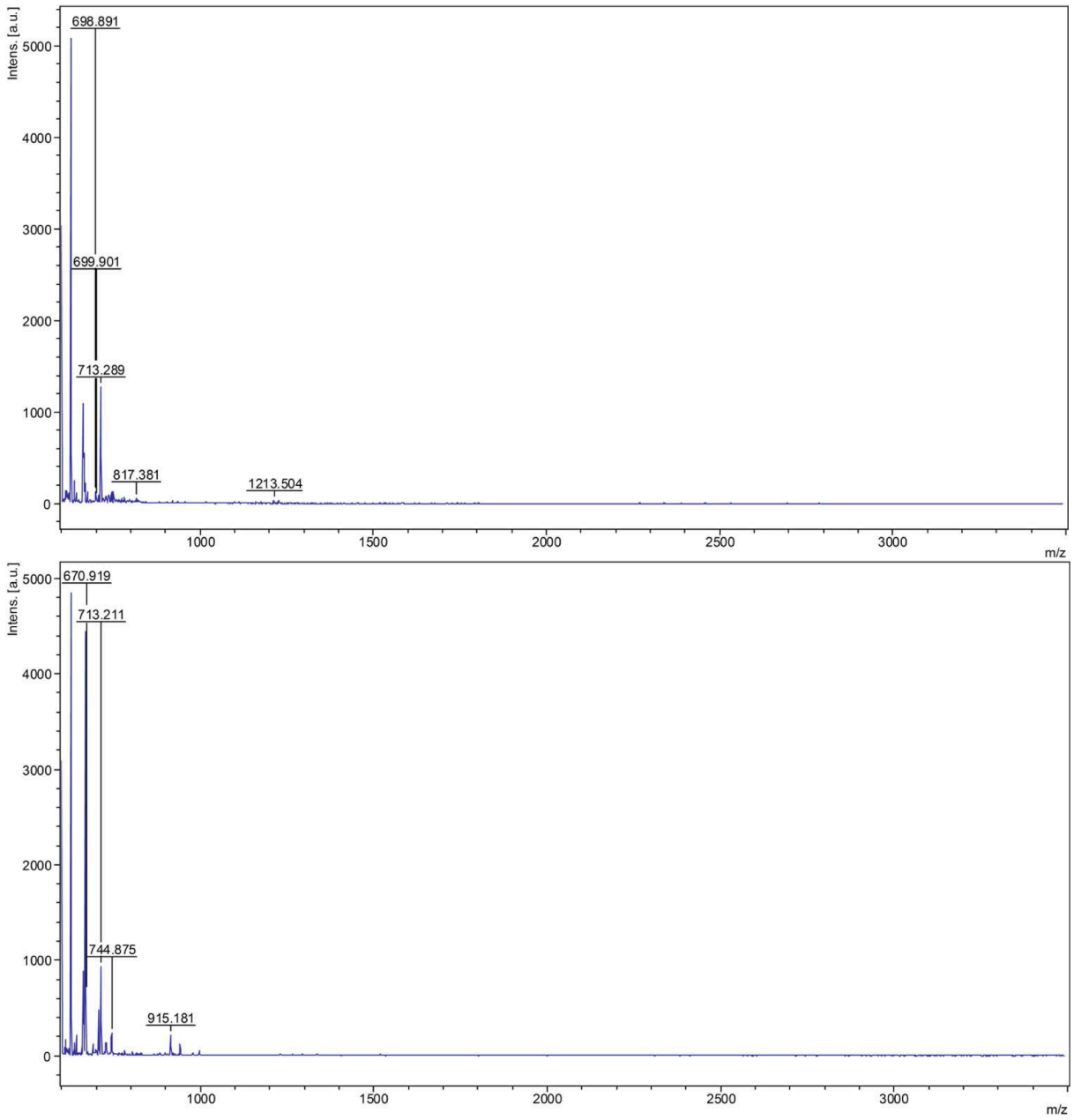


Figure 27. MALDI-MS of the mixture of **6a−c**, **26**, **27**, **35**, **50** in the presence of $d(CTG)_{16}$ (top) and $r(CUG)_{16}$ (bottom) after 7 d incubation at 37 °C, pH 7.0. MALDI-MS were obtained from the Mass Spectrometry Laboratory, UIUC.

It is important to note that the relative intensity of these peaks varies between runs, but remains the primary products of the *in situ* template click reaction as detected by MALDI-MS in independent selections. These results suggest that binding affinity and alkyne reactivity is the

main determinant of the selection process. The apparent lack of a requirement for precise positioning of the alkyne and azide also supports this conclusion. Thus, the increasingly longer chain connecting the azide to the acridine ring along the series **26** → **27** → **28** did not affect the ability of the click product to form. Liu and coworkers reported an extensive series of distance-independent, DNA-templated synthetic reactions, to which they attributed hybridization kinetics being rate determining in many instances.[83,84] Although the dataset here is not expansive enough to reach firm conclusions, these examples show that there does not appear to be a strict distance or geometric requirement for the template-assisted click reaction.

As discussed above, only known rCUG binding molecules participated in the templated click reaction. In addition, no products from the Michael addition of amine-containing compounds and activated alkynes were observed. Therefore, to allow other ligands to bind rCUG repeats, amine- and azide-containing ligands **23**, **26**, **28**−**32** were excluded in another selection experiment. On the other hand, the second selection included more general DNA/RNA binders, including electron-deficient alkynes **36**, **37**, **39**, azide **40**, and aminoglycosides **41**−**44**. After 2 days of incubation with DNA or RNA, reaction samples were mixed with either 2,5-dihydroxybenzoic acid (DHB), or α-cyano-4-hydroxycinnamic (CHCA), and analyzed using a positive mode. Using DHB matrix, distinct peaks were observed in different samples (Figure 28). A small peak corresponding to the click product of **25** and **27** was observed in the reactions with $r(CUG)_{16}$, 24mer-DNA, and $r(CCUG)_{12}$. Although the incubation with $d(CTG)_{16}$ showed an intense peak of **25**+**27** at 1041 in DHB matrix, this peak did not appear in CHCA matrix, suggesting the advantage of DHB over CHCA in detecting this kind of small molecules. Peaks from 800 to 1500 seen in $r(CCCCGG)_3$, HIV-1 FS RNA, and tRNA matched with those from the degradation of RNA using RNase A, and therefore not identified (Figure 28 and 29).

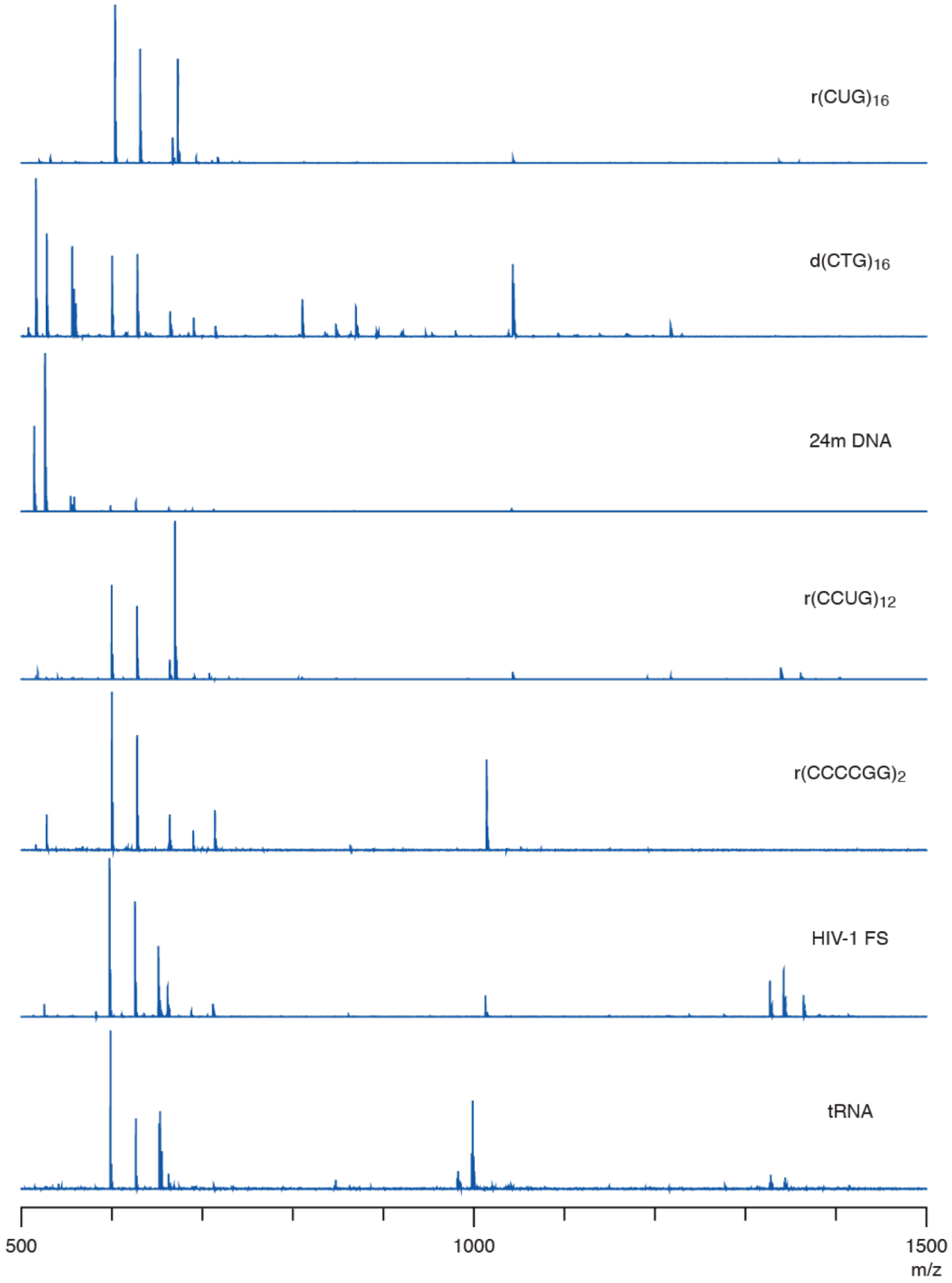


Figure 28. MALDI-MS of the selection experiment containing **6a**−**c**, **24**, **25**, **27**, and **35**−**59** after 2 d using DHB matrix. The mass spectra were formatted using Adobe Illustrator CS6.

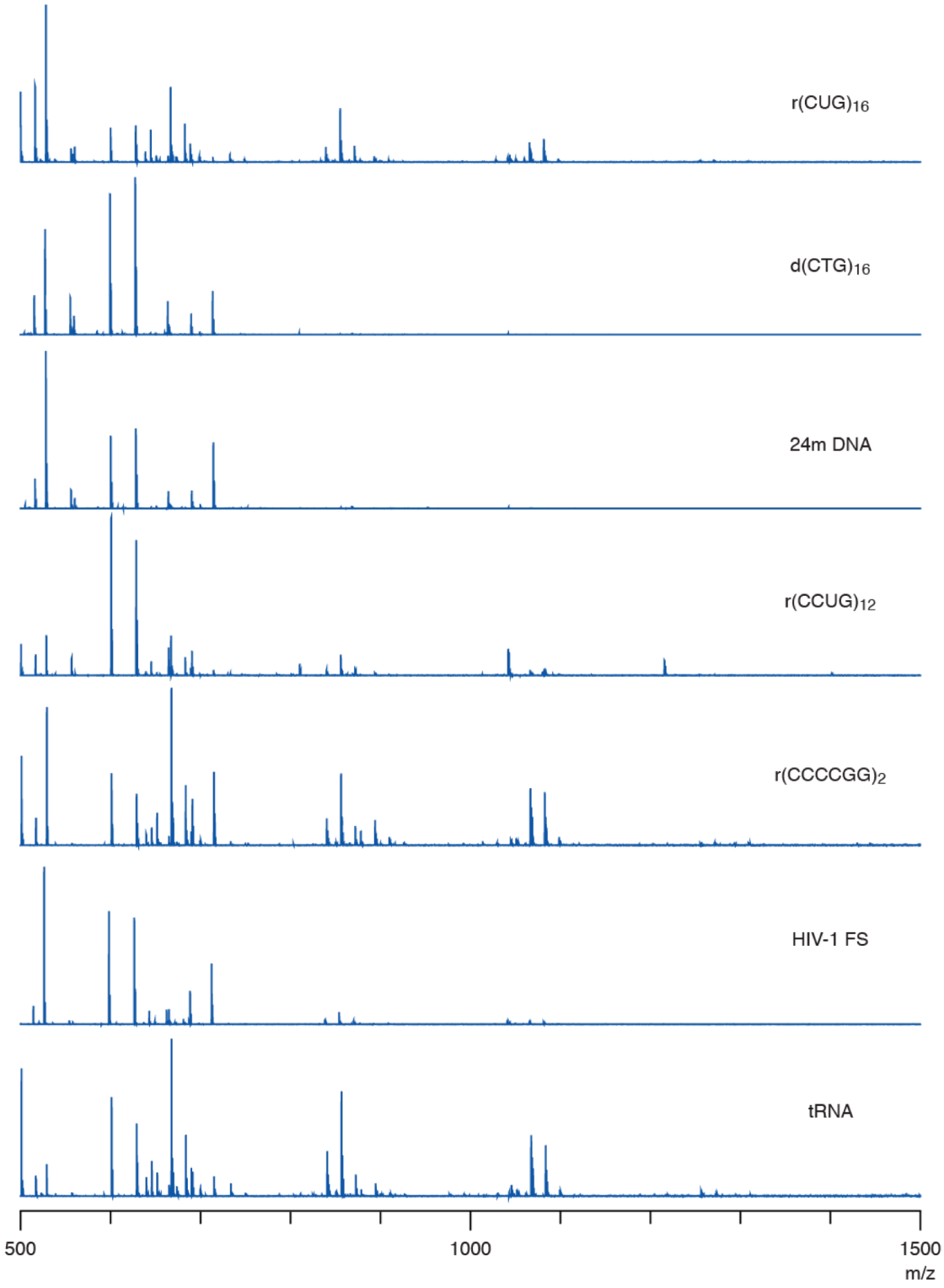


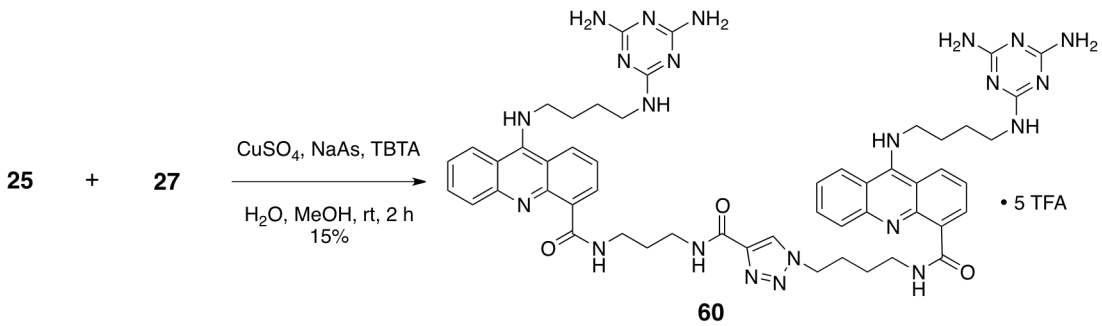
Figure 29. MALDI-MS of the selection experiment containing **6a**−**c**, **24**, **25**, **27**, and **35**−**59** after 2 d using CHCA matrix. The mass spectra were formatted using Adobe Illustrator CS6.

4.5 Analysis of Click Reaction of 25 and 27 Using HPLC

A more quantitative view of the RNA-templated click reaction was obtained using HPLC analysis, which is capable of monitoring the consumption of reactants, and the formation of product (Figure 30a). Alkyne **25** and azide **27** were chosen for this experiment because their click product was regularly one of the most predominant peaks observed across all of the MALDI-MS selection assays. To obtain detectable signals in HPLC, concentrations of both ligands and r(CUG)$_{16}$ were raised to 100 μM and 25 μM, whereas the pH and salt concentration held constant to those in the selection experiment. As seen in Figure 30a, compounds **25** and **27** were rapidly consumed, and a peak corresponding to the click product was observed after 1 d. In contrast, the size of the HPLC peak corresponding to alkyne **25** and azide **27** remained minimally changed after 5 d in the absence of r(CUG)$_{16}$. By using an internal standard in the HPLC analysis, the consumption of alkyne **25** was quantitatively determined (Figure 30b). In the presence of r(CUG)$_{16}$, ca. 40% of **25** remained after 1 d, and both reactants were almost consumed after 4 d. Without the RNA template, only 20% of **25** reacted with the azide after 5 d.

The [3+2] cycloaddition of **25** and **27** can produce 1,4- and 1,5-disubstituted 1,2,3-triazole products. The 1,4-isomer **60**, prepared using a system containing copper (II), sodium

Scheme 7. Synthesis of dimer **60**.



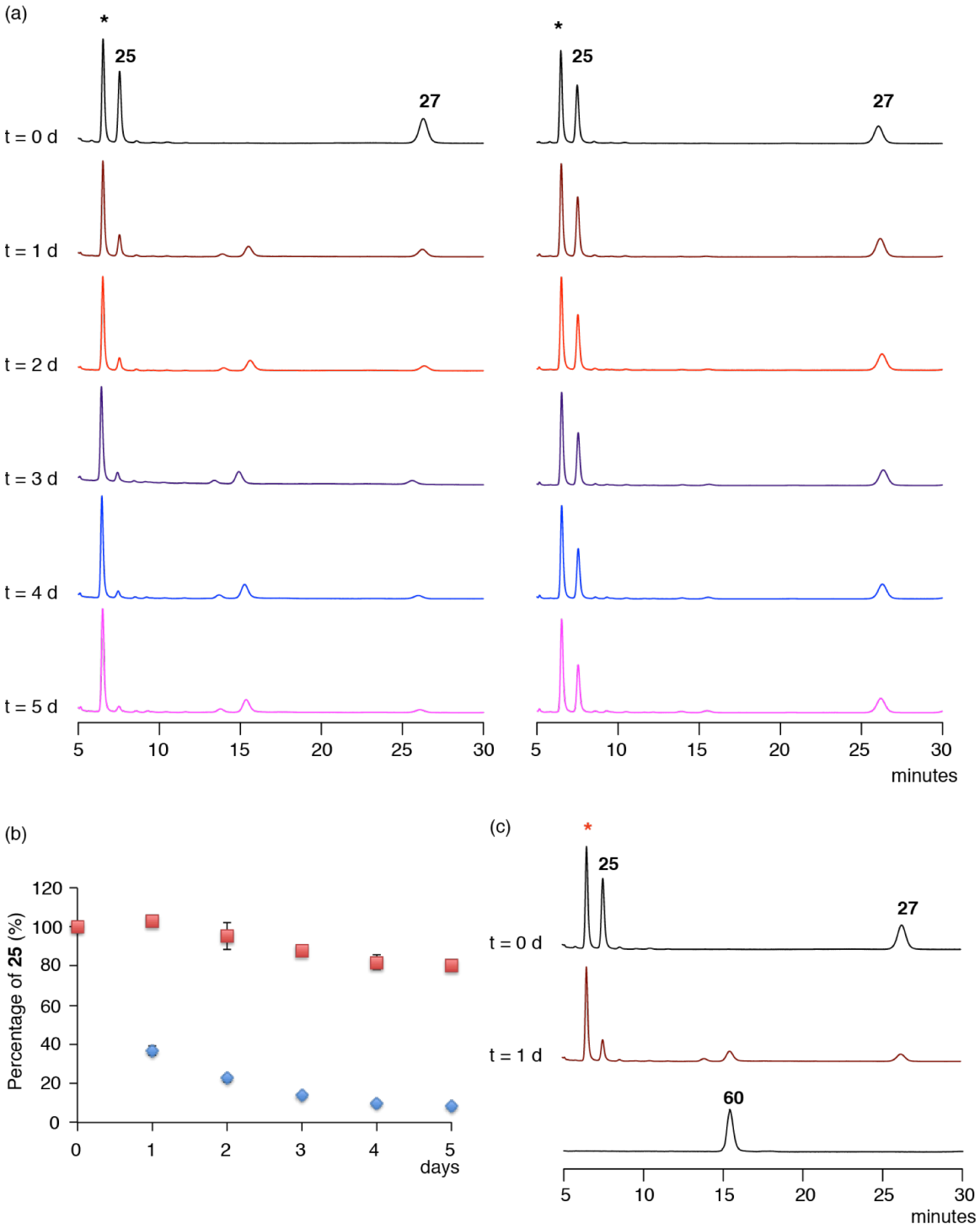


Figure 30. (a) Monitoring the reaction of **25** and **27** in the presence (left) and absence (right) of r(CUG)$_{16}$. Internal standard (compound **28**) was marked as an asterisk. (b) The consumption of **25** in the presence (blue diamond) and absence (red square) of r(CUG)$_{16}$. The error bars were obtained from two independent experiments. (c) Determination of the click product isomer by HPLC.

ascorbate and TBTA (Scheme 7), served as the authentic sample in the HPLC analysis. It was found that the RNA-templated click reaction mainly affords 1,4-isomer (Figure 30c). In fact, the thermal cycloaddition of azide and electron deficient acetylene favors the formation of 1,4-isomer.[66]

4.6 Biological Activities of the Click Product

4.6.1 Inhibition of the Transcription of dCTG Repeats *In Vitro*

The transcription inhibition of dCTG repeats has emerged as a potential therapeutic strategy to treat DM1 in recent years. Pearson *et al.* discovered that 67−70% of the DNA containing 50 dCTG-dCAG repeats exist as slipped-strand conformations, in which dCTG and

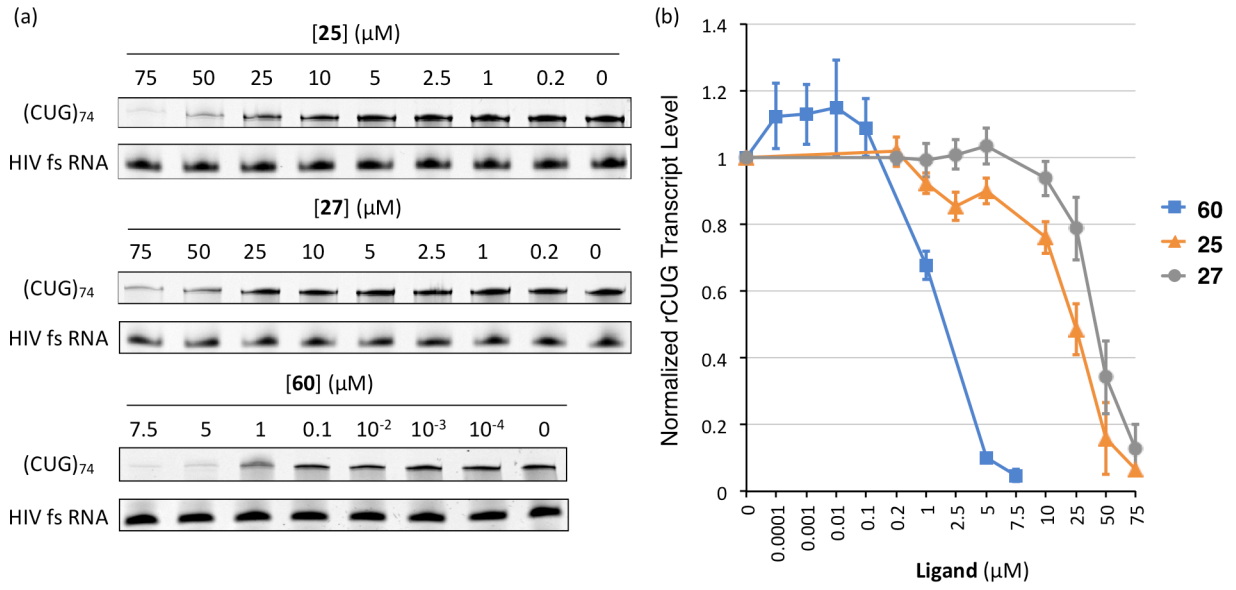


Figure 31. *In vitro* transcription inhibition of d(CTG)$_{74}$-d(GAC)$_{74}$ by monomers **25**, **27**, and dimer **60** at 37 °C, pH 8.3. (a) Representative transcription gels. The reaction mixture was separated on a 8% denaturing PAGE gel. (b) Plot of the normalized level of r(CUG)$_{74}$ transcript versus concentration of ligands. Error bars represent the standard error of mean from three independent experiments. The intensity of rCUG product was normalized to HIV-1 FS RNA as an internal standard. To access the transcription inhibition potency, the ratio of r(CUG)$_{74}$/HIV-1 FS RNA in the ligand-treated sample was compared with that from the untreated one. (This experiment was conducted by JuYeon Lee)

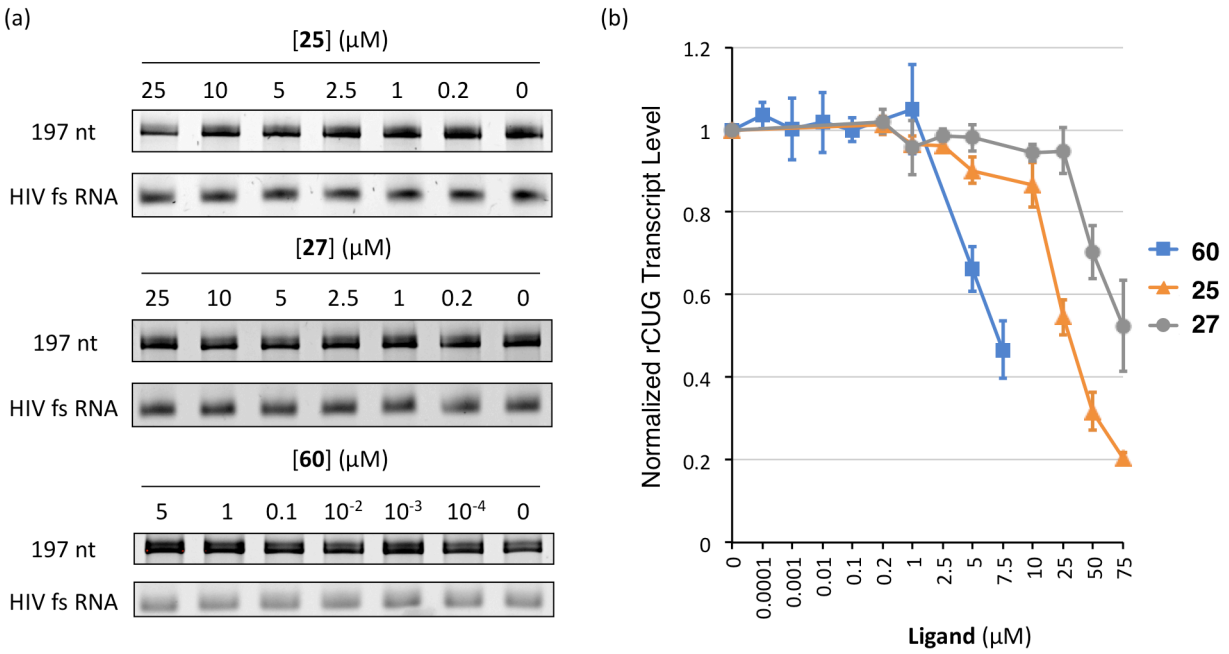


Figure 32. *In vitro* transcription inhibition of the control DNA by monomers **25**, **27**, and dimer **60**. (a) Representative transcription gels. (b) Plot of the normalized level of r(CUG)$_{74}$ transcript versus concentration of ligands. Error bars represent the standard error of mean from three independent experiments. (This experiment was conducted by JuYeon Lee)

dCAG repeats form hairpin and unpaired loop structures, respectively.[79,85] Recently, we[49,86] and Berglund[36,37] reported that targeting DNA with a dCTG binding ligand could inhibit the transcription to form rCUG repeats both *in vitro* and in DM1 model cells.

To determine whether compound **60**, that was discovered from the dCTG template-mediated selection, could function in this manner, the linearized T7-promoter containing (CTG)$_{74}$ plasmids were incubated with **60**, followed by the *in vitro* transcription reaction with T7 polymerase.[13,49] For comparison, the transcription inhibition by monomers **25** and **27** was also performed under similar conditions. The rCUG transcript and HIV-1 FS RNA (internal standard) were stained with ethidium bromide (EtBr). It was found that incubation of the dimeric ligand **60** at 5−7.5 μM with the d(CTG)$_{74}$ template resulted in more than 90% suppression of the

production of CUG-RNA, whereas treatment with monomeric ligands **25** and **27** in concentrations up to 10 μM only slightly reduced the r(CUG)$_{74}$ levels (Figure 31). Although monomers exhibited a nearly complete transcription inhibition at 75 μM, this study indicates an improved potency of dimer **60** at a low micro-molar concentration. However, it was found that dimer **60** also inhibited 34−50% of transcription of the control DNA duplex at 5−7.5 μM (Figure 32). In fact, a small amount of dimer **60** was also observed in the selection experiment involving fully paired DNA sequences.

4.6.2 Inhibition of the MBNL1-rCUG^{exp} *In Vitro*

Because dimer **60** was formed in both dCTG and rCUG templated click reaction, the ability of **60** as well as alkyne **25** to inhibit the MBNL1 sequestration by rCUG^{exp} was evaluated using EMSA. In this experiment, the truncated MBNL1 and radiolabeled-r(CUG)$_{12}$ formed a tight complex with a K_D value of 9 ± 1 nM. An IC$_{50}$ value of 26000 ± 2000 nM and a calculated K_I of 3343 ± 452 nM were determined for monomer **25**, whereas dimer **60** was ca. 144-fold more potent, with IC$_{50}$ and K_I values of 180 ± 5 nM and 23 ± 3 nM, respectively (Figure 33).

The improvement in inhibition potency of **60** is likely a result of its enhanced binding to both d(CTG)$_{16}$ and r(CUG)$_{16}$. Both compounds **25** and **60** are fluorescent and have an observable fluorescence intensity decrease upon binding to DNA/RNA (Figure 34). Thus, the apparent binding constants for complexation to d(CTG)$_{16}$, and r(CUG)$_{16}$ were determined using a reverse fluorescence titration (Figure 35). It was found that the binding affinities of alkyne **25** to d(CTG)$_{16}$ and r(CUG)$_{16}$ were 4810 ± 314 nM and 350 ± 27 nM, respectively, which are similar to K_D values of **1** determined by ITC.[40,41] In comparison to **25**, dimer **60** exhibited ca. 21- and

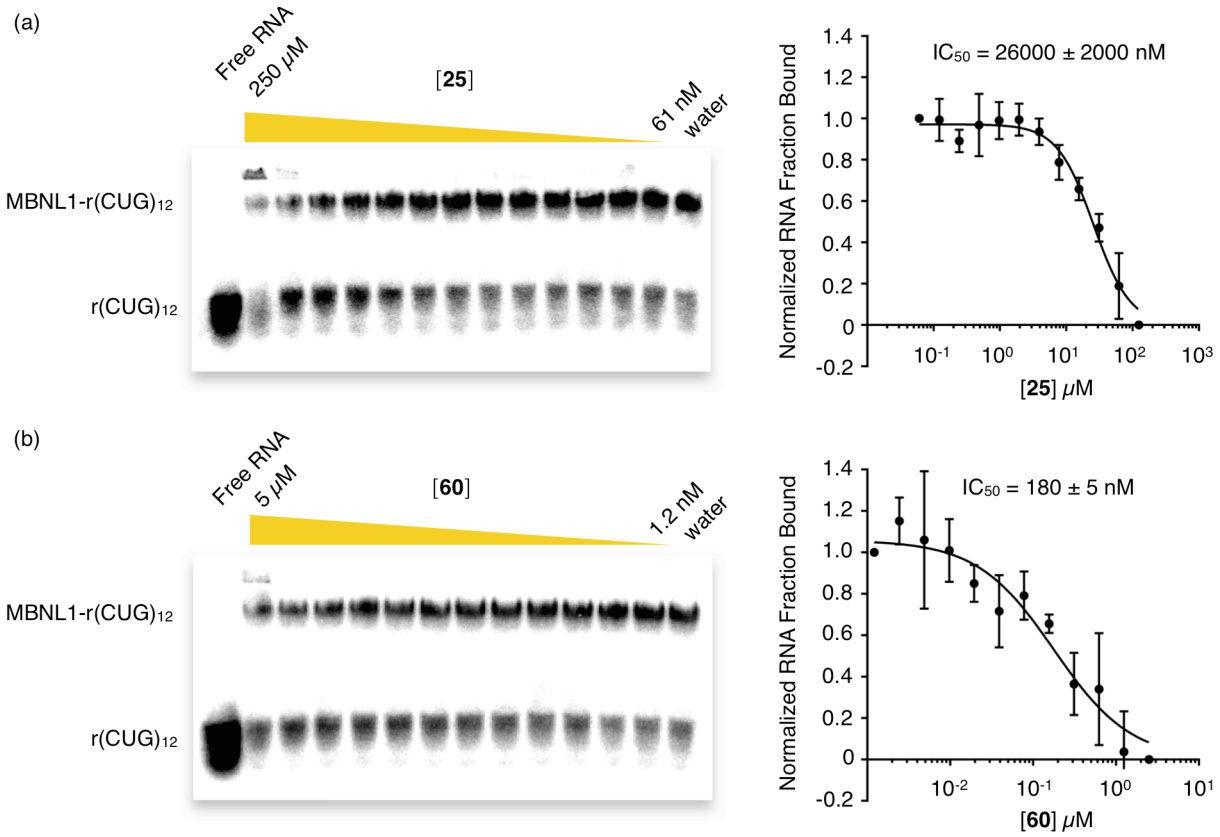


Figure 33. Inhibition of MBNL1-r(CUG)$_{12}$ by (a) monomer **25** and (b) dimer **60** by EMSA. Left: Representative EMSA gel. The compound concentration was serially diluted by the factor of 2. Right: Determination of the IC_{50} value. The points at 250 μM of **25** and 5 μM of **60** were not included in the plot because of aggregation in some gels. Standard error of the mean was obtained from three independent experiments. All data were plotted and fitted using Prism.

15-fold stronger affinities to d(CTG)$_{16}$ (K_D = 23 ± 1 nM) and r(CUG)$_{16}$ (K_D = 230 ± 10 nM). Apart from the increased binding affinities, the improved inhibition potency of **60** may in part result from the steric hindrance of the triazole linker.

4.7 Conclusions and Outlook

Herein, we have described an accelerated strategy for discovering agents that target the DNA and RNA that cause myotonic dystrophy type 1 (DM1). The approach combines elements of FBDD and TGS with a positive cross screen to discover dual-targeting lead agents for DM1.

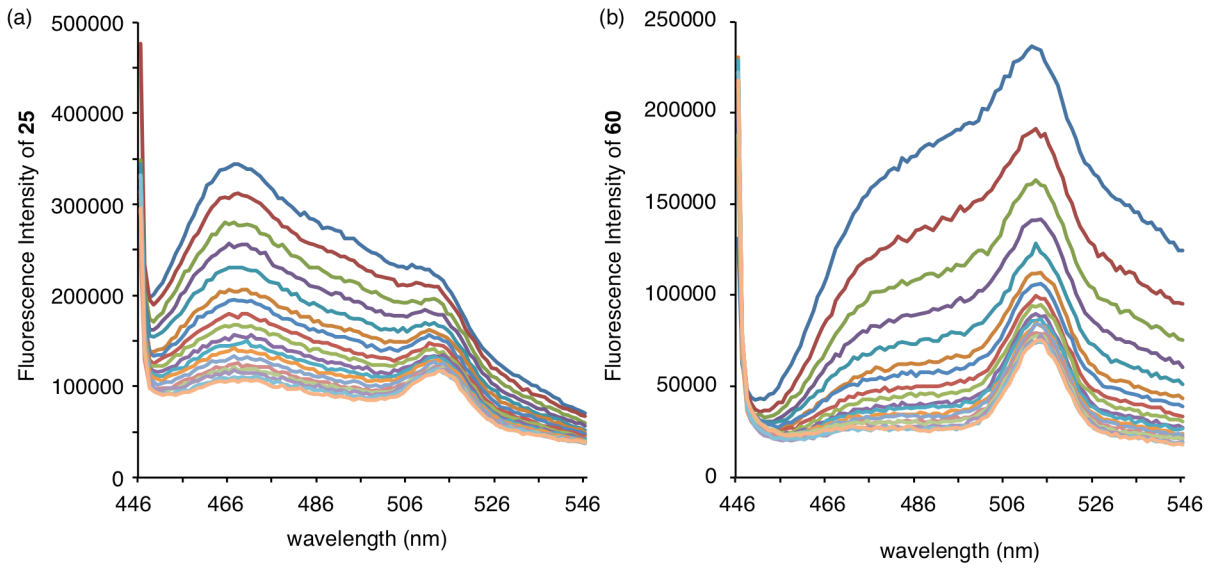


Figure 34. Fluorescence spectra of **25** and **60** as the r(CUG)$_{16}$ solution was titrated to the ligand solutions.

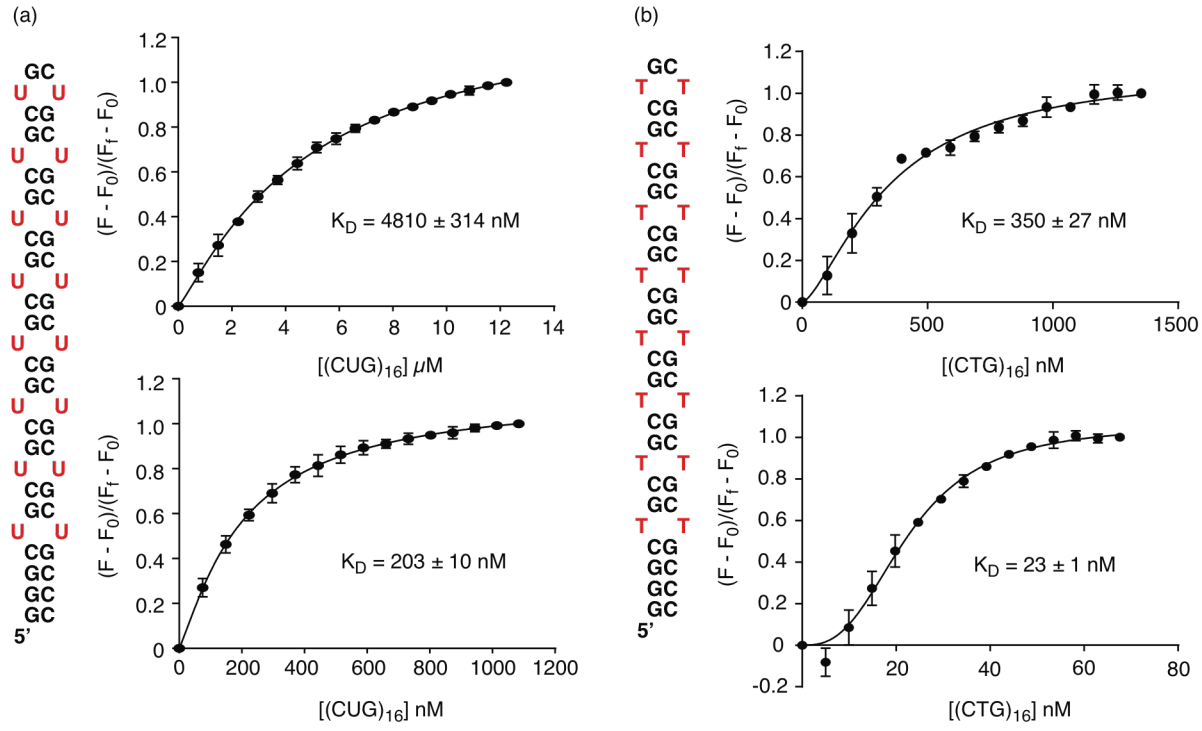


Figure 35. Determination of binding affinities of **25** (top) and **60** (bottom) to r(CUG)$_{16}$ and d(CTG)$_{16}$.

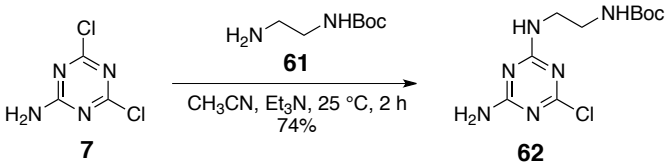
Template-assisted selections produced only a small number of products that formed only

in the presence of the target DNA or RNA and not with other oligonucleotides, such as tRNA and a generic dsDNA. The selection appeared to favor tighter binding ligands and those with more reactive clickable functionality (i.e., activated alkyne) and was less discriminating with respect to the length of the linker between the ligand and the click-functional group.

Beyond the general approach described here for the discovery of more potent *in vitro* inhibitors and multi-targeting agents, this same general protocol could be applied to DM1 cells. A cell-based approach adds variables of cellular and nuclear permeability, critical aspect for drug development, to the selection. 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, which then facilitates the synthesis of the active agent.[67,87] Such an approach will require a new method of product analysis and a faster biocompatible click reaction. Efforts to accomplish these goals are currently underway. The success of the templated-click selection using the toxic RNA in DM1 suggests that it may have applications to other diseases caused by unstable trinucleotide repeats such as Huntington's disease, fragile X syndrome, and Friederich's ataxia.

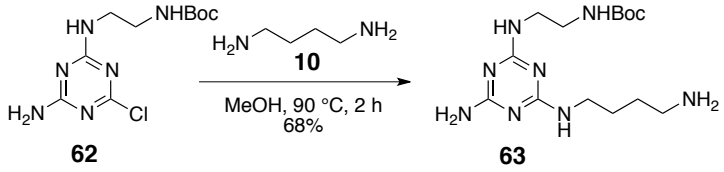
4.8 Materials and Methods

4.8.1 Synthesis

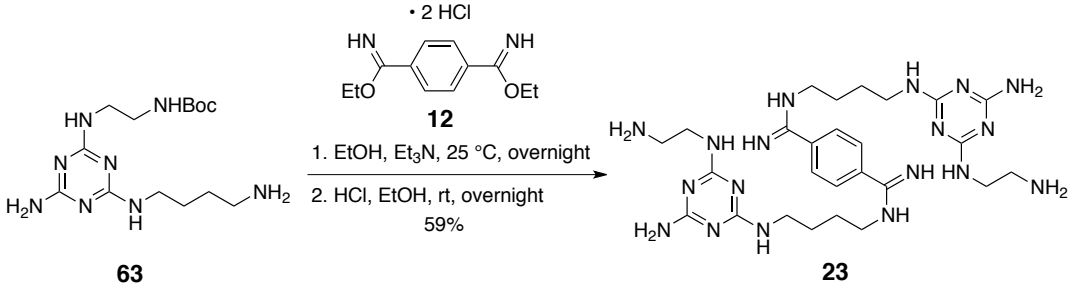


Compound **62**: Using the similar procedure as described in Chapter 2, 2.9 g (17.6 mmol) of **7**, 3 mL (21.5 mmol) of Et_3N, and 3.1 g (19.3 mmol) of *N*-Boc-ethylenediamine **61** afforded

3.75 g (74%) of **62** as a white solid. ^{1}H NMR (DMSO-d_6): δ 7.65−7.50 (m, 1H), 7.31−7.11 (m, 2H), 6.85−6.81 (m, 1H), 3.25−3.20 (m, 2H), 3.07−3.01 (m, 2H), 1.36 (s, 9H). Note: *N*-Boc-ethylenediamine **61** is available from Sigma-Aldrich or can be easily prepared using reported procedures.

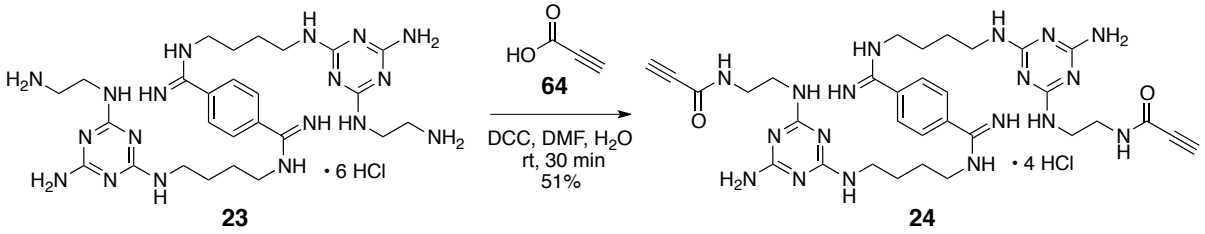


Compound **63**: Using the similar procedure as described above, 3.13 g (10.8 mmol) of **62**, 5.5 mL (54.7 mmol) of 1,4-diaminobutane **10** afforded 2.51 g (68%) of **63** as a white solid. ^{1}H NMR (DMSO-d_6): δ 6.8 (s, 1H), 6.53−5.88 (m, 4H), 3.24−3.02 (m, 8H), 1.88 (s, 2H), 1.75 (s, 2H), 1.37 (s, 9H).



Compound **23**: To a 200 mL 24/40 one neck round bottom flask was added 1.0 g (3.44 mmol) of **12**. The solid was suspended in 30 mL of anhydrous EtOH. To the white suspension was added 1.2 mL (8.6 mmol) of anhydrous Et_3N. To the obtaining clear solution was added 2.9 g (8.2 mmol) of **63** in 5 mL of anhydrous EtOH. The mixture was stirred at room temperature overnight. Ethanol was removed using a rotary evaporator. The crude was dissolved in 50 mL of ethanolic solution of HCl 2 M. The reaction was stirred at room temperature overnight. The

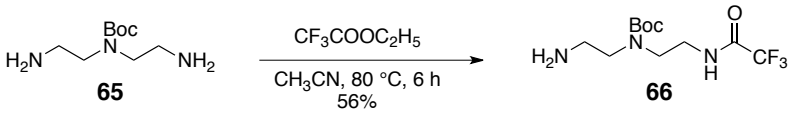
solution was concentrated using a rotary evaporator. The crude was purified by Sephadex CM-25 column using aqueous solution of NH_4HCO_3 (gradient from 0.3 M to 1 M). Fractions containing products were combined, concentrated using a rotary evaporator to give 1.23 g (59%) of **23** as a white solid. The solid was dissolved in a methanolic solution of HCl, then concentrated to give a white HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.22 (s, 2H), 9.77 (s, 2H), 9.38 (s, 2H), 8.15 (s, 4H), 7.97(s, 4H), 6.84−6.65 (m, 4H), 6.31−6.09 (m, 4H), 3.48−3.46 (m, 4H), 3.38−3.34 (m, 4H), 3.24 (s, 4H), 2.91 (s, 4H), 1.69−1.57 (m, 8H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 609.4075; found 609.4071.



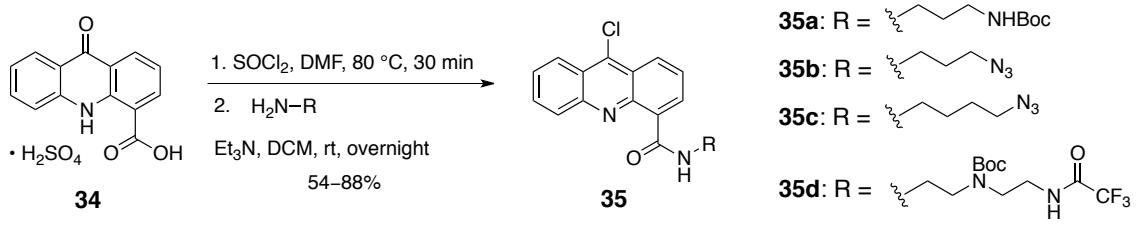
23 **24**

Compound **24**: To a 15 mL 14/20 round bottom flask was added 207 mg (0.25 mmol) of **23**. The solid was dissolved in 0.5 mL of DI water and 3 mL of DMF. The mixture was stirred at room temperature for 5 min. To a 5 mL vial was added 516 mg (2.50 mmol) of DCC, and 2 mL of CH_2Cl_2. To the resulting solution was added slowly 181 mg (2.58 mmol) of propiolic acid **64**. The resulting brown suspension was added to the reaction flask at once. The reaction was stirred at room temperature for 30 min. The mixture was diluted with 10 mL of water and 8 mL of CH_2Cl_2, and mixed thoroughly. The aqueous layer was isolated, extracted with 10 mL of CH_2Cl_2 one more time. The aqueous layer was purified by C18 column chromatography using a CombiFlash with 0.1% of HCl in MeOH/H_2O (1/9, v/v) at a flow rate of 5 mL/min. The fractions containing product was combined, and concentrated using a rotary evaporator to give 132 mg (51%) of **24** as a tetra-HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.27−10.20 (m, 2H), 9.80 (s, 2H),

9.44−9.41 (m, 2H), 8.93−8.83 (m, 2H), 8.50−7.91 (m, 12H), 4.18 (s, 2llH), 3.49−3.46 (m, 4H), 3.39−3.33 (m, 8H), 3.27−3.25 (m, 4H), 1.71−1.63 (m, 8H). ^{13}C NMR (DMSO-d_6): δ 161.74, 155.92, 151.99, 151.95, 132.93, 128.69, 128.66, 78.25, 75.99, 75.92, 48.61, 42.54, 42.47, 39.52, 38.32, 26.00, 24.65. HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 713.3973; found 713.3950. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 4.7 min (86%).



Compound **66**: To a 300 mL round bottom flask was added 7.35 g (36.2 mmol) of **65** and 100 mL of CH_3CN. To the reaction flask was added slowly 5 ml (42.0 mmol) of ethyl trifluoroacetate over 5 h at 80 °C. The mixture was stirred at 80 °C for an additional 1 h. The solvent was removed using a rotary evaporator. The crude was purified by silica gel column chromatography with a gradient mixture of CH_2Cl_2/MeOH from 19/1 to 9/1 (v/v). The fractions containing product were combined and concentrated to give 6.1 g (56%) of **66** as a pale yellow oil. ^{1}H NMR (DMSO-d_6): δ 3.32−3.28 (m, 6H), 3.14−3.10 (m, 2H), 2.65−2.63 (m, 2H), 1.38 (s, 9H). ESI-MS (*m/z*) calculated for $[M + H]^+$: 300.2; found 300.1.



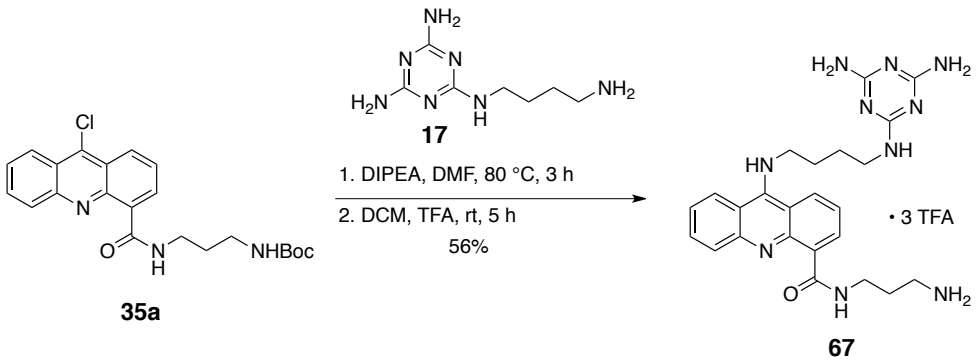
Compound **35a**: To a 100 mL 14/20 one neck oven-dried round bottom flask was added 3.1 g (9.2 mmol) of acridone **34**. The solid was dissolved in 20 mL of $SOCl_2$ and 1 mL of anhydrous DMF. The mixture was stirred at 80 °C for 30 min. Thionyl chloride was removed by

distillation. The crude was dissolved in 30 mL of anhydrous CH_2Cl_2, the solvent was removed by distillation. This step was repeated two times. The yellow solid was dried under high vacuum for at least 1 hour. The flask was placed in an ice bath. The yellow solid was dissolved by 50 mL of anhydrous CH_2Cl_2. To the resulting solution was added anhydrous Et_3N until pH 11. To the reaction mixture was added 2.0 g (11.5 mmol) of *N*-Boc-1,3-propanediamine. The temperature was raised slowly to 25 °C, and the mixture was stirred overnight. The reaction was concentrated using a rotary evaporator. The crude was purified by silica gel column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 100:0 to 19:1 (v/v). Fractions containing product were combined, concentrated using a rotary evaporator to give 3.2 g (84%) of compound **35a** as a yellow solid. ^{1}H NMR ($CDCl_3$): δ 11.75 (s, 1H), 9.00 (dd, *J* = 7.5, 1.5, 1H), 8.65−8.63 (m, 1H), 8.47 (d, *J* = 8.5, 1H), 8.32 (d, *J* = 8.7, 2H), 7.91−7.87 (m, 1H), 7.78−7.74 (m, 1H), 7.73−7.70 (m, 1H), 5.19 (s, 1H), 3.75 (q, *J* = 6.3, 2H), 3.34 (q, *J* = 6.3, 2H), 1.95 (quint, *J* = 6.4, 2H), 1.46 (s, 9H).

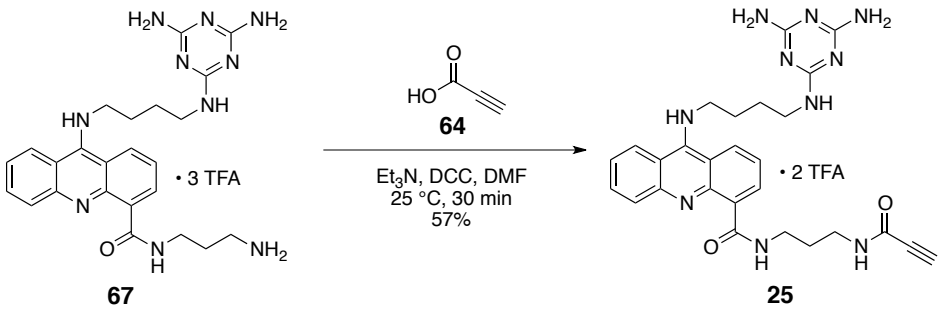
Compound **35b**: Using the similar procedure as described above, 1.0 g (3.0 mmol) of acridone **34** and 450 mg (4.5 mmol) of 3-azido-propylamine gave 898 mg (88%) of **35b** as a yellow solid. ^{1}H NMR (CD_3OD-d_4): δ 8.52 (dd, *J* = 7.0, 1.4, 1H), 8.18 (dd, *J* = 8.6, 1.5, 1H), 8.06−8.04 (m, 1H), 7.81−7.78 (m, 1H), 7.75−7.71 (m, 1H), 7.55−7.47 (m, 2H), 3.63 (t, *J* = 6.8, 2H), 3.56 (t, *J* = 6.6, 2H), 2.02 (quint, *J* = 6.7, 2H).

Compound **35c**: Using the similar procedure as described above, 5.15 g (15.3 mmol) of acridone **34** and 2.65 g (23.2 mmol) of 4-azido-butylamine gave 3.75 g (69%) of **35c** as a yellow solid. ^{1}H NMR (CD_3OD-d_4): δ 8.47 (dd, *J* = 7.1, 1.5, 1H), 8.07 (dd, *J* = 8.6, 1.5, 1H), 7.95 (dt, *J* = 8.7, 1.0, 1H), 7.69−7.64 (m, 2H), 7.50−7.41 (m, 2H), 3.56−3.45 (m, 4H), 1.85−1.77 (m, 4H).

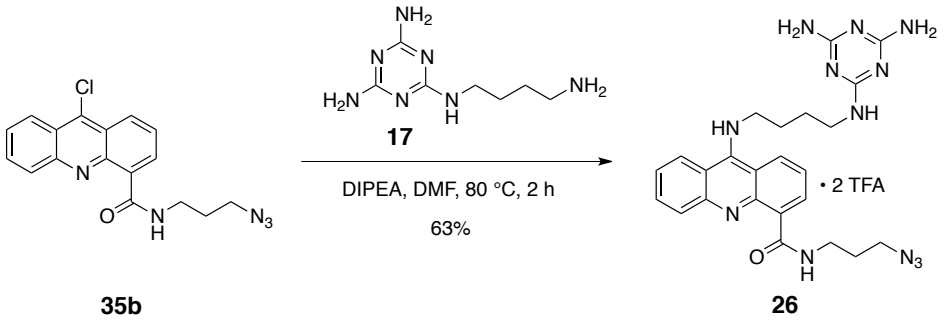
Compound **35d**: Using the similar procedure as described above, 3.0 g (8.9 mmol) of acridone **34** and 3.46 g (10.64 mmol) of **66** gave 2.57 g (54%) of **35d** as a yellow solid.



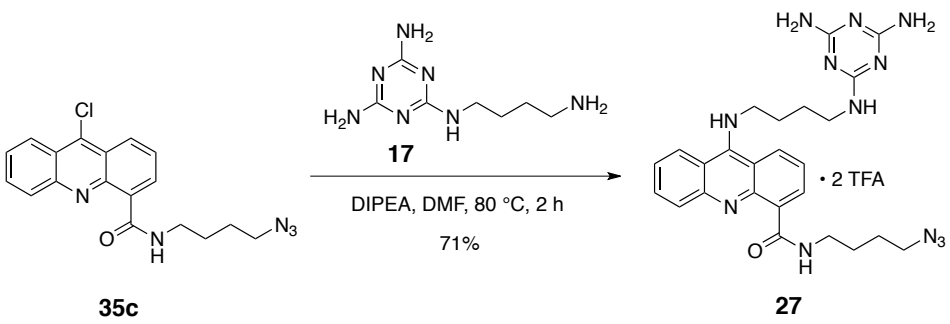
Compound **67**: To a 100 mL round bottom flask was added 1.53 g (3.7 mmol) of **35a**, 0.95 g (4.8 mmol) of **17**, 20 mL of DMF, and 0.9 mL (5.2 mmol) of DIPEA. The mixture was stirred at 80 °C in an oil bath for 3 h. The reaction was concentrated using a rotary evaporator. The crude was purified by basic Al_2O_3 column chromatography with a gradient mixture of CH_2Cl_2/MeOH from 100/0 to 90/10 (v/v). Fractions containing product were combined and concentrated using a rotary evaporator. The yellow solid was dissolved in a mixture of 50 mL of TFA and 50 mL of CH_2Cl_2. The mixture was stirred at room temperature for 5 h. The reaction was concentrated using a rotary evaporator, the crude was dried under high vacuum to give 1.68 g (56%) of **67** as a yellow solid. ^{1}H NMR (CD_3OD-d_4): δ 8.70 (d, J = 8.6, 1H), 8.52 (d, J = 8.7, 1H), 8.46 (d, J = 7.3, 1H), 7.99 (m, 1H), 7.90 (d, J = 8.4, 1H), 7.64−7.59 (m, 2H), 4.21 (t, J = 7.3, 2H), 3.63 (t, J = 6.7, 2H), 3.45 (t, J = 6.6, 2H), 3.09 (t, J = 7.5, 2H), 2.06 (m, 4H), 1.78 (m, 2H). ESI-MS (*m/z*) calculated for $[M + H]^+$: 475.3; found 475.1.



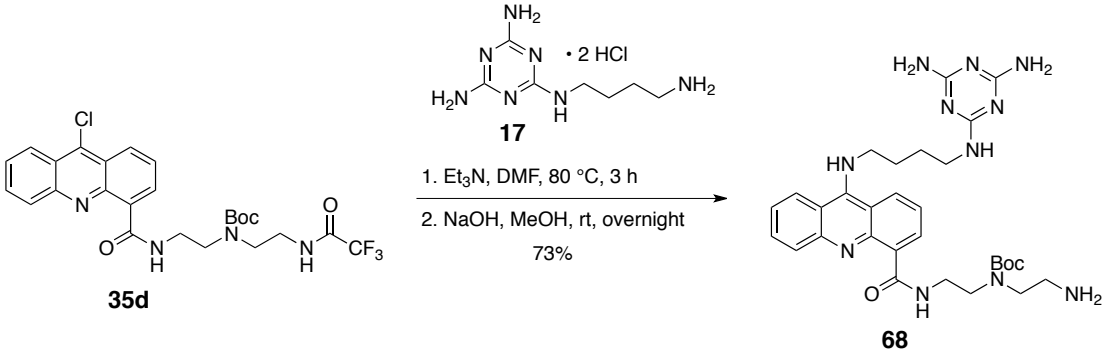
Compound **25**: To a 50 mL centrifuge tube containing 30 mL of anhydrous DMF was added 1.50 g (7.3 mmol) of DCC. To the resulting solution was added 0.47 g (6.7 mmol) of propiolic acid **64**. The mixture was vortexed until the suspension was formed. The solid was filtered, the filtrate was transferred to a 200 mL round bottom flask. To the separate 50 mL centrifuge tube was added 1.05 g (1.3 mmol) of **67**, 20 mL of anhydrous DMF, and 1 mL (7.2 mmol) of anhydrous Et_3N. The resulting suspension was added to the reaction flask slowly. The reaction was stirred at room temperature for 30 min. The mixture was concentrated using a rotary evaporator. The crude was purified by C18 column chromatography using a gradient mixture of H_2O/MeOH from 100/0 to 50/50 (v/v) with 0.1% TFA. Fractions containing product were combined and concentrated using a rotary evaporator to give 0.55 g (57%) of **25** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.64 (dd, J = 8.7, 2.3, 1H), 8.49 (dd, J = 8.8, 2.7, 1H), 8.39 (dd, J = 7.4, 2.3, 1H), 7.99−7.95 (m, 1H), 7.90 (dd, J = 8.6, 2.3, 1H), 7.61−7.56 (m, 2H), 4.20 (dt, J = 7.3, 2.7, 2H), 3.63 (d, J = 3.2, 1H), 3.56 (dt, J = 6.8, 2.7, 2H), 3.44 (dt, J = 6.7, 2.9, 2H), 3.38 (dt, J = 6.9, 2.8, 2H), 2.05 (quint, J = 8.0, 2H), 1.92 (dquint, J = 6.8, 2.6, 2H), 1.77 (quint, J = 7.2, 2H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 527.2631; found 527.2635. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 7.5 min (100%).



Compound **26**: To a 100 mL round bottom flask was added 201 mg (0.59 mmol) of **35b**, 142 mg (0.71 mmol) of **17**, 10 mL of DMF, and 0.15 mL (0.86 mmol) of DIPEA. The mixture was stirred at 80 °C in an oil bath for 2 h. The reaction mixture was concentrated using a rotary evaporator. The crude was purified by basic Al_2O_3 column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 100/0 to 90/10 (v/v). Fractions containing product were combined and concentrated using a rotary evaporator to give a yellow solid. The solid was re-purified by C18 column chromatography using a gradient mixture of H_2O/MeOH from 90/10 to 80/20 (v/v) with 0.1% TFA at a flow rate of 5 mL/min. Fractions containing product was combined and concentrated using a rotary evaporator to give 271 mg (63%) of compound **26** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.63 (s, 1H), 8.46 (dd, J = 8.6, 1.5, 1H), 8.31 (d, J = 8.7, 1H), 7.93−7.91 (m, 1H), 7.72 (t, J = 7.5, 1H), 7.41−7.37 (m, 2H), 3.92 (t, J = 7.1, 2H), 3.69 (t, J = 6.7, 2H), 3.55 (t, J = 6.6, 2H), 3.29 (t, J = 6.9, 2H), 2.04 (quint, J = 6.7, 2H), 1.85 (quint, J = 7.6, 2H), 1.62 (quint, J = 7.4, 2H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 501.2587; found 501.2593. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 17.8 min (100%).

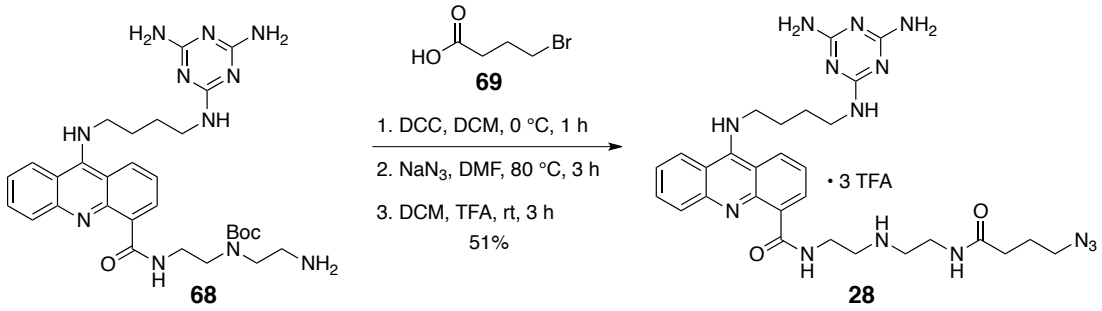


Compound **27**: Using the similar procedure as described above, 3.75 g (10.6 mmol) of **35c** and 2.15 g (10.9 mmol) of **17** gave 5.57 g (71%) of **27** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.65 (dd, J = 8.6, 1.2, 1H), 8.49 (d, J = 8.6, 1H), 8.41 (dd, J = 7.4, 1.2, 1H), 7.99−7.96 (m, 1H), 7.87 (dd, J = 8.6, 1.1, 1H), 7.61−7.56 (m, 2H), 4.20 (t, J = 7.2, 2H), 3.54 (t, J = 6.9, 2H), 3.44 (t, J = 6.6, 2H), 3.40 (t, J = 6.7, 2H), 2.08 −2.02 (m, 2H), 1.82−1.69 (m, 6H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 515.2744; found 515.2741. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 26.4 min (100%).



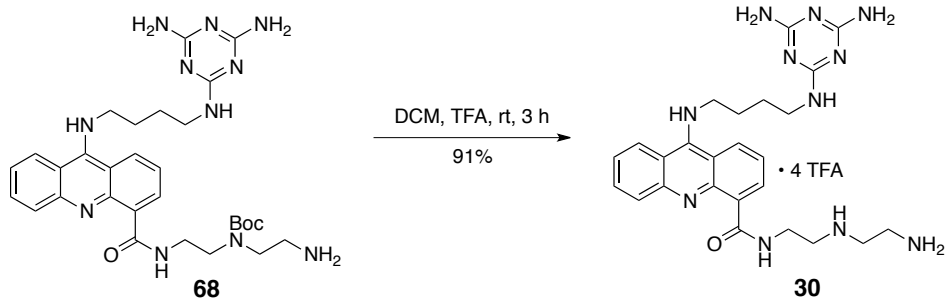
Compound **68**: To a 100 mL round bottom flask was added 2.57 g (4.8 mmol) of **35d**, 1.51 g (5.6 mmol) of N^2-(4-aminobutyl)-1,3,5-triazine-2,4,6-triamine hydrochloride salt **17**, 50 mL of DMF, and 3.3 mL (23.7 mmol) of Et_3N. The mixture was stirred at 80 °C in an oil bath for 3 h. The reaction mixture was concentrated using a rotary evaporator. The crude was purified by silica gel column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 19:1 to 4:1

(v/v). Fractions containing product were combined and concentrated using a rotary evaporator. The yellow solid was dissolved in 50 mL of MeOH. To the reaction solution was added slowly 10 g of pellet NaOH. The reaction was stirred at room temperature overnight. The reaction was concentrated using a rotary evaporator. The crude was purified by silica gel column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 19:1 to 4:1 (v/v). Fractions containing product were combined, concentrated using a rotary evaporator, and dried under high vacuum to give 2.1 g (73%) of **68** as a yellow solid.



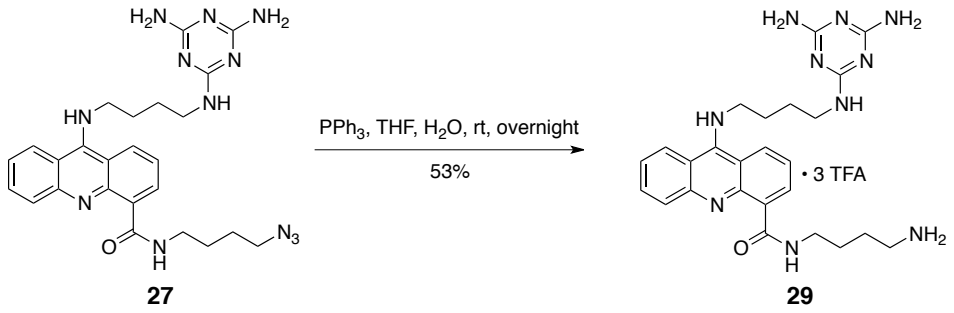
Compound **28**: To a 100 mL 14/20 1 neck round-bottom flask was added 110 mg (0.18 mmol) of **68** and 5 mL of CH_2Cl_2. The flask was placed in an ice bath. A mixture of 130 mg (0.63 mmol) of DCC and 95 mg (0.57 mmol) of 4-bromobutanoic acid **69** was dissolved in 5 mL of CH_2Cl_2. The resulting suspension was added to the reaction flask at once. The reaction was stirred at 0 °C for 1 h. The solvent was removed using a rotary evaporator. To the reaction flask was added 400 mg (6.15 mmol) of NaN_3. The solid was dissolved in 10 mL of DMF. The suspension was stirred at 80 °C in an oil bath for 3 h. The mixture was concentrated using a rotary evaporator. The crude was dissolved in a mixture containing 15 mL of TFA and 15 mL of CH_2Cl_2. The solution was stirred at room temperature for 3 h. The reaction was concentrated using a rotary evaporator. The crude was purified by C18 column chromatography using 0.1% TFA in H_2O/MeOH mixture (8:2, v/v), at a flow rate of 10 mL/min. Fractions containing product

were combined, concentrated using a rotary evaporator, and dried under high vacuum to give product 90 mg (51%) of **28** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.71 (dd, J = 8.6, 0.6, 1H), 8.53−8.48 (m, 2H), 8.03−7.98 (m, 2H), 7.65−7.60 (m, 2H), 4.21 (t, J = 7.3, 2H), 3.85 (t, J = 5.4, 2H), 3.53 (t, J = 5.4, 2H), 3.45 (t, J = 6.6, 2H), 3.38 (t, J = 5.4, 2H), 3.25 (t, J = 5.4, 2H), 3.15 (t, J = 6.8, 2H), 2.23 (t, J = 7.5, 2H), 2.06 (quint, J = 7.4, 2H), 1.77 (quint, J = 7.2, 2H), 1.63 (quint, J = 7.1, 2H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 615.3380; found 615.3380. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 6.5 min (100%).



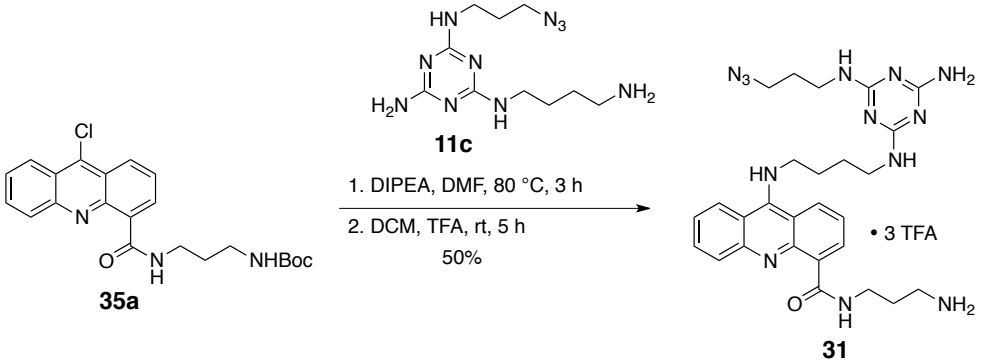
68 **30**

Compound **30**: To a 25 mL 14/20 1 neck round bottom flask was added 95 mg (0.16 mmol) of **68**. The solid was dissolved in a mixture of 5 mL of TFA and 5 mL of CH_2Cl_2. The mixture was stirred at room temperature for 3 h. Solvents were removed using a rotary evaporator, and dried overnight to give 138 mg (91%) of compound **30** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.69 (dd, J = 8.6, 1.2, 1H), 8.52−8.48 (m, 2H), 8.02−7.98 (m, 1H), 7.91 (dd, J = 8.6, 1.1, 1H), 7.60 (t, J = 8.0, 2H), 4.2 (t, J = 7.2, 2H), 3.89 (t, J = 5.5, 2H), 3.51−3.38 (m, 8H), 2.05 (quint, J = 7.3, 2H), 1.77 (quint, J = 7.3, 2H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 504.2948; found 504.2942. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 5.0 min (100%).



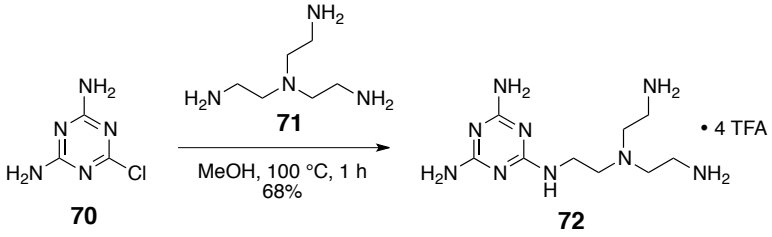
27 29

Compound **29**: To a 50 mL 14/20 round bottom flask was added 111 mg (0.22 mmol) of **27**. The yellow solid was suspended in 10 mL of THF and 1 mL of DI water. To the stirred reaction mixture was added 68 mg (0.26 mmol) of PPh_3 in several portions over 1 h. The mixture was stirred at room temperature overnight. The mixture was concentrated using a rotary evaporator. The crude was purified by C18 column chromatography using 0.1% TFA in H_2O/MeOH (8/2, v/v). The fractions containing product was combined, concentrated using a rotary evaporator to give 95 mg (53%) of **29** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.68 (dd, J = 8.7, 1.2, 1H), 8.52 (d, J = 8.5, 1H), 8.44 (dd, J = 7.5, 1.2, 1H), 8.02−7.99 (m, 1H), 7.89 (dd, J = 8.6, 1.1, 1H), 7.63−7.59 (m, 2H), 4.21 (t, J = 7.2, 2H), 3.58−3.55 (m, 2H), 3.45 (t, J = 6.6, 2H), 3.04−3.01 (m, 2H), 2.05 (quint, J = 7.5, 2H), 1.82−1.74 (m, 6H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 489.2839; found 489.2840. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 5.5 min (100%).

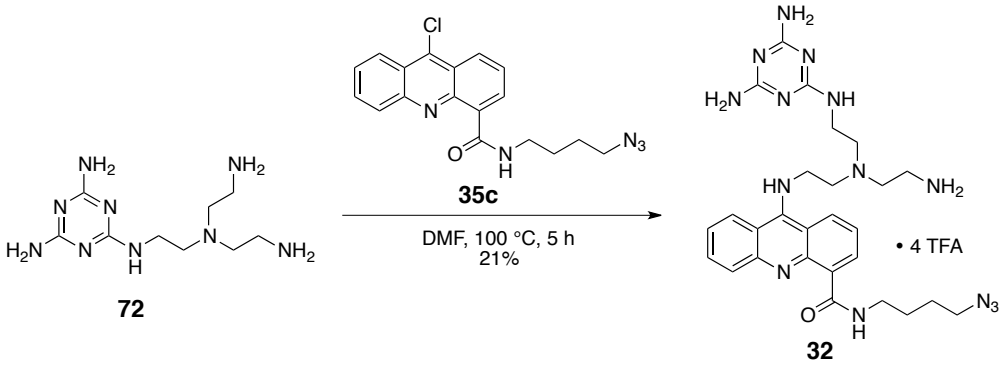


35a 31

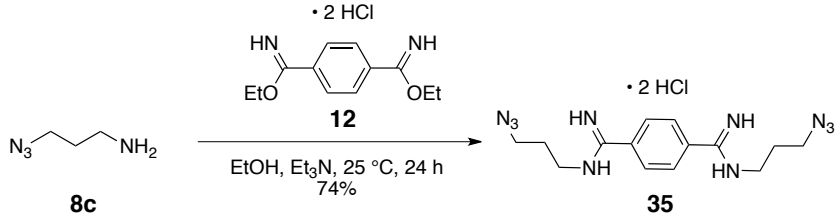
Compound **31**: Using the similar procedure described above, 219 mg (0.53 mmol) of **35a**, and 179 mg (0.64 mmol) of **11c** gave 238 mg (50%) of **31** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.69 (d, J = 8.6, 1H), 8.51 (d, J = 8.7, 1H), 8.47 (d, J = 7.5, 1H), 7.99 (t, J = 7.7, 1H), 7.88−7.86 (m, 1H), 7.62−7.57 (m, 2H), 4.22 (q, J = 6.5, 2H), 3.63 (t, J = 6.7, 2H), 3.50−3.29 (m, 6H), 3.10 (t, J = 7.4, 2H), 2.10−2.03 (m, 4H), 1.85−1.72 (m, 4H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 558.3166; found 558.3162. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 9.0 min (93%).



Compound **72**: To a 100 mL 24/40 round bottom flask was added 15.5 g (103.5 mmol) of tris(2-aminoethyl)amine **71**, and 30 mL of MeOH. The reaction flask was placed in an oil bath and heated to 100 °C. To the reaction flask was added 3.15 g (21.6 mmol) of 2-chloro-4,6-diamino-1,3,5-triazine **70** slowly over 30 min. The reaction was stirred at 100 °C for 30 min. The reaction was concentrated using a rotary evaporator. The residue was suspended in 200 mL of acetone. The solid was filtered, purified by C18 column chromatography using 0.1 % TFA in H_2O/MeOH mixture (9:1, v/v) at a flow rate of 5 mL/min. Fractions containing product were combined, concentrated using a rotary evaporator to give 3.75 g (68%) of **72** as a white TFA salt. ^{1}H NMR (DMSO-d_6): δ 8.20−7.85 (m, 11H), 3.39−3.35 (m, 2H), 2.91 (br s, 4H), 2.69 (t, J = 5.7, 4H), 2.60 (t, J = 6.3, 2H). LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 256.2; found 256.2.

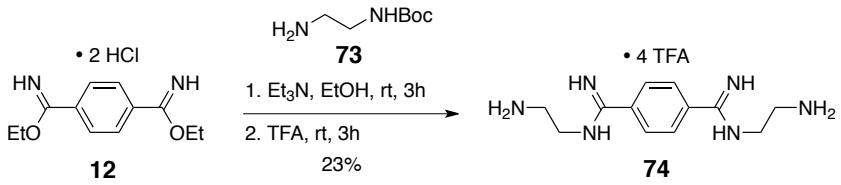


Compound **32**: To a 100 mL 14/20 1 neck round bottom flask was added 350 mg (0.99 mmol) of **35c**, 750 mg (2.94 mmol) of **72**, and 15 mL of DMF. The mixture was stirred at 100 °C for 5 h in an oil bath. The solvent was removed using a rotary evaporator. The crude was purified by C18 column chromatography using 0.1% TFA in H_2O/MeOH mixture (gradient from 100:0 to 10:90 (v/v) at a flow rate of 5 mL/min. Fractions containing product were combined and concentrated using a rotary evaporator to give 218 mg (21%) of **32** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.83 (d, J = 8.5, 1H), 8.62 (d, J = 8.7, 1H), 8.48 (d, J = 7.3, 1H), 8.05−8.03 (m, 1H), 7.95−7.94 (d, J = 8.5, 1H), 7.68−7.63 (m, 2H), 4.52 (s, 2H), 3.66 (s, 2H), 3.56 (t, J = 6.9, 4H), 3.40 (t, J = 6.6, 4H), 3.22 (q, J = 7.3, 2H), 1.83−1.70 (m, 4H), 1.32 (t, J = 7.3, 2H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 573.3273; found 573.3275. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 12.9 min (97%).



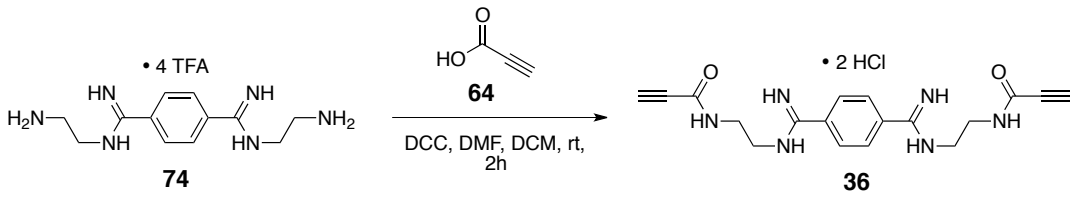
Compound **35:** To a 25 mL 14/20 round bottom flask was added 0.5 g (1.7 mmol) of **12**. The solid was suspended in 10 mL of anhydrous EtOH. To the resulting white suspension was

added 0.5 mL (3.6 mmol) of Et_3N, resulting in a clear solution. To the reaction flask was added 0.37 g (mmol) of 3-azido-propylamine **8c**. The mixture was stirred at room temperature for 20 h. The white solid was obtained by vacuum filtration, washed with a mixture of EtOH and CH_2Cl_2 (1/1, v/v) and with CH_2Cl_2. After drying under high vacuum, the solid was dissolved in 25 mL of MeOH containing 0.8 mL of concentrated HCl 37%. The resulting solution was stirred at room temperature for 2 h. The solvent was removed using a rotary evaporator. The solid was dried under high vacuum to give 0.41 g (74%) of **35** as a white HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.26−10.23 (m, 2H), 9.85 (s, 2H), 9.49 (s, 2H), 7.99 (s, 4H), 3.55−3.49 (m, 8H), 1.93 (quint, J = 6.8, 4H). HR-ESI-MS (m/z) calculated for $[M + H]^+$: 329.1951; found 329.1944. HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 5.6 min (97%).

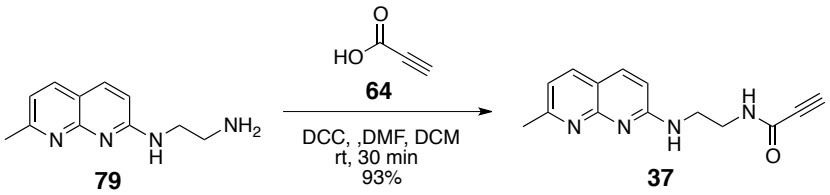


Compound **74**: To a 100 mL round bottom flask was added 0.57 g (1.9 mmol) of **12**, and 20 mL of anhydrous EtOH. To the white suspension was added 0.75 mL (5.4 mmol) of anhydrous Et_3N at once. The reaction solution was stirred at room temperature, and became clear after 10 min. To the resulting solution was added 0.66 g (4.1 mmol) of *N*-Boc-ethylenediamine **73**. The mixture was stirred at room temperature for 3 h. The solvent was removed using a rotary evaporator. The crude was dissolved in 10 mL of TFA, and stirred at room temperature for 3 h. TFA was removed using a rotary evaporator. The crude was dissolved in 5 mL of MeOH, and suspended in 200 mL of acetone. The white solid was filtered, washed with acetone, and dried under high vacuum overnight to give 0.31 (23%) of **74** as a white TFA salt. ^{1}H NMR (DMSO-

d_6): δ 10.30 (s, 2H), 9.99 (s, 2H), 9.63 (s, 2H), 8.38 (s, 6H), 8.08 (s, 4H), 3.78 (t, J = 5.9, 4H), 3.19 (t, J = 5.8, 4H).

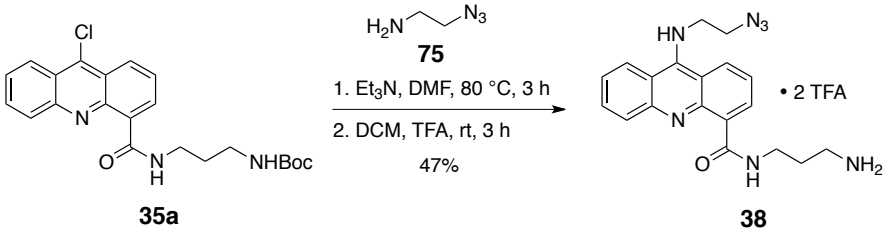


Compound **36**: To a 25 mL round bottom flask was added 0.24 g (0.34 mmol) of **74** and 5 mL of DMF. To a separated vial was added 5 mL of CH_2Cl_2, 0.24 g (1.2 mmol) of DCC, and 0.11 g (1.6 mmol) of propiolic acid **64**. The solid was filtered, the filtrate was added to the reaction flask at once. The mixture was stirred at room temperature for 30 min. The reaction was concentrated using a rotary evaporator. The crude was purified by C18 column chromatography using a gradient mixture of H_2O/MeOH from 100/0 to 90/10 (v/v) with 0.1% HCl 10M. Fractions containing products were combined and concentrated using a rotary evaporator to give 0.13 g (90%) of **36** as a white HCl salt. ^{1}H NMR (DMSO-d_6): δ 10.14 (s, 2H), 9.86 (s, 2H), 9.50 (s, 2H), 9.07 (t, J = 6.0, 2H), 7.96 (s, 4H), 4.23 (s, 2H), 3.57 (q, J = 5.5, 4H), 3.44 (q, J = 5.8, 4H). LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 353.2; found 353.3.



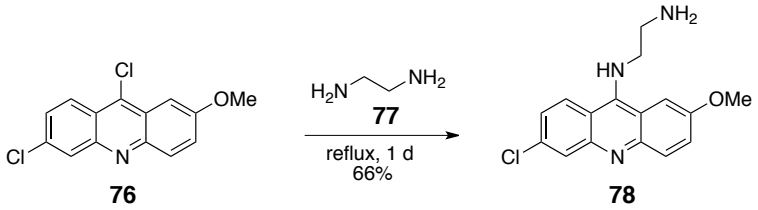
Compound **37**: To a 50 mL round bottom flask was added 0.06 g (0.3 mmol) of **79**[88] and 10 mL of DMF. To a separated vial was added 10 mL of CH_2Cl_2, 0.12 g (0.06 mmol) of DCC, and 0.04 g (0.06 mmol) of propiolic acid **64**. The resulting solution was added to the reaction

flask at once. The mixture was stirred at room temperature for 30 min. The reaction was concentrated using a rotary evaporator. The crude was purified by silica gel column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 98/2 to 95/5 (v/v). Fractions containing products were combined and concentrated using a rotary evaporator to give 0.07 g (93%) of **37** as a brown solid. ^{1}H NMR (DMSO-d_6): δ 8.91 (d, J = 5.8, 1H), 7.92 (d, J = 8.2, 1H), 7.84 (d, J = 8.8, 1H), 7,43 (t, J = 5.4, 1H), 7.05 (d, J = 7.9, 1H), 6.73 (d, J = 8.9, 1H), 4.14−4.10 (m, 1H), 3.50−3.47 (m, 2H), 3.37−3.33 (m, overlap with H_2O, 2H), 2.53 (s, 3H). HR-ESI-MS (m/z) calculated for $C_{14}H_{15}N_4O$: 255.1246; found 255.1238.

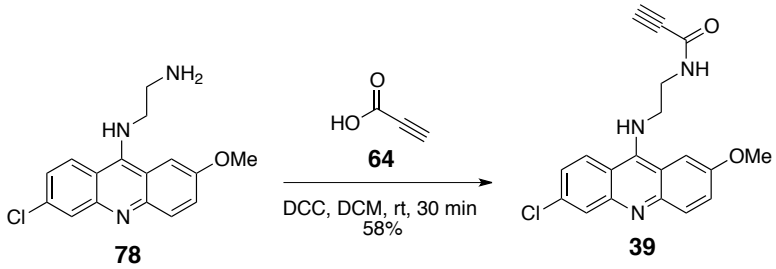


Compound **38**: To a 25 mL round bottom flask was added 990 mg (2.4 mmol) of **35a** and 10 mL DMF. To the resulting mixture was added 0.5 mL (3.6 mmol) of Et_3N, and 310 mg (3.6 mmol) of 2-azido-ethylamine **75**. The mixture was stirred at 80 °C in an oil bath for 3 h. The solvent was removed using a rotary evaporator. The crude was purified by silica column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 100/0 to 90/0 (v/v). Fractions containing product were combined and concentrated using a rotary evaporator. The crude was dissolved in 15 mL of CH_2Cl_2 and 15 mL of TFA. The mixture was stirred at room temperature for 3 h. The solvent was removed using a rotary evaporator. The crude was purified by silica column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 100/0 to 75/25 (v/v). Fractions containing product were combined and concentrated using a rotary evaporator to give 670 mg (47%) of **38** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.73 (dd, J = 8.7, 1.2, 1H),

8.56 (dd, *J* = 8.7, 0.5, 1H), 8.50 (dd, *J* = 7.5, 1.2, 1H), 8.05−8.02 (ddd, *J* = 8.3, 6.9, 1.3, 1H), 7.95−7.93 (m, 1H), 7.68−7.63 (m, 2H), 4.35 (t, *J* = 5.7, 2H), 3.95 (t, *J* = 5.7, 2H), 3.64 (t, *J* = 6.7, 2H), 3.10 (t, *J* = 7.5, 2H), 2.07 (quint, *J* = 7.1, 2H). HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 364.1886; found 364.1877.



Compound **78**: To a 100 mL round bottom flask containing 1.03g (3.7 mmol) of **76** was added 20 mL (300 mmol) of ethylenediamine **77**. The mixture was refluxed at 120 °C for 10 h, and at room temperature overnight. To the reaction mixture was added 40 mL of DI water. The yellow solid was filtered, and washed with 20 mL of DI water. The yellow solid was suspended in 100 mL Et_2O, sonicated for 1-2 min. The solid was filtered, washed with 50 mL Et_2O to give 0.74 g (66%) of **78** as a yellow solid. ^{1}H NMR ($CDCl_3$): δ 8.10 (d, *J* = 9.3, 1H), 8.05 (d, *J* = 2.1, 1H). 7.98 (d, *J* = 9.3, 1H), 7.40 (dd, *J* = 9.4, 2.6, 1H), 7.35 (d, *J* = 2.7, 1H), 7.27 (dd, *J* = 9.2, 2.2, 1H), 3.95 (s, 3H), 3.68 (m, 2H), 2.99 (m, 2H).



Compound **39**: To a 100 mL round bottom flask was added 0.35 g (1.2 mmol) of **78** and 15 mL of CH_2Cl_2. To a separated vial was added 5 mL of CH_2Cl_2, 0.51 g (2.5 mmol) of DCC,

and 0.21 g (3.0 mmol) of propiolic acid **64**. The solid was filtered, the filtrate was added to the reaction flask at once. The mixture was stirred at room temperature for 30 min. The reaction was concentrated using a rotary evaporator. The crude was purified by silica gel column chromatography using a gradient mixture of CH_2Cl_2/MeOH from 98/2 to 90/10 (v/v). Fractions containing products were combined and concentrated using a rotary evaporator to give 0.24 g (58%) of **39** as a yellow solid. ^{1}H NMR (DMSO-d_6): δ 9.10 (t, J = 4.9, 1H), 8.49 (d, J = 9.3, 1H), 7.94 (d, J = 2.2, 1H), 7.87 (d, J = 9.3, 1H), 7.81 (s, 1H), 7.61 (dd, J = 9.3, 2.4, 1H), 7.44 (dd, J = 9.4, 2.2, 1H), 4.20 (s, 1H), 4.08 (t, J = 6.5, 2H), 3.97 (s, 3H), 3.58 (q, J = 6.3, 2H). ^{13}C NMR (CD_3OD-d_4): δ 158.05, 156.40, 155.62, 143.80, 140.04, 128.70, 128.15, 124.86, 124.25, 121.23, 116.36, 112.55, 102.99, 77.72, 77.33, 56.68, 51.36, 40.89. LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 354.1; found 354.2.

Azide **40** was obtained from Julio Serrano. Azide-containing aminoglycosides **41**−**44** were prepared using a published protocol.[75]

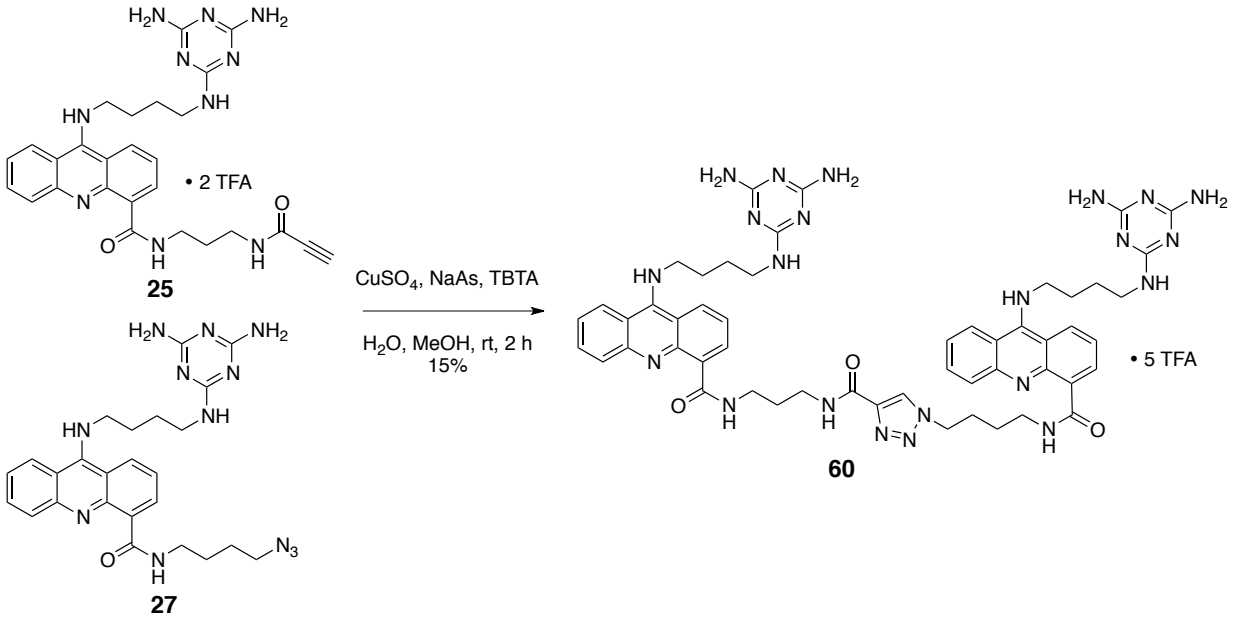
Kanamycin **41**: ^{1}H NMR (D_2O): δ 5.46−5.45 (m, 1H), 5.03 (m , 1H), 4.00−3.94 (m, 2H), 3.91−3.81 (m, 3H), 3.74−3.67 (m, 2H), 3.64−3.27 (m, 10H), 3.15−3.09 (m, 1H), 2.50−2.45 (m, 1H), 1.89−1.81 (q, J = 12.5, 1H) ; ^{13}C NMR (D_2O): δ 163.39−162.54 (q, TFA), 119.92−112.96 (q, TFA), 100.83, 96.20, 84.08, 78.92, 72.45, 72.34, 71.60, 70.91, 70.85, 68.89, 68.26, 66.57, 54.88, 50.82, 50.15, 47.71, 40.43, 30.37, 27.63.

Apramycin **42**: ^{1}H NMR (D_2O): δ 5.76 (d, J = 3.8, 1H), 5.55 (d, J = 3.9, 1H), 5.25 (d, J = 8.5, 1H), 4.59 (t, J = 2.8, 1H), 4.15−4.11 (m, 1H), 4.00−3.92 (m, 3H), 3.82−3.76 (m, 3H), 3.70−3.66 (m, 2H), 3.62−3.53 (m, 3H), 3.42−3.30 (m, 3H), 2.81 (s, 3H), 2.54−2.50 (m, 1H),

2.42−2.38 (m, 1H), 2.06 (q, J = 12.0, 1H), 1.89 (q, J = 12.6, 1H); HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 565.2946, found 565.2946.

Amikacin **43**: ^{1}H NMR (D_2O): δ 5.47 (d, J = 3.9, 1H), 5.10 (d, J = 3.8, 1H), 4.22−4.18 (m, 2H), 4.07−3.96 (m, 2H), 3.86−3.77 (m, 3H), 3.74−3.70 (m, 2H), 3.64−3.59 (m, 3H), 3.55−3.48 (m, 3H), 3.40−3.30 (m, 3H), 3.14−3.10 (m, 3H), 2.19−2.08 (m, 2H), 1.94−1.86 (m, 1H), 1.75 (q, J = 12.6, 1H), 1.12 (t, J = 7.1, 1H); ^{13}C NMR (D_2O): δ 175.61, 163.42−162.56 (q, TFA), 119.94−112.97 (q, TFA), 98.09, 95.90, 80.22, 79.54, 72.53, 72.42, 70.95, 70.87, 69.72, 68.88, 68.12, 66.55, 66.13, 55.24, 50.50, 48.97, 47.96, 40.42, 37.08, 30.92, 30.23, 14.24; HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 611.3000, found 611.2985.

Neomycin **44**: ^{1}H NMR (D_2O): δ 6.06 (d, J = 4.0, 1H), 5.37 (d, J = 2.3, 1H), 5.29 (d, J = 1.8, 1H), 4.49−4.44 (m, 1H), 4.40−4.38 (m, 1H), 4.31−4.25 (m, 2H), 4.21 (t, J = 3.1, 1H), 4.08 (dd, J = 10.4, 9, 1H), 4.01 (dd, J = 10.6, 8.9, 1H), 3.97−3.79 (m, 4H), 3.69−3.24 (m, 12H), 2.50−2.45 (m, 1H), 1.90 (q, J = 12.7, 1H), 1.14 (t, J = 7.1, 2H); ^{13}C NMR (D_2O): δ 163.32−162.48 (q, TFA), 119.95−112.99 (q, TFA), 110.86, 95.55, 95.13, 85.36, 79.54, 76.25, 75.26, 73.44, 72.52, 70.72, 70.14, 69.87, 68.18, 67.77, 67.71, 66.10, 53.53, 51.54, 50.97, 49.90, 48.58, 40.69, 40.32, 28.08, 14.24.; HR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 640.3266, found 640.3251.



Compound **60**: To a 100 mL 14/20 round bottom flask was added 210 mg (0.41 mmol) of **27**, 300 mg (0.40 mmol) of **25**, and 13 mg (0.024 mmol) of TBTA. The solid was dissolved in 40 mL of MeOH. The mixture was stirred at room temperature until a clear solution was obtained. To the reaction mixture was added 11 mg (0.044 mmol) of $CuSO_4.5H_2O$ in 10 mL of water, and 35 mg (0.18 mmol) of sodium ascorbate in 10 mL of water. The resulting mixture was stirred at room temperature for 2 h. The mixture was concentrated using a rotary evaporator. The crude was purified by silica gel column chromatography using a gradient mixture of acetone/NH_4OH from 100/0 to 95/5 (v/v). Fractions containing product were combined, and concentrated using a rotary evaporator. The crude was dissolved in MeOH, and concentrated again to remove ammonium hydroxide. The yellow solid was dissolved in a mixture of 20 mL of MeOH and 0.5 mL of TFA. The mixture was concentrated using a rotary evaporator to give 98 mg (15%) of **60** as a yellow TFA salt. ^{1}H NMR (CD_3OD-d_4): δ 8.66−8.63 (m, 2H), 8.49−8.45 (m, 3H), 8.41−8.38 (m, 2H), 7.98−7.94 (m, 2H), 7.90−7.85 (m, 4H), 7.61−7.54 (m, 4H), 4.57 (t, J = 6.9, 2H), 4.19 (q, J = 7.8, 4H), 3.61 (t, J = 6.5, 2H), 3.55 (t, J = 6.6, 4H), 3.44−3.41 (m, 4H), 2.11−1.97 (m,

8H), 1.80−1.70 (m, 6H). LR-ESI-MS (*m/z*) calculated for $[M + H]^+$: 1041.5; found 521.3 ($[M + 2H]^{2+}$, 25%), 348.0 ($[M + 3H]^{3+}$, 98%), 261.3 ($[M + 4H]^{4+}$, 100%). HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min): t = 15.4 min (96%).

4.8.2 Synthesis of r(CUG)$_{16}$ by *In Vitro* Transcription

r(CUG)$_{16}$ was prepared from the *in vitro* transcription of d(CTG)$_{16}$ using a MEGAscript T7 transcription kit (Ambion). The DNA duplex containing d(CTG)$_{16}$ and d(GAC)$_{16}$ with a T7 promoter was purchased from Integrated DNA Technology (Coralville, IA). The DNA solids were dissolved in DNase/RNase-free water. Their concentrations were determined using the Beer's law with the absorbance at 260 nm measured on a Shimadzu UV-2501PC spectrophotometer (Kyoto, Japan), and the extinction coefficient at 25 °C provided from IDT.

Structure of the DNA duplex: 5'-TTCTAATACGACTCACTATAGGG(CTG)$_{16}$CCC-3'
3'-AAGATTATGCTGAGTGATATCCC(GAC)$_{16}$GGG-5'

The DNA duplex solution containing 50 μM of d(CTG)$_{16}$, 50 μM of d(GAC)$_{16}$ was prepared in 1X transcription buffer. To a 500 μL minitube was added 35 μL of RNase-, DNase-free water, 10 μL of 10X reaction buffer, 10 μL of ATP (100 mM), 10 μL of CTP (100 mM), 10 μL of GTP (100 mM), 10 μL of UTP (100 mM), 5 μL of DNA solution, and 10 μL of T7 enzyme mix. The mixture was mixed by inverting the tube two times, and incubated at 37 °C in a water bath overnight. To degrade the excess DNA, the resulting solution was mixed with 5 μL of TURBO DNase (2U/μL), and incubated at 37 °C for 15 min. The protein was removed by phenol-chloroform extraction. To the reaction mixture was added 575 μL DNase-, RNase-free water and 75 μL of ammonium acetate solution. The solution was extracted with 750 μL of phenol:chloroform (5:1, v/v) solution. The aqueous layer was transferred to a 2 mL minitube,

and extracted with 750 μL chloroform:isoamyl alcohol (24:1, v/v) solution. The aqueous layer was transferred to two 2 mL minitubes. The unpurified RNA was recovered by ethanol precipitation. To each 2 mL minitube was added 1200 μL of cold absolute ethanol. The resulting solutions were incubated at -80 °C for 2 h. The tubes were centrifuged at 13000 rpm for 15 min at 4 °C. The solvent was removed to give the RNA pellet. RNA pellets were purified by Urea gel 20% using a UreaGel System (National Diagnostics). The gel was run at 300 V for 60 min. The gel was stained with ethidium bromide solution for 10 min, and washed with Milli-Q water for 10 min. The gel was visualized under a UV shadowing. The r(CUG)$_{16}$ band was isolated, cut into small pieces, and soaked in 500 μL of DNase-, RNase-free water at 37 °C for 2 h. The solution was collected. The RNA was recovered by ethanol precipitation. RNA pellets were dissolved in RNA storage buffer, the concentration was determined by UV-Vis.

Other DNA/RNA sequences were purchased from Integrated DNA Technology (Coralville, IA). DNA samples were stored in DNase/RNase-free water, whereas RNA was dissolved in RNA storage buffer containing 1 mM sodium citrate pH 6.4 (Ambion® #AM7001).

4.8.3 Selection Experiments Using DNA and RNA

The selection experiment containing compounds 6a−c, 23−32, 35, 38, and 45−59. Structures of r(CUG)$_{16}$, 24m-DNA, d(CTG)$_{16}$-d(GAC)$_{16}$, r(CCUG)$_8$ were shown in Figure 24. The structure of d(CTG)$_{16}$ used in this experiment is: 5'-TTCTAATACGACTCACTATAGGG(CTG)$_{16}$CCC-3'. 24m-DNA, d(CTG)$_{16}$-d(GAC)$_{16}$, r(CCUG)$_8$ were purchased from Integrated DNA Technology (Coralville, IA); yeast tRNA was purchased from Invitrogen (Carlsbad, CA); r(CUG)$_{16}$ was obtained the *in vitro* transcription as described above. The DNA and RNA solids were dissolved in DNase/RNase-free water, and the

RNA storage buffer, respectively. Their concentrations were determined using the Beer's law with the absorbance at 260 nm measured on a Shimadzu UV-2501PC spectrophotometer (Kyoto, Japan), and the extinction coefficient at 25 °C provided from IDT.

The DNA/RNA solutions (400 μL, 12.5 μM DNA/RNA, 2 mM Tris-HCl pH 7.0, 2 mM KCl, 2 mM $MgCl_2$, and 1 mM $CaCl_2$) were heated at 95 °C for 5 min, and cooled slowly to room temperature for 1 h. To the DNA/RNA solutions were added ligands to the final concentration of 25 μM each compound. The DNA/RNA solutions were shaked at 37 °C over a week. Everyday, 10 μL of the RNA-templated reaction was withdrawn, and mixed with 10 μL of Urea 8 M, and 1 μL of RNase A (10 mg/mL). Similarly, 10 μL of the DNA-templated reaction was withdrawn daily, and mixed with 1 μL of TURBO-DNase A 10X. The resulting samples were incubated at 37 °C in a water bath for 20 min, then analyzed by MALDI-MS technicians in the Mass Spectrometry Laboratory, UIUC. The reaction without DNA/RNA was conducted using the similar procedure. The experiment was triplicated.

The selection experiment containing compounds 6a−c, 24, 25, 27, and 35−59. Structures of $r(CUG)_{16}$, $d(CTG)_{16}$, 24m-DNA, $r(CCCCGG)_3$, and HIV-1 FS RNA were shown in Figure 24. The structure of $r(CCUG)_{12}$ used in this experiment is: 5'-GGG$(CCUG)_{12}$CCC-3'. $d(CTG)_{16}$, 24m-DNA and HIV-1 FS RNA were purchased from Integrated DNA Technology (Coralville, IA); yeast tRNA was purchased from Invitrogen (Carlsbad, CA); $r(CCCCGG)_3$ was obtained from Prof. Keith T. Gagnon, Southern Illinois University.

Generally, the selection experiment was conducted as described previously. For the samples containing $r(CUG)_{16}$, $r(CCUG)_{12}$, $r(CCCCGG)_3$, and tRNA: 10 μL of the reaction was mixed with 10 μL of Urea 8 M, and 2.5 μL of RNase A (10 mg/mL). Similarly, 10 μL of the

reaction using $d(CTG)_{16}$ and 24m-DNA was withdrawn daily, and mixed with 2.5 μL of TURBO-DNase A 10X. The resulting samples were incubated at 37 °C in a water bath for 30 min, then analyzed by MALDI-MS. For the HIV-1 FS RNA templated reaction: 5 μL of the reaction was mixed with 5 μL of Urea 8 M, and 2.5 μL of RNase A (10 mg/mL). The mixture was incubated at 37 °C for 60 min.

Preparation of MALDI-MS samples: 1 μL of the $r(CUG)_{16}$, $r(CCUG)_{12}$, $r(CCCCGG)_3$, tRNA, $d(CTG)_{16}$ and 24m-DNA samples or 2 μL of the HIV-1 FS RNA sample was mixed with 10 μL of the DHB matrix solution (25 mg in 1 mL of MeOH) or 10 μL of the CHCA matrix solution (10 mg in 1 mL of H_2O/acetone (1/1, v/v) with 0.1% TFA). The resulting solution (1 μL) was spotted on a MALDI-MS plate and dried slowly. The MS was obtained using a positive mode.

Monitoring $r(CUG)_{16}$-templated click reaction of 25 and 27 by HPLC. RNA-templated reactions (200 μL, 100 μM each compound, 25 μM $r(CUG)_{16}$, 2 mM Tris-HCl pH 7.0, 2 mM KCl, 2 mM $MgCl_2$, and 1 mM $CaCl_2$) were incubated at 37 °C over a week. Everyday, 15 μL of reaction was withdrawn, mixed with 15 μL of Urea 8 M, and 2 μL of RNase A (10 mg/mL). The resulting sample was incubated at 37 °C for 20 min. The sample was mixed with 68 μL of **28** (40 μM), analyzed by HPLC (0.1% TFA in H_2O/MeOH (1:1, v/v), flow rate = 3 mL/min). A parallel experiment without RNA was conducted using the similar procedure.

4.8.4 Biological Activities of the Click Product

***In vitro* transcription inhibition.** Briefly, for *in vitro* transcription inhibition experiments, ligands at desired concentrations were incubated with 15 ng of linearized $d(CTG)_{74}$ plasmid in 1X T7 transcription buffer (80 mM Tris pH 8.3, 10 mM $MgCl_2$, 2 mM spermine,

0.1% Triton-X, 10 mM NaCl) containing 0.5 mM of rATP, rCTP, rGTP, and rUTP (Promega) at 37 °C. After 3.5 h, T7 polymerase (New England Biolab) was added to the final concentration of 0.5 U, and the transcription reaction was allowed to process for 1.5 h. The reaction was quenched by adding 6.5 μL of Urea 8 M, 2 μL of denaturing dye (95% formamide, 5 mM EDTA, 0.025% xylene cyanol and bromophenol blue), and 1.5 μL of HIV-1 FS RNA (10 μM) as an internal standard and heating to 95 °C for 5 min. The solution (15 μL) was loaded on a 8% denaturing PAGE gel and the RNA was detected by EtBr post-staining. Bands were quantified using the ImageJ (NIH). The intensity ratio of r(CUG)$_{74}$/HIV FS RNA from ligand-treated sample was compared with that from the untreated sample, providing the percentage of the transcription inhibition of ligand.

Inhibition of MBNL1-r(CUG)$_{12}$ by EMSA. The IC_{50} and K_I values were determined as described in chapter 3, with $[MBNL1]_{total} = 70$ nM.

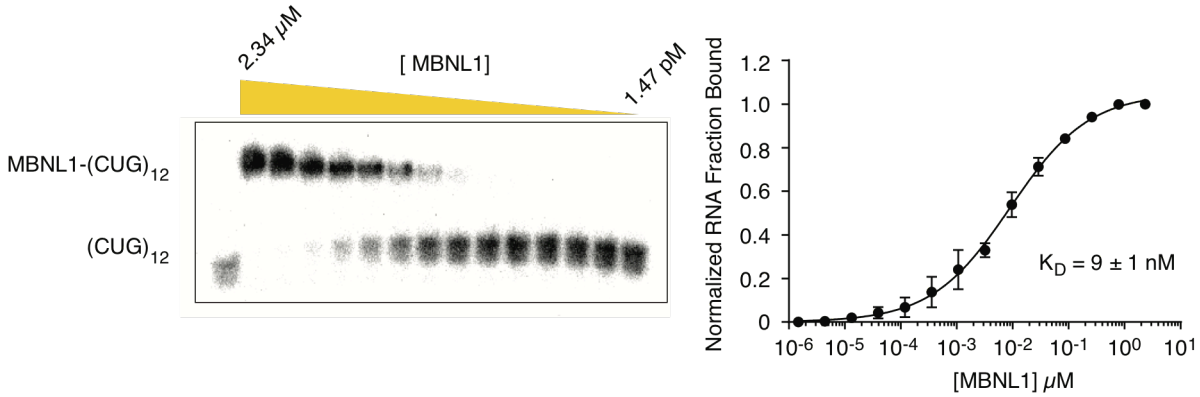


Figure 36. Determination of the binding constant of MBNL1-r(CUG)$_{12}$ by EMSA. Left: A representative gel from the EMSA experiments. The concentration of MBNL1 was diluted by a factor of 3 from 2.34 μM to 1.47 pM. Right: Plot of normalized RNA fraction bound vs. concentration of MBNL1. The standard deviation was obtained from three independent experiments.

Determination of binding affinity by fluorescence binding experiments. The fluorescence binding experiment was carried out using a Horiba Jobin Yvon fluorometer with a 1 mL cell (0.5 cm wide) at 25 °C. The RNA sample (300 μM r(CUG)$_{16}$, 20 mM MOPS pH 7.0, and 300 mM NaCl) were heated in a water bath at 95 °C for 5 min, then allowed to cool slowly to room temperature. The ligand solution (400 μL, 0.2 μM compound **25**, 20 mM MOPS pH 7.0, and 300 mM NaCl) was excited at 438 nm (excitation slit of 5 nm). Emission was recorded between 446−546 nm (emission slit of 5 nm), monitoring at 468 nm for **25**. To the ligand solution was added the freshly annealed r(CUG)$_{16}$ as 1 μL injections via pipette. The sample was then allowed to equilibrate for 15 min and the fluorescence emission was recorded. The titration was performed until the fluorescence change was no longer evident. The normalized fluorescence intensity at 468 nm versus the concentration of RNA was plotted and fit using the one binding site model with a Hill slope in Prism.

Similarly, the ligand solution for **60** (400 μL, 0.2 μM compound **60**, 20 mM MOPS pH 7.0, and 300 mM NaCl) and the RNA solution (30 μM r(CUG)$_{16}$, 20 mM MOPS pH 7.0, and 300 mM NaCl) were prepared. The ligand solution was excited at 438 nm, emission was recorded between 446−546 nm, monitoring at 513 nm.

For the binding experiment with monomer **25** and d(CTG)$_{16}$: The DNA sample (40 μM d(CTG)$_{16}$, 20 mM MOPS pH 7.0, and 300 mM NaCl) and the ligand solution (400 μL, 0.2 μM compound **25** 20 mM MOPS pH 7.0, and 300 mM NaCl) were prepared. For the binding experiment with **60** and d(CTG)$_{16}$: The DNA sample (2 μM d(CTG)$_{16}$, 20 mM MOPS pH 7.0, and 300 mM NaCl) and the ligand solution (400 μL, 0.2 μM compound **60**, 20 mM MOPS pH 7.0, and 300 mM NaCl) were prepared. The experiment was conducted as described above.

REFERENCES

(1) Buxton, J.; Shelbourne, P.; Davies, J.; Jones, C.; Van Tongeren, T.; Aslanidis, C.; de Jong, P.; Jansen, G.; Anvret, M.; Riley, B. Detection of an Unstable Fragment of DNA Specific to Individuals with Myotonic Dystrophy. *Nature* **1992**, *355*, 547–548.

(2) Lee, J. E.; Cooper, T. A. Pathogenic Mechanisms of Myotonic Dystrophy. *Biochem. Soc. Trans.* **2009**, *37*, 1281–1286.

(3) Wheeler, T. M. Myotonic Dystrophy: Therapeutic Strategies for the Future. *Neurotherapeutics* **2008**, *5*, 592–600.

(4) Cho, D. H.; Tapscott, S. J. Myotonic Dystrophy: Emerging Mechanisms for DM1 and DM2. *Biochim. Biophys. Acta* **2007**, *1772*, 195–204.

(5) Mahadevan, M.; Tsilfidis, C.; Sabourin, L.; Shutler, G.; Amemiya, C.; Jansen, G.; Neville, C.; Narang, M.; Barceló, J.; O'Hoy, K. Myotonic Dystrophy Mutation: an Unstable CTG Repeat in the 3' Untranslated Region of the Gene. *Science* **1992**, *255*, 1253–1255.

(6) Brook, J. D.; McCurrach, M. E.; Harley, H. G.; Buckler, A. J.; Church, D.; Aburatani, H.; Hunter, K.; Stanton, V. P.; Thirion, J. P.; Hudson, T. Molecular Basis of Myotonic Dystrophy: Expansion of a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member. *Cell* **1992**, *68*, 799–808.

(7) Tsilfidis, C.; MacKenzie, A. E.; Mettler, G.; Barceló, J.; Korneluk, R. G. Correlation Between CTG Trinucleotide Repeat Length and Frequency of Severe Congenital Myotonic Dystrophy. *Nat. Genet.* **1992**, *1*, 192–195.

(8) Ranum, L. P. W.; Day, J. W. Myotonic Dystrophy: RNA Pathogenesis Comes Into Focus. *Am. J. Hum. Genet.* **2004**, *74*, 793–804.

(9) Berul, C. I.; Maguire, C. T.; Aronovitz, M. J.; Greenwood, J.; Miller, C.; Gehrmann, J.; Housman, D.; Mendelsohn, M. E.; Reddy, S. DMPK Dosage Alterations Result in Atrioventricular Conduction Abnormalities in a Mouse Myotonic Dystrophy Model. *J. Clin. Invest.* **1999**, *103*, R1–7.

(10) Klesert, T. R.; Cho, D. H.; Clark, J. I.; Maylie, J.; Adelman, J.; Snider, L.; Yuen, E. C.; Soriano, P.; Tapscott, S. J. Mice Deficient in Six5 Develop Cataracts: Implications for Myotonic Dystrophy. *Nat. Genet.* **2000**, *25*, 105–109.

(11) Korade-Mirnics, Z. Myotonic Dystrophy: Tissue-Specific Effect of Somatic CTG Expansions on Allele-Specific DMAHP/SIX5 Expression. *Hum. Mol. Genet.* **1999**, *8*, 1017–1023.

(12) Narang, M. A.; Waring, J. D.; Sabourin, L. A.; Korneluk, R. G. Myotonic Dystrophy (DM) Protein Kinase Levels in Congenital and Adult DM Patients. *Eur. J. Hum. Genet.* **2000**, *8*, 507–512.

(13) Miller, J. W.; Urbinati, C. R.; Teng-Umnuay, P.; Stenberg, M. G.; Byrne, B. J.; Thornton, C. A.; Swanson, M. S. Recruitment of Human Muscleblind Proteins to $(CUG)_n$ Expansions Associated with Myotonic Dystrophy. *EMBO J.* **2000**, *19*, 4439–4448.

(14) Warf, M. B.; Berglund, J. A. MBNL Binds Similar RNA Structures in the CUG Repeats of Myotonic Dystrophy and Its Pre-mRNA Substrate Cardiac Troponin T. *RNA* **2007**, *13*, 2238–2251.

(15) Fardaei, M.; Larkin, K.; Brook, J. D.; Hamshere, M. G. *In Vivo* Co-Localisation of MBNL Protein with DMPK Expanded-Repeat Transcripts. *Nucleic Acids Res.* **2001**, *29*, 2766–2771.

(16) Du, H.; Cline, M. S.; Osborne, R. J.; Tuttle, D. L.; Clark, T. A.; Donohue, J. P.; Hall, M. P.; Shiue, L.; Swanson, M. S.; Thornton, C. A.; Ares, M. Aberrant Alternative Splicing and Extracellular Matrix Gene Expression in Mouse Models of Myotonic Dystrophy. *Nat. Struct. Mol. Biol.* **2010**, *17*, 187–193.

(17) Gatchel, J. R.; Zoghbi, H. Y. Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nat. Rev. Genet.* **2005**, *6*, 743–755.

(18) Wheeler, T. M.; Lueck, J. D.; Swanson, M. S.; Dirksen, R. T.; Thornton, C. A. Correction of ClC-1 Splicing Eliminates Chloride Channelopathy and Myotonia in Mouse Models of Myotonic Dystrophy. *J. Clin. Invest.* **2007**, *117*, 3952–3957.

(19) Savkur, R. S.; Philips, A. V.; Cooper, T. A. Aberrant Regulation of Insulin Receptor Alternative Splicing Is Associated with Insulin Resistance in Myotonic Dystrophy. *Nat. Genet.* **2001**, *29*, 40–47.

(20) Hino, S. I.; Kondo, S.; Sekiya, H.; Saito, A.; Kanemoto, S.; Murakami, T.; Chihara, K.; Aoki, Y.; Nakamori, M.; Takahashi, M. P.; Imaizumi, K. Molecular Mechanisms Responsible for Aberrant Splicing of SERCA1 in Myotonic Dystrophy Type 1. *Hum. Mol. Genet.* **2007**, *16*, 2834–2843.

(21) Ho, T. H.; Bundman, D.; Armstrong, D. L.; Cooper, T. A. Transgenic Mice Expressing CUG-BP1 Reproduce Splicing Mis-Regulation Observed in Myotonic Dystrophy. *Hum. Mol. Genet.* **2005**, *14*, 1539–1547.

(22) Kuyumcu-Martinez, N. M.; Wang, G.-S.; Cooper, T. A. Increased Steady-State Levels of CUGBP1 in Myotonic Dystrophy 1 Are Due to PKC-Mediated Hyperphosphorylation. *Mol. Cell* **2007**, *28*, 68–78.

(23) Kalsotra, A.; Singh, R. K.; Gurha, P.; Ward, A. J.; Creighton, C. J.; Cooper, T. A. The

Mef2 Transcription Network Is Disrupted in Myotonic Dystrophy Heart Tissue, Dramatically Altering miRNA and mRNA Expression. *Cell Rep.* **2014**, *6*, 336–345.

(24) Zu, T.; Gibbens, B.; Doty, N. S.; Gomes-Pereira, M.; Huguet, A.; Stone, M. D.; Margolis, J.; Peterson, M.; Markowski, T. W.; Ingram, M. A. C.; Nan, Z.; Forster, C.; Low, W. C.; Schoser, B.; Somia, N. V.; Clark, H. B.; Schmechel, S.; Bitterman, P. B.; Gourdon, G.; Swanson, M. S.; Moseley, M.; Ranum, L. P. W. Non-ATG-Initiated Translation Directed by Microsatellite Expansions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 260–265.

(25) Kanadia, R. N.; Johnstone, K. A.; Mankodi, A.; Lungu, C.; Thornton, C. A.; Esson, D.; Timmers, A. M.; Hauswirth, W. W.; Swanson, M. S. A Muscleblind Knockout Model for Myotonic Dystrophy. *Science* **2003**, *302*, 1978–1980.

(26) Kanadia, R. N.; Shin, J.; Yuan, Y.; Beattie, S. G.; Wheeler, T. M.; Thornton, C. A.; Swanson, M. S. Reversal of RNA Missplicing and Myotonia After Muscleblind Overexpression in a Mouse Poly(CUG) Model for Myotonic Dystrophy. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11748–11753.

(27) Mooers, B. H. M.; Logue, J. S.; Berglund, J. A. The Structural Basis of Myotonic Dystrophy From the Crystal Structure of CUG Repeats. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16626–16631.

(28) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. Structural Insights Into CUG Repeats Containing the "Stretched U-U Wobble": Implications for Myotonic Dystrophy. *Nucleic Acids Res.* **2009**, *37*, 4149–4156.

(29) 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 **2009**, *106*, 18551–18556.

(30) Childs-Disney, J. L.; Stepniak-Konieczna, E.; Tran, T.; Yildirim, I.; Park, H.; Chen, C. Z.; Hoskins, J.; Southall, N.; Marugan, J. J.; Patnaik, S.; Zheng, W.; Austin, C. P.; Schatz, G. C.; Sobczak, K.; Thornton, C. A.; Disney, M. D. Induction and Reversal of Myotonic Dystrophy Type 1 Pre-mRNA Splicing Defects by Small Molecules. *Nat. Commun.* **2013**, *4*, 2044.

(31) Pushechnikov, A.; Lee, M. M.; Childs-Disney, J. L.; Sobczak, K.; French, J. M.; Thornton, C. A.; Disney, M. D. 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.* **2009**, *131*, 9767–9779.

(32) Childs-Disney, J. L.; Parkesh, R.; Nakamori, M.; Thornton, C. A.; Disney, M. D. Rational Design of Bioactive, Modularly Assembled Aminoglycosides Targeting the RNA That Causes Myotonic Dystrophy Type 1. *ACS Chem. Biol.* **2012**, *7*, 1984–1993.

(33) Guan, L.; Disney, M. D. Small-Molecule-Mediated Cleavage of RNA in Living Cells. *Angew. Chem. Int. Ed.* **2013**, *52*, 1462–1465.

(34) Guan, L.; Disney, M. D. Covalent Small-Molecule-RNA Complex Formation Enables Cellular Profiling of Small-Molecule-RNA Interactions. *Angew. Chem. Int. Ed.* **2013**, *52*, 10010–10013.

(35) Mulders, S. A. M.; van den Broek, W. J. A. A.; Wheeler, T. M.; Croes, H. J. E.; van Kuik-Romeijn, P.; de Kimpe, S. J.; Furling, D.; Platenburg, G. J.; Gourdon, G.; Thornton, C. A.; Wieringa, B.; Wansink, D. G. Triplet-Repeat Oligonucleotide-Mediated Reversal of RNA Toxicity in Myotonic Dystrophy. *Proc. Natl. Acad. Sci. USA*

2009, *106*, 13915–13920.

(36) Coonrod, L. A.; Nakamori, M.; Wang, W.; Carrell, S.; Hilton, C. L.; Bodner, M. J.; Siboni, R. B.; Docter, A. G.; Haley, M. M.; Thornton, C. A.; Berglund, J. A. Reducing Levels of Toxic RNA with Small Molecules. *ACS Chem. Biol.* **2013**, *8*, 2528–2537.

(37) Siboni, R. B.; Nakamori, M.; Wagner, S. D.; Struck, A. J.; Coonrod, L. A.; Harriott, S. A.; Cass, D. M.; Tanner, M. K.; Berglund, J. A. Actinomycin D Specifically Reduces Expanded CUG Repeat RNA in Myotonic Dystrophy Models. *Cell Rep.* **2015**, *13*, 2386–2394.

(38) Asamitsu, S.; Kawamoto, Y.; Hashiya, F.; Hashiya, K.; Yamamoto, M.; Kizaki, S.; Bando, T.; Sugiyama, H. Sequence-Specific DNA Alkylation and Transcriptional Inhibition by Long-Chain Hairpin Pyrrole-Imidazole Polyamide-Chlorambucil Conjugates Targeting CAG/CTG Trinucleotide Repeats. *Bioorg. Med. Chem.* **2014**, *22*, 4646–4657.

(39) Chen, G.; Masuda, A.; Konishi, H.; Ohkawara, B.; Ito, M.; Kinoshita, M.; Kiyama, H.; Matsuura, T.; Ohno, K. Phenylbutazone Induces Expression of MBNL1 and Suppresses Formation of MBNL1-CUG RNA Foci in a Mouse Model of Myotonic Dystrophy. *Sci. Rep.* **2016**, 25317.

(40) Wong, C.-H.; Richardson, S. L.; Ho, Y.-J.; Lucas, A. M. H.; Tuccinardi, T.; Baranger, A. M.; Zimmerman, S. C. Investigating the Binding Mode of an Inhibitor of the MBNL1·RNA Complex in Myotonic Dystrophy Type 1 (DM1) Leads to the Unexpected Discovery of a DNA-Selective Binder. *ChemBioChem* **2012**, *13*, 2505–2509.

(41) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein

Binding. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16068–16073.

(42) Haghighat Jahromi, A.; Honda, M.; Zimmerman, S. C.; Spies, M. Single-Molecule Study of the CUG Repeat-MBNL1 Interaction and Its Inhibition by Small Molecules. *Nucleic Acids Res.* **2013**, *41*, 6687–6697.

(43) Jahromi, A. H.; Nguyen, L.; Fu, Y.; Miller, K. A.; Baranger, A. M.; Zimmerman, S. C. A Novel CUG^{exp}·MBNL1 Inhibitor with Therapeutic Potential for Myotonic Dystrophy Type 1. *ACS Chem. Biol.* **2013**, *8*, 1037–1043.

(44) Jahromi, A. H.; Fu, Y.; Miller, K. A.; Nguyen, L.; Luu, L. M.; Baranger, A. M.; Zimmerman, S. C. Developing Bivalent Ligands to Target CUG Triplet Repeats, the Causative Agent of Myotonic Dystrophy Type 1. *J. Med. Chem.* **2013**, *56*, 9471–9481.

(45) Wong, C.-H.; Nguyen, L.; Peh, J.; Luu, L. M.; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. Targeting Toxic RNAs That Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *J. Am. Chem. Soc.* **2014**, *136*, 6355–6361.

(46) Marcheschi, R. J.; Tonelli, M.; Kumar, A.; Butcher, S. E. Structure of the HIV-1 Frameshift Site RNA Bound to a Small Molecule Inhibitor of Viral Replication. *ACS Chem. Biol.* **2011**, *6*, 857–864.

(47) Salim, N. N.; Feig, A. L. Isothermal Titration Calorimetry of RNA. *Methods* **2009**, *47*, 198–205.

(48) Molinaro, M.; Tinoco, I. Use of Ultra Stable UNCG Tetraloop Hairpins to Fold RNA Structures: Thermodynamic and Spectroscopic Applications. *Nucleic Acids Res.* **1995**, *23*, 3056–3063.

(49) Nguyen, L.; Luu, L. M.; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C.

Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc.* **2015**, *137*, 14180–14189.

(50) Meldal, M.; Tornøe, C. W. Cu-Catalyzed Azide−Alkyne Cycloaddition. *Chem. Rev.* **2008**, *108*, 2952–3015.

(51) Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. Polytriazoles as Copper(I)-Stabilizing Ligands in Catalysis. *Org. Lett.* **2004**, *6*, 2853–2855.

(52) Hellman, L. M.; Fried, M. G. Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions. *Nat. Protoc.* **2007**, *2*, 1849–1861.

(53) Yuan, Y.; Compton, S. A.; Sobczak, K.; Stenberg, M. G.; Thornton, C. A.; Griffith, J. D.; Swanson, M. S. Muscleblind-Like 1 Interacts with RNA Hairpins in Splicing Target and Pathogenic RNAs. *Nucleic Acids Res.* **2007**, *35*, 5474–5486.

(54) Garcia-Lopez, A.; Monferrer, L.; Garcia-Alcover, I.; Vicente-Crespo, M.; Alvarez-Abril, M. C.; Artero, R. D. Genetic and Chemical Modifiers of a CUG Toxicity Model in *Drosophila*. *PLoS ONE* **2008**, *3*, e1595.

(55) Garcia-Alcover, I.; Colonques-Bellmunt, J.; Garijo, R.; Tormo, J. R.; Artero, R.; Alvarez-Abril, M. C.; López Castel, A.; Perez-Alonso, M. Development of a *Drosophila* Melanogaster Spliceosensor System for *In Vivo* High-Throughput Screening in Myotonic Dystrophy Type 1. *Dis. Model. Mech.* **2014**, *7*, 1297–1306.

(56) Garcia-Lopez, A.; Llamusi, B.; Orzaez, M.; Perez-Paya, E.; Artero, R. D. In Vivo Discovery of a Peptide That Prevents CUG-RNA Hairpin Formation and Reverses RNA Toxicity in Myotonic Dystrophy Models. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11866–11871.

(57) Nichols, C. D.; Becnel, J.; Pandey, U. B. Methods to Assay Drosophila Behavior. *J. Vis.*

Exp. **2012**, *61*, e3795.

(58) Lipinski, C. A. Drug-Like Properties and the Causes of Poor Solubility and Poor Permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.

(59) Bai, Y.; Feng, X.; Xing, H.; Xu, Y.; Kim, B. K.; Baig, N.; Zhou, T.; Gewirth, A. A.; Lu, Y.; Oldfield, E.; Zimmerman, S. C. A Highly Efficient Single-Chain Metal-Organic Nanoparticle Catalyst for Alkyne-Azide "Click" Reactions in Water and in Cells. *J. Am. Chem. Soc.* **2016**, *138*, 11077–11080.

(60) Antonow, D. Fragment-Based Approaches and the Prospect of Fragmented Prodrugs. *Drug Discov. Today* **2010**, *15*, 801–803.

(61) Rideout, D.; Calogeropoulou, T.; Jaworski, J.; McCarthy, M. Synergism Through Direct Covalent Bonding Between Agents: a Strategy for Rational Design of Chemotherapeutic Combinations. *Biopolymers* **1990**, *29*, 247–262.

(62) Ramström, O.; Lehn, J.-M. Drug Discovery by Dynamic Combinatorial Libraries. *Nat. Rev. Drug Discov.* **2002**, *1*, 26–36.

(63) Hu, X.; Manetsch, R. Kinetic Target-Guided Synthesis. *Chem. Soc. Rev.* **2010**, *39*, 1316–1324.

(64) Thirumurugan, P.; Matosiuk, D.; Jozwiak, K. Click Chemistry for Drug Development and Diverse Chemical–Biology Applications. *Chem. Rev.* **2013**, *113*, 4905–4979.

(65) Oueis, E.; Sabot, C.; Renard, P.-Y. New Insights Into the Kinetic Target-Guided Synthesis of Protein Ligands. *Chem. Commun.* **2015**, *51*, 12158–12169.

(66) Poulin-Kerstien, A. T.; Dervan, P. B. DNA-Templated Dimerization of Hairpin Polyamides. *J. Am. Chem. Soc.* **2003**, *125*, 15811–15821.

(67) Rzuczek, S. G.; Park, H.; Disney, M. D. A Toxic RNA Catalyzes the in Cellulo

Synthesis of Its Own Inhibitor. *Angew. Chem. Int. Ed.* **2014**, *53*, 10956–10959.

(68) Imoto, S.; Hirohama, T.; Nagatsugi, F. DNA-Templated Click Chemistry for Creation of Novel DNA Binding Molecules. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5660–5663.

(69) Di Antonio, M.; Biffi, G.; Mariani, A.; Raiber, E.-A.; Rodriguez, R.; Balasubramanian, S. Selective RNA Versus DNA G-Quadruplex Targeting by *In Situ* Click Chemistry. *Angew. Chem. Int. Ed.* **2012**, *51*, 11073–11078.

(70) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. 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.* **2008**, *130*, 16254–16261.

(71) Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. Finding the Right (Bioorthogonal) Chemistry. *ACS Chem. Biol.* **2014**, *9*, 592–605.

(72) Wong, C.-H.; Zimmerman, S. C. Orthogonality in Organic, Polymer, and Supramolecular Chemistry: From Merrifield to Click Chemistry. *Chem. Commun.* **2013**, *49*, 1679–1695.

(73) Kislukhin, A. A.; Hong, V. P.; Breitenkamp, K. E.; Finn, M. G. Relative Performance of Alkynes in Copper-Catalyzed Azide-Alkyne Cycloaddition. *Bioconjugate Chem.* **2013**, *24*, 684–689.

(74) Peng, T.; Nakatani, K. Binding of Naphthyridine Carbamate Dimer to the $(CGG)_n$ Repeat Results in the Disruption of the G-C Base Pairing. *Angew. Chem. Int. Ed.* **2005**, *44*, 7280–7283.

(75) Tsitovich, P. B.; Pushechnikov, A.; French, J. M.; Disney, M. D. A Chemoenzymatic Route to Diversify Aminoglycosides Enables a Microarray-Based Method to Probe

Acetyltransferase Activity. *ChemBioChem* **2010**, *11*, 1656–1660.

(76) Thomas, J. R.; Liu, X.; Hergenrother, P. J. Size-Specific Ligands for RNA Hairpin Loops. *J. Am. Chem. Soc.* **2005**, *127*, 12434–12435.

(77) Thomas, J. R.; Hergenrother, P. J. Targeting RNA with Small Molecules. *Chem. Rev.* **2008**, *108*, 1171–1224.

(78) Trimpin, S.; Inutan, E. D.; Herath, T. N.; McEwen, C. N. Laserspray Ionization, a New Atmospheric Pressure MALDI Method for Producing Highly Charged Gas-Phase Ions of Peptides and Proteins Directly From Solid Solutions. *Mol. Cell Proteomics* **2010**, *9*, 362–367.

(79) Pearson, C. E.; Wang, Y. H.; Griffith, J. D.; Sinden, R. R. Structural Analysis of Slipped-Strand DNA (S-DNA) Formed in $(CTG)_n \cdot (CAG)_n$ Repeats From the Myotonic Dystrophy Locus. *Nucleic Acids Res.* **1998**, *26*, 816–823.

(80) Wong, C. H.; Fu, Y.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. Selective Inhibition of MBNL1-CCUG Interaction by Small Molecules Toward Potential Therapeutic Agents for Myotonic Dystrophy Type 2 (DM2). *Nucleic Acids Res.* **2011**, *39*, 8881–8890.

(81) Nguyen, L.; Lee, J.; Wong, C.-H.; Zimmerman, S. C. Small Molecules That Target the Toxic RNA in Myotonic Dystrophy Type 2. *ChemMedChem* **2014**, *9*, 2455–2462.

(82) Dodd, D. W.; Tomchick, D. R.; Corey, D. R.; Gagnon, K. T. Pathogenic C9ORF72 Antisense Repeat RNA Forms a Double Helix with Tandem C:C Mismatches. *Biochemistry* **2016**, *55*, 1283–1286.

(83) Li, X.; Liu, D. R. DNA-Templated Organic Synthesis: Nature's Strategy for Controlling Chemical Reactivity Applied to Synthetic Molecules. *Angew. Chem. Int. Ed.* **2004**, *43*,

4848–4870.

(84) Gartner, Z. J.; Kanan, M. W.; Liu, D. R. Expanding the Reaction Scope of DNA-Templated Synthesis. *Angew. Chem. Int. Ed.* **2002**, *41*, 1796–1800.

(85) Pearson, C. E.; Tam, M.; Wang, Y. H.; Montgomery, S. E.; Dar, A. C.; Cleary, J. D.; Nichol, K. Slipped-Strand DNAs Formed by Long (CAG)·(CTG) Repeats: Slipped-Out Repeats and Slip-Out Junctions. *Nucleic Acids Res.* **2002**, *30*, 4534–4547.

(86) Bai, Y.; Nguyen, L.; Song, Z.; Peng, S.; Lee, J.; Zheng, N.; Kapoor, I.; Hagler, L. D.; Cai, K.; Cheng, J.; Chan, H. Y. E.; Zimmerman, S. C. Integrating Display and Delivery Functionality with a Cell Penetrating Peptide Mimic as a Scaffold for Intracellular Multivalent Multitargeting. *J. Am. Chem. Soc.* **2016**, *138*, 9498–9507.

(87) Rideout, D. Self-Assembling Cytotoxins. *Science* **1986**, *233*, 561–563.

(88) Meng, L.; Wu, Y.; Yi, T. A Ratiometric Fluorescent Probe for the Detection of Hydroxyl Radicals in Living Cells. *Chem. Commun.* **2014**, *50*, 4843–4845.

APPENDIX A

LIST OF PUBLICATIONS

(1) **Luu, L. M.**; Lee, J.; Nguyen, L.; Vergara, J. I.; Zimmerman, S. C. A Fragment-Based, Multi-target Drug Discovery Approach for Both DNA and RNA Targets Using a Template-Assisted Selection. (Manuscript in preparation)

(2) **Luu, L. M.**; Nguyen, L.; Peng, S.; Lee, J.; Lee, H. Y.; Wong, C.-H.; Hergenrother, P. J.; Chan, H. Y. E.; Zimmerman, S. C. A Potent Inhibitor of Protein Sequestration by Expanded Triplet (CUG) Repeats that Shows Phenotypic Improvements in a *Drosophila* Model of Myotonic Dystrophy. *ChemMedChem* **2016**, *11*, 1428−1435.

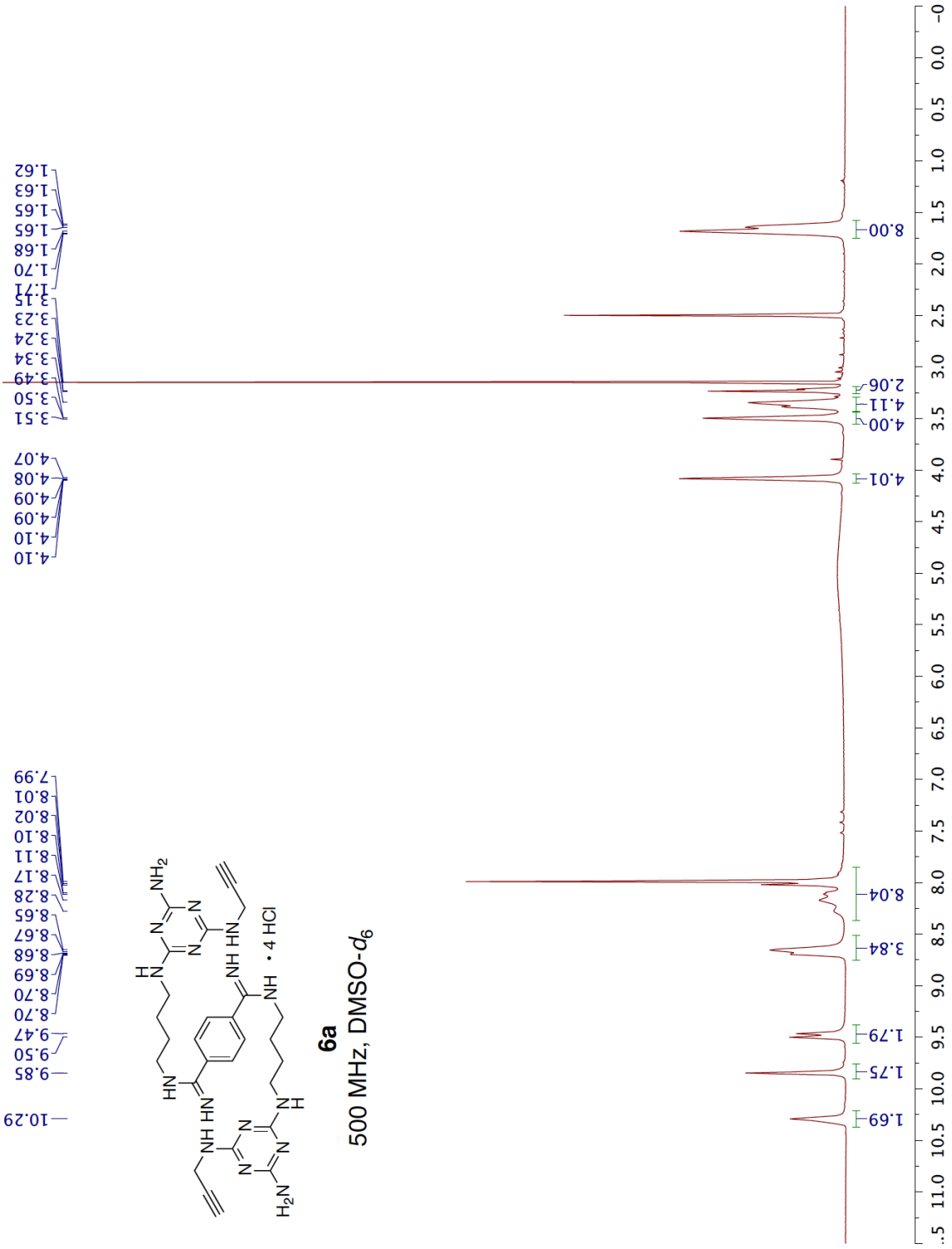
(3) Nguyen, L.; **Luu, L. M.**; Peng, S.; Serrano, J. F.; Chan, H. Y. E.; Zimmerman, S. C. Rationally Designed Small Molecules That Target Both the DNA and RNA Causing Myotonic Dystrophy Type 1. *J. Am. Chem. Soc.* **2015**, *137*, 14180−14189.

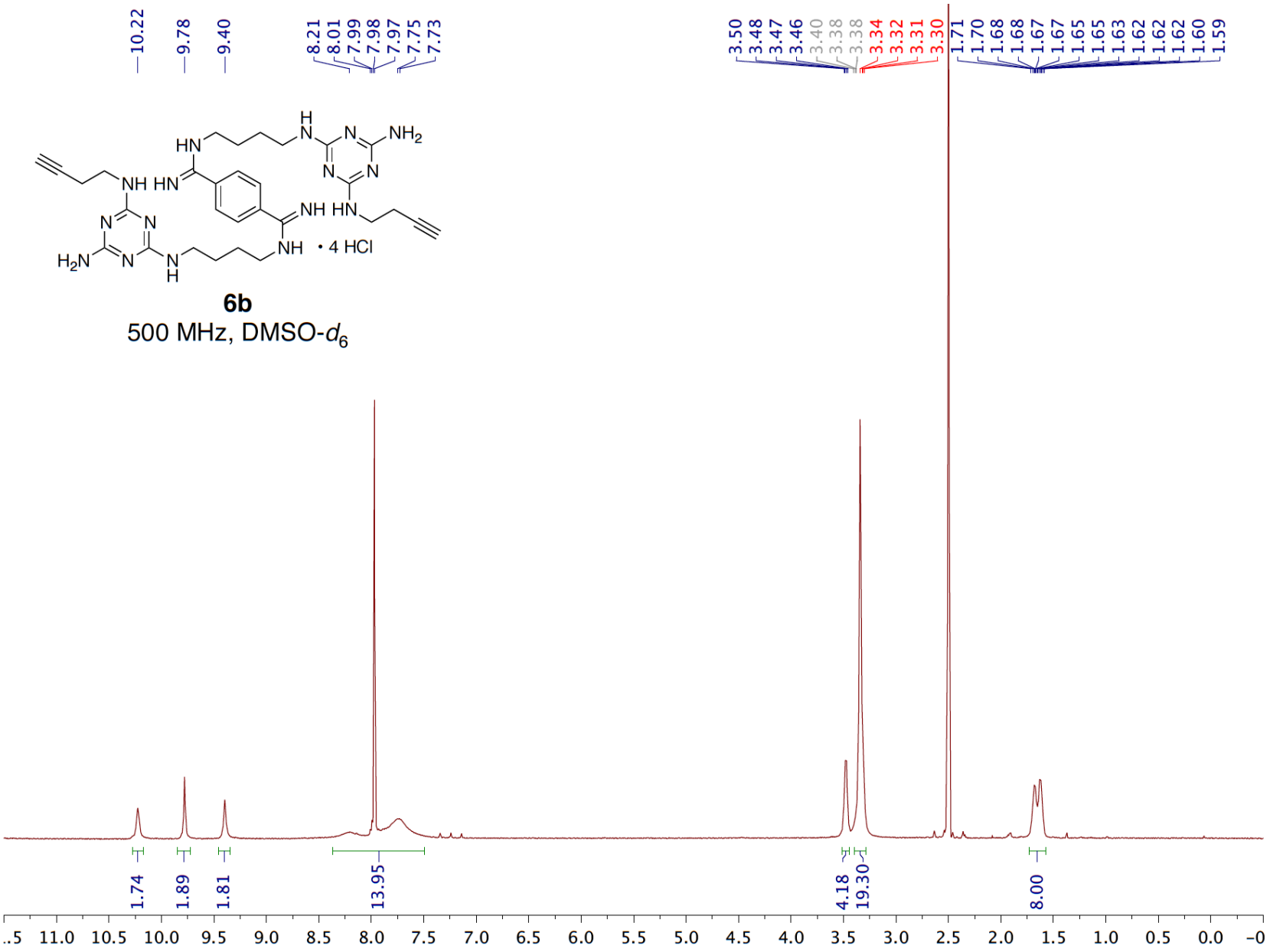
(4) Wong, C.-H.; Nguyen, L.; Peh, J.; **Luu, L. M.**; Sanchez, J. S.; Richardson, S. L.; Tuccinardi, T.; Tsoi, H.; Chan, W. Y.; Chan, H. Y. E.; Baranger, A. M.; Hergenrother, P. J.; Zimmerman, S. C. Targeting Toxic RNAs that Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *J. Am. Chem. Soc.* **2014**, *136*, 6355−6361.

(5) Jahromi, A. H.; Fu, Y.; Miller, K. A.; Nguyen, L.; **Luu, L. M.**; Baranger, A. M.; Zimmerman, S. C. Developing Bivalent Ligands to Target CUG Triplet Repeats, the Causative Agent of Myotonic Dystrophy Type 1. *J. Med. Chem.* **2013**, *56*, 9471−9481.

APPENDIX B

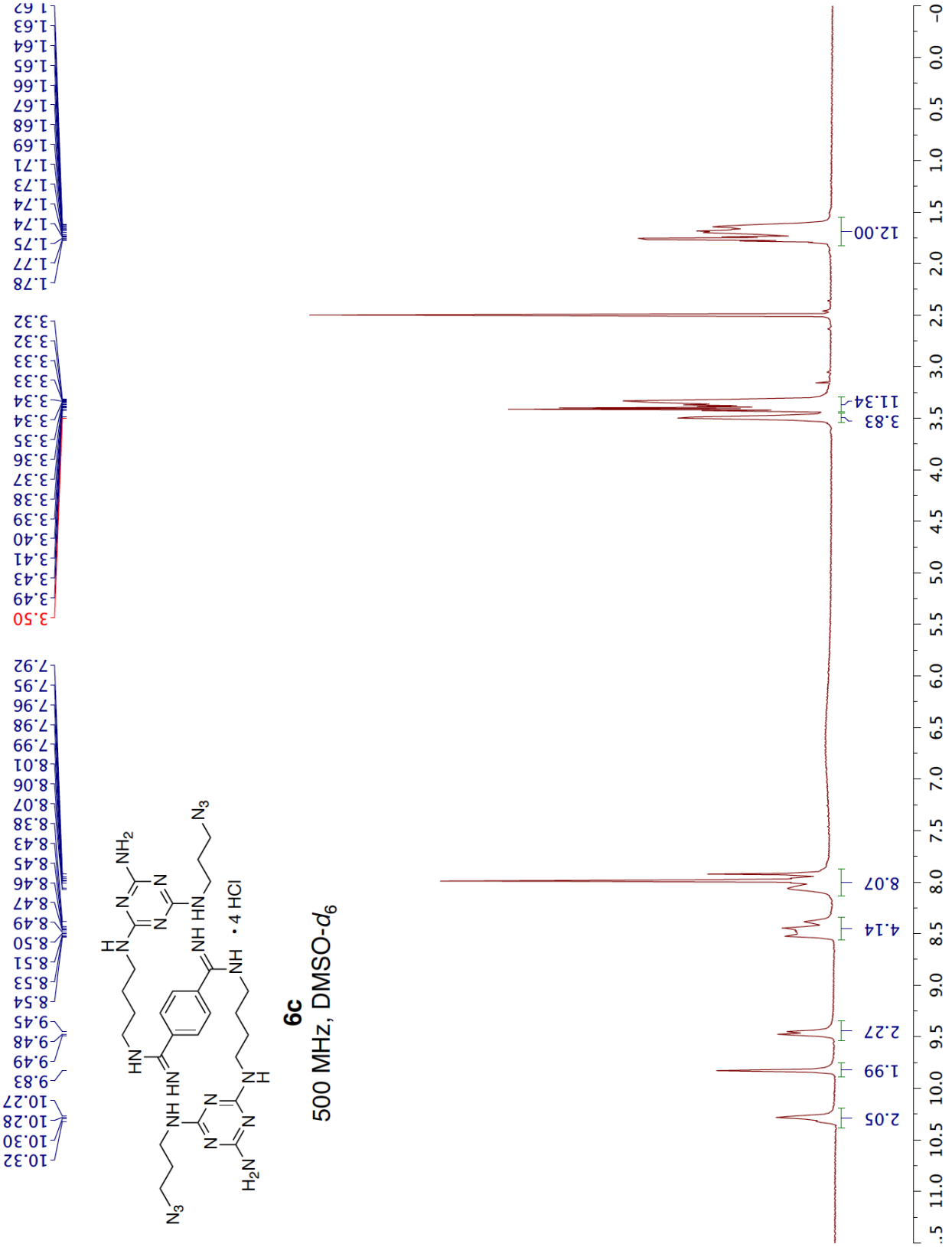
NMR SPECTRA

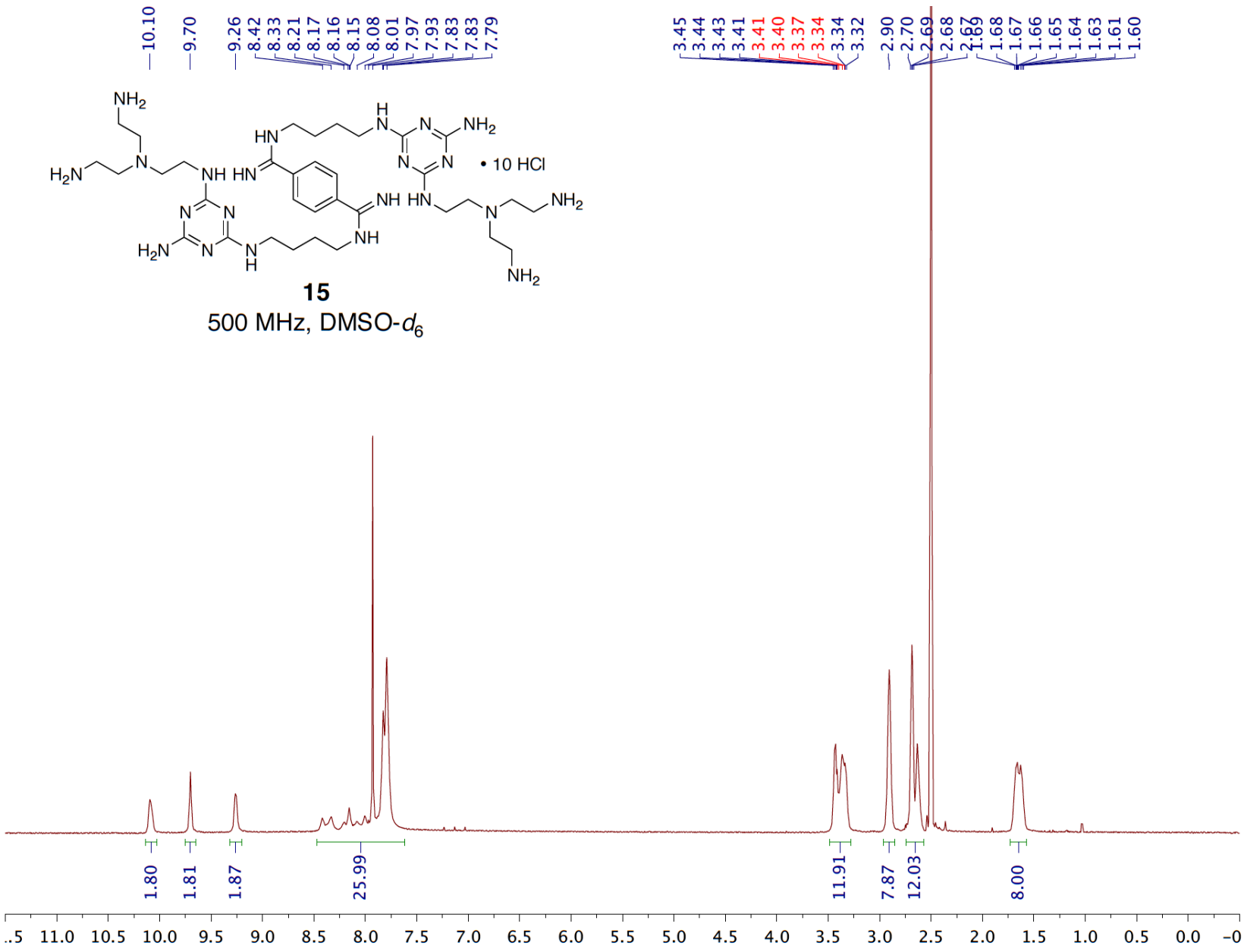


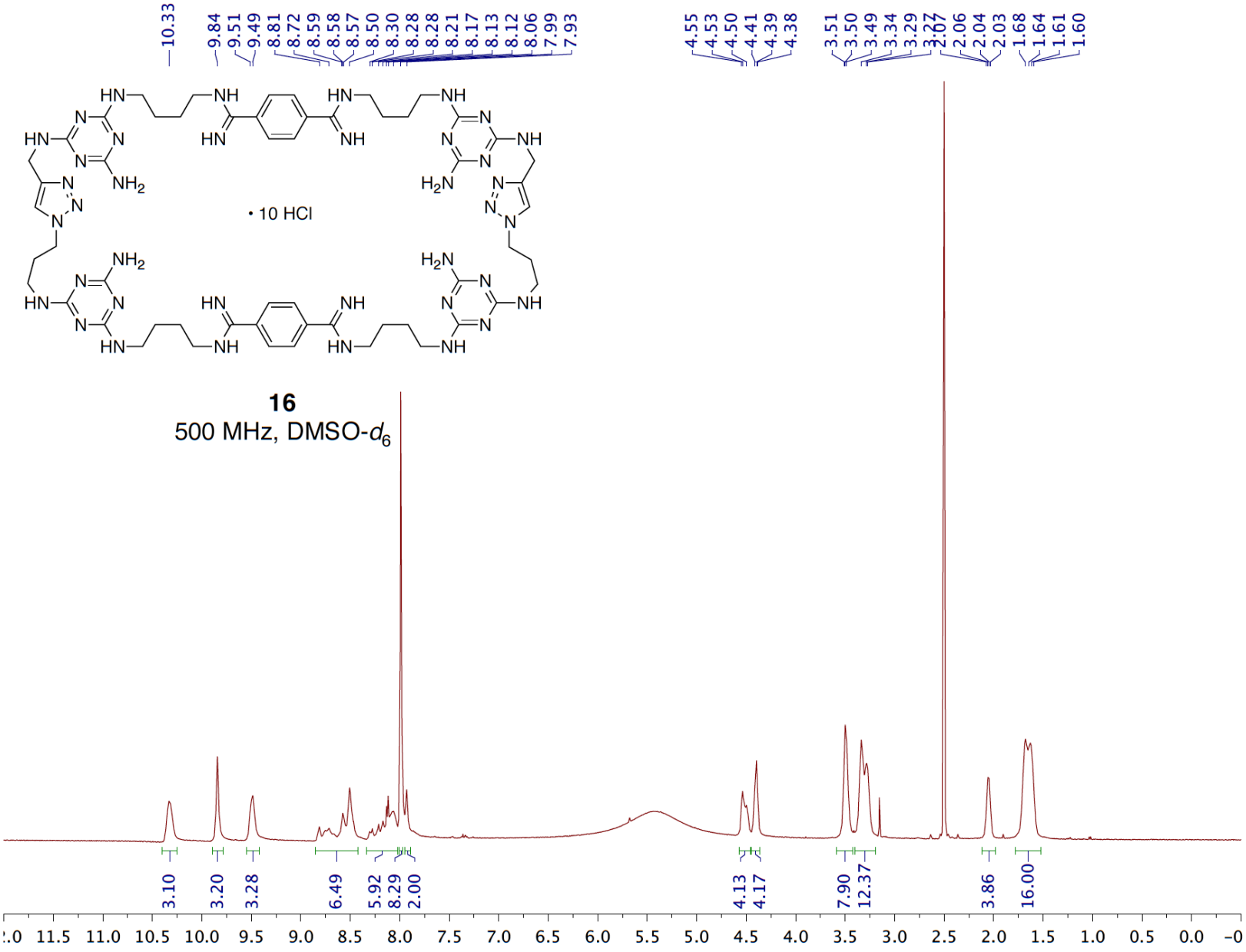


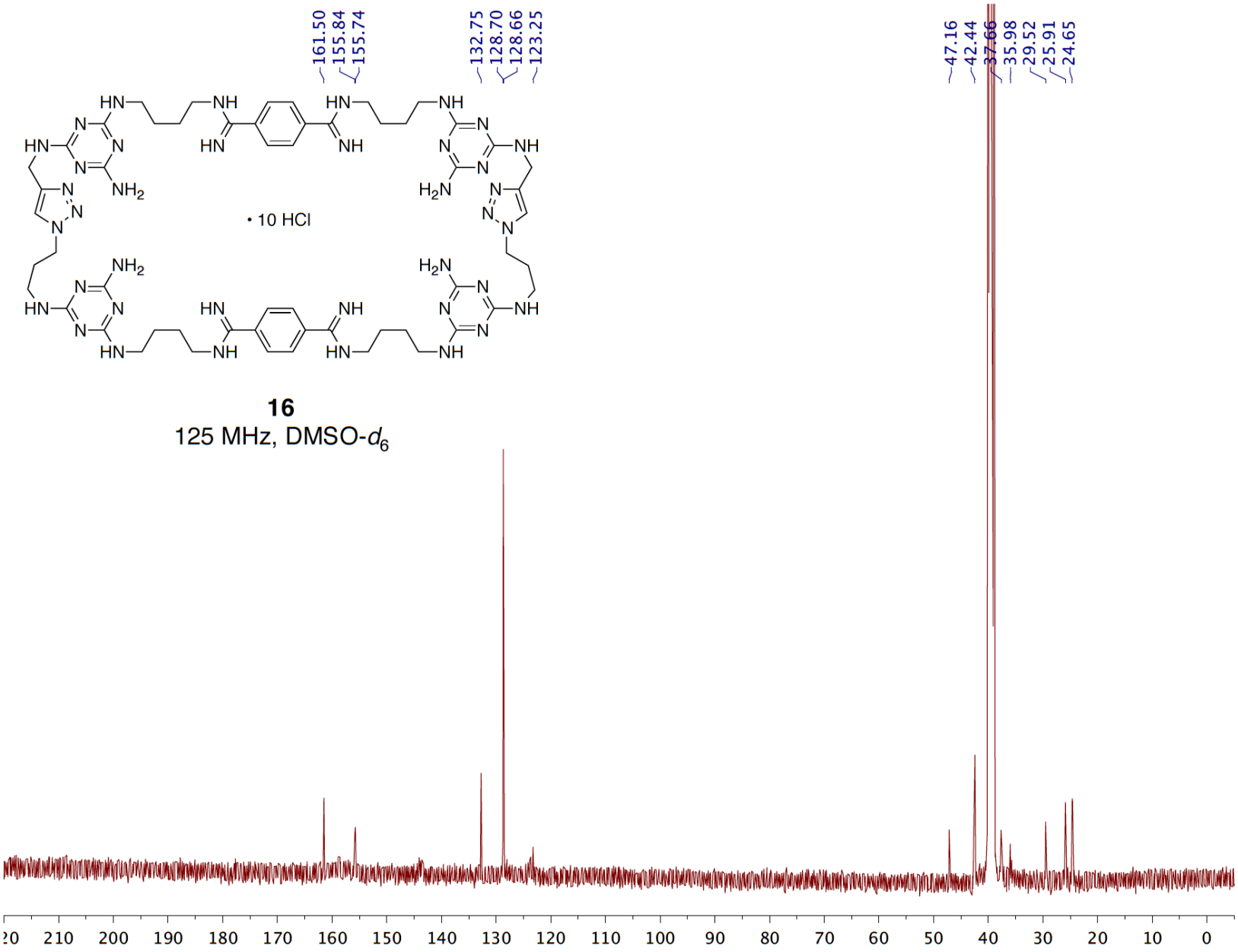
6b

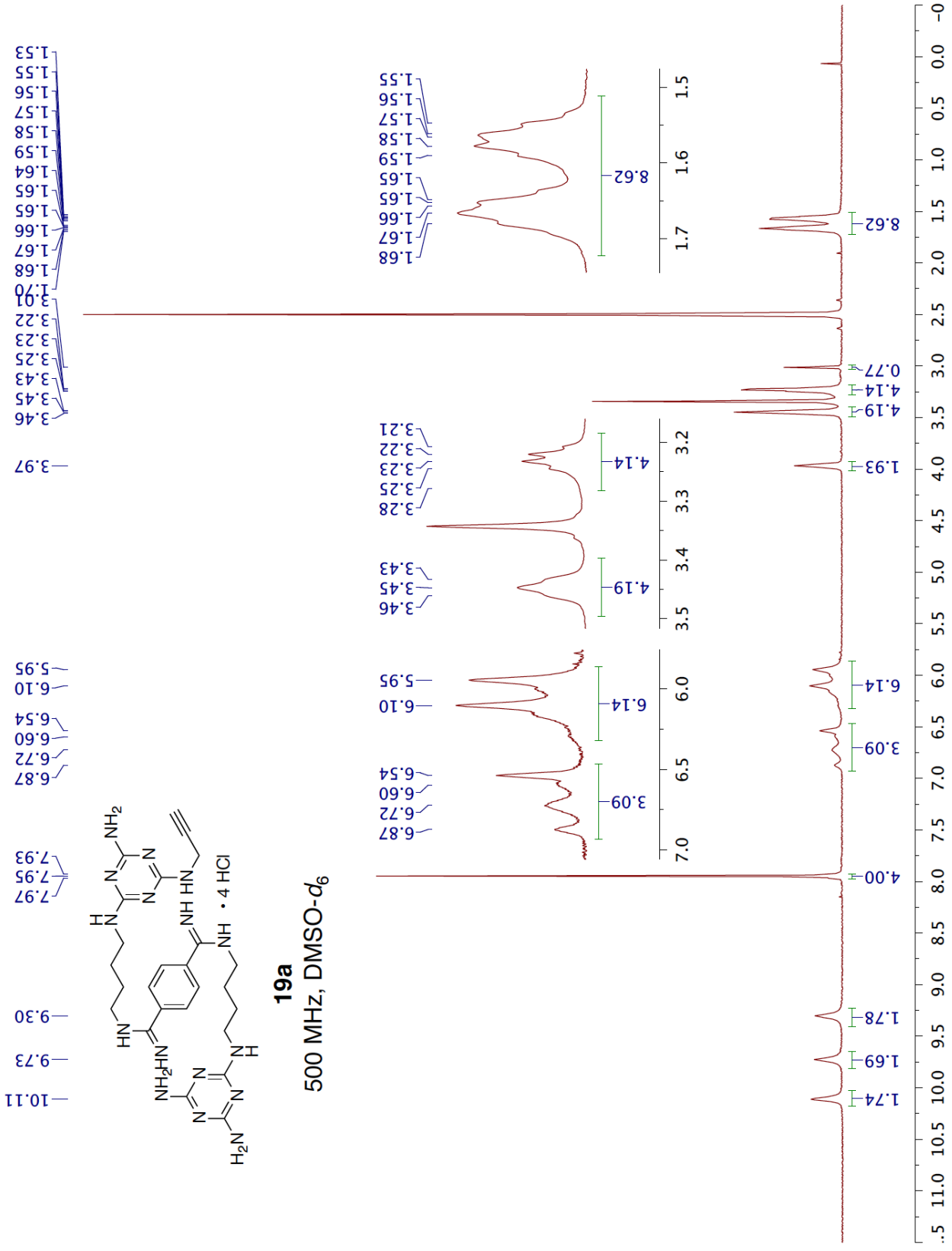
500 MHz, DMSO-d_6

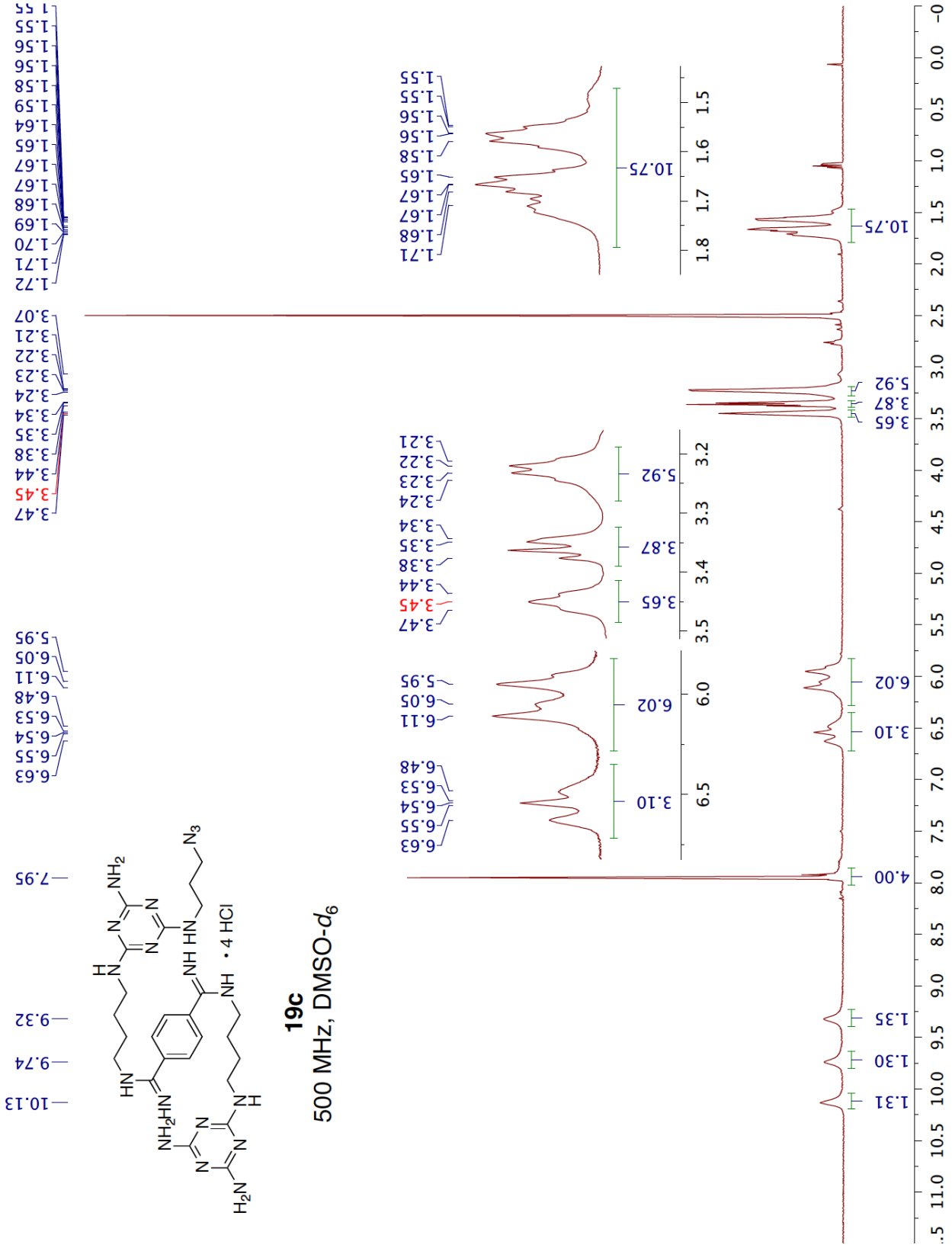

6c
500 MHz, DMSO-d_6
12.00
3.83
11.34
8.07
4.14
2.27
1.99
2.05

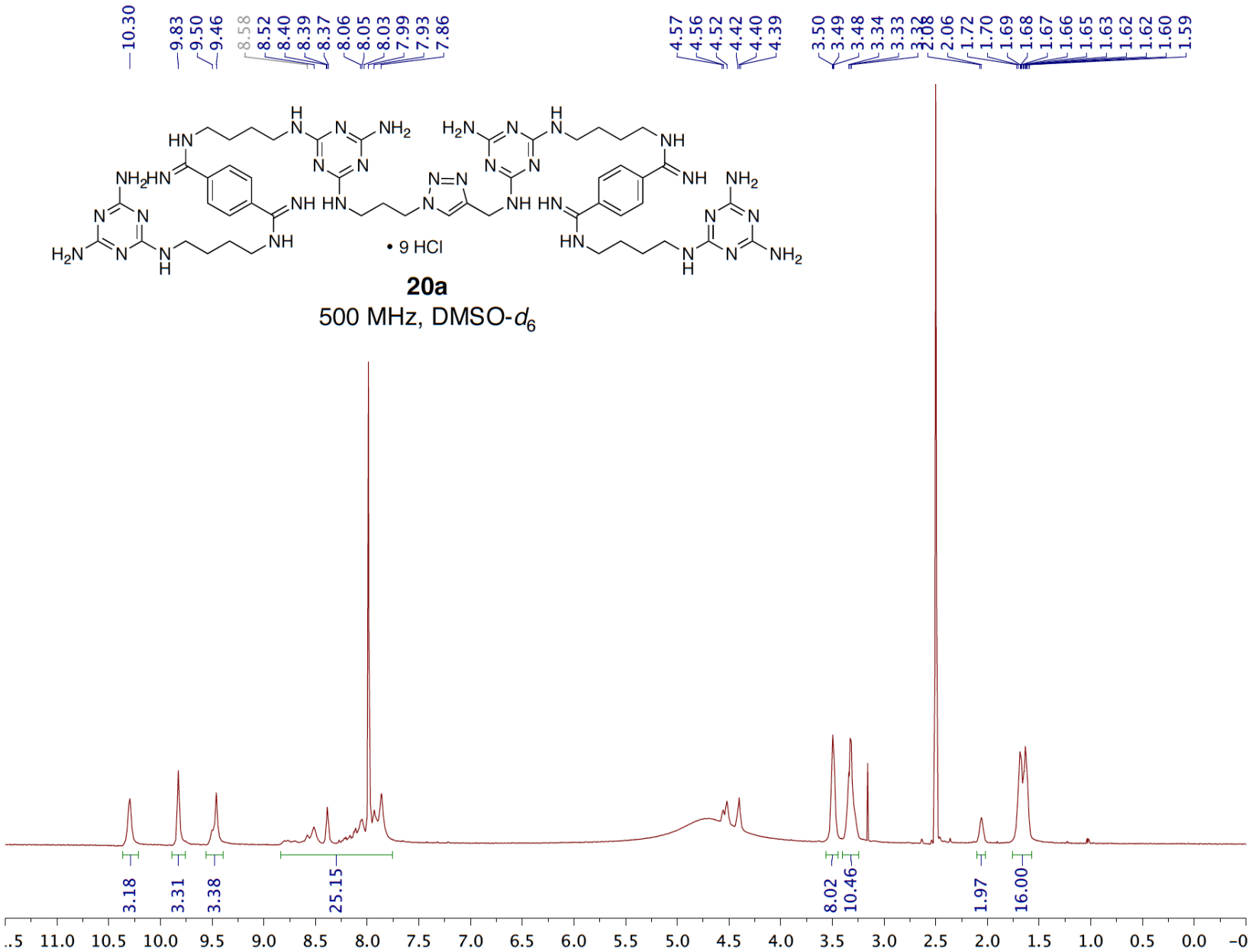

10.10
9.70
9.26
8.42
8.33
8.21
8.17
8.16
8.15
8.08
8.01
7.97
7.93
7.83
7.83
7.79
3.45
3.44
3.43
3.41
3.41
3.40
3.37
3.34
3.34
3.32
2.90
2.70
2.69
2.68
1.69
1.68
1.67
1.66
1.65
1.64
1.63
1.61
1.60
• 10 HCl
15
500 MHz, DMSO-d_6
1.80
1.81
1.87
25.99
11.91
7.87
12.03
8.00
11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0

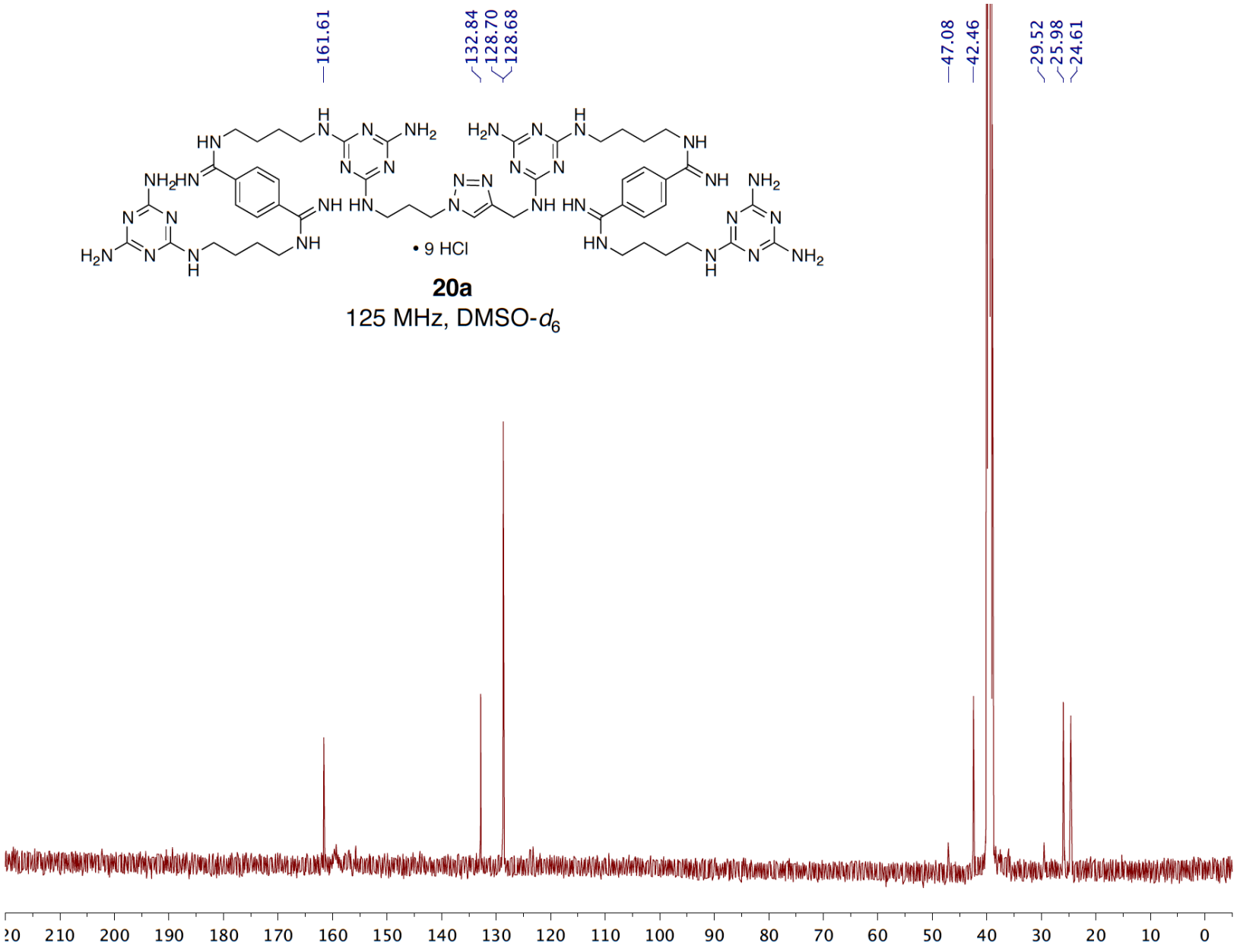

10.33
9.84
9.51
9.49
8.81
8.72
8.59
8.58
8.57
8.50
8.30
8.28
8.28
8.21
8.17
8.13
8.12
8.06
7.99
7.93
4.55
4.53
4.50
4.41
4.39
4.38
3.51
3.50
3.49
3.34
3.29
2.07
2.06
2.04
2.03
1.68
1.64
1.61
1.60
• 10 HCl
16
500 MHz, DMSO-d6
3.10
3.20
3.28
6.49
5.92
8.29
2.00
4.13
4.17
7.90
12.37
3.86
16.00

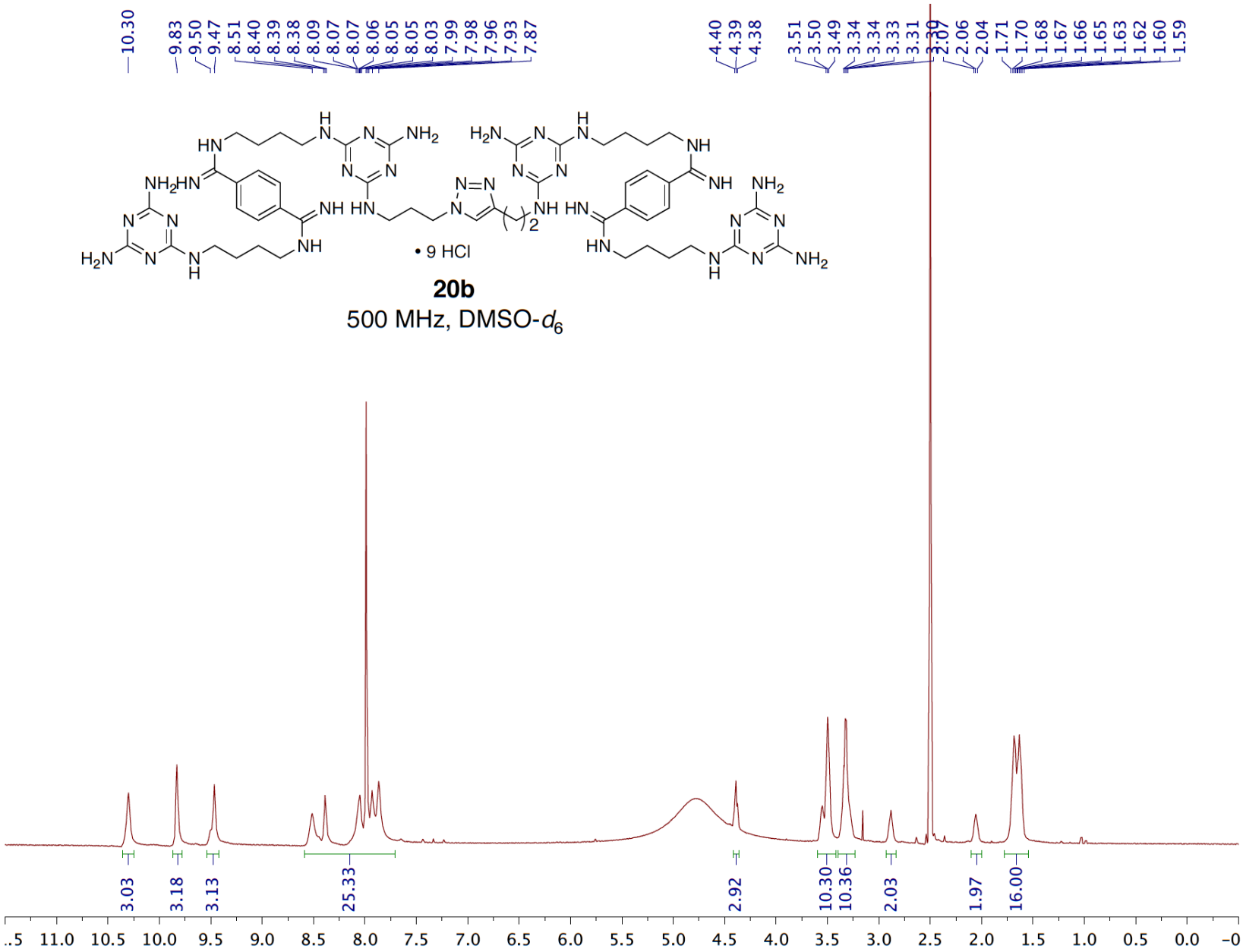

161.50
155.84
155.74
132.75
128.70
128.66
123.25
47.16
42.44
37.66
35.98
29.52
25.91
24.65
• 10 HCl
16
125 MHz, DMSO-d_6
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

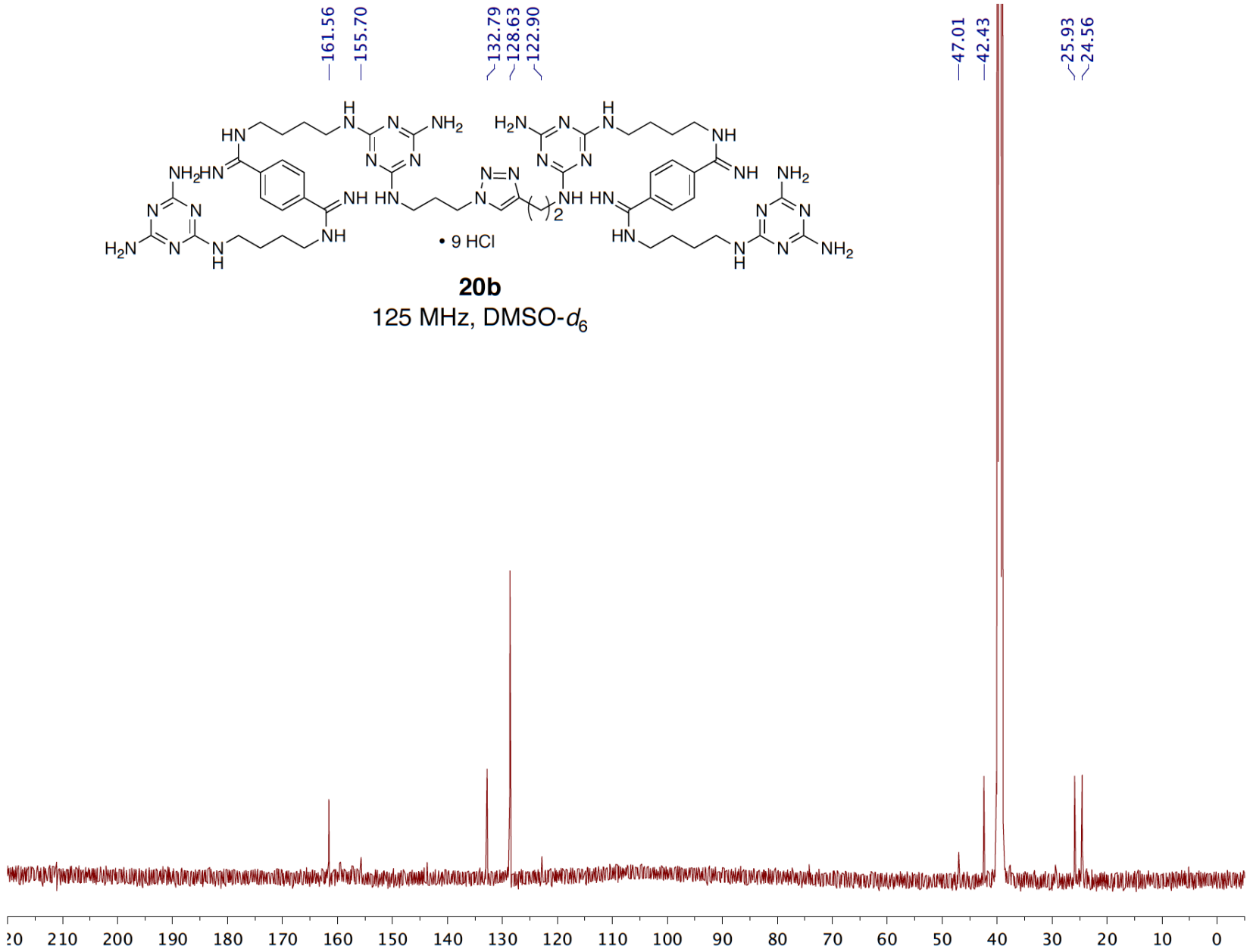

19a
500 MHz, DMSO-d_6
4 HCl

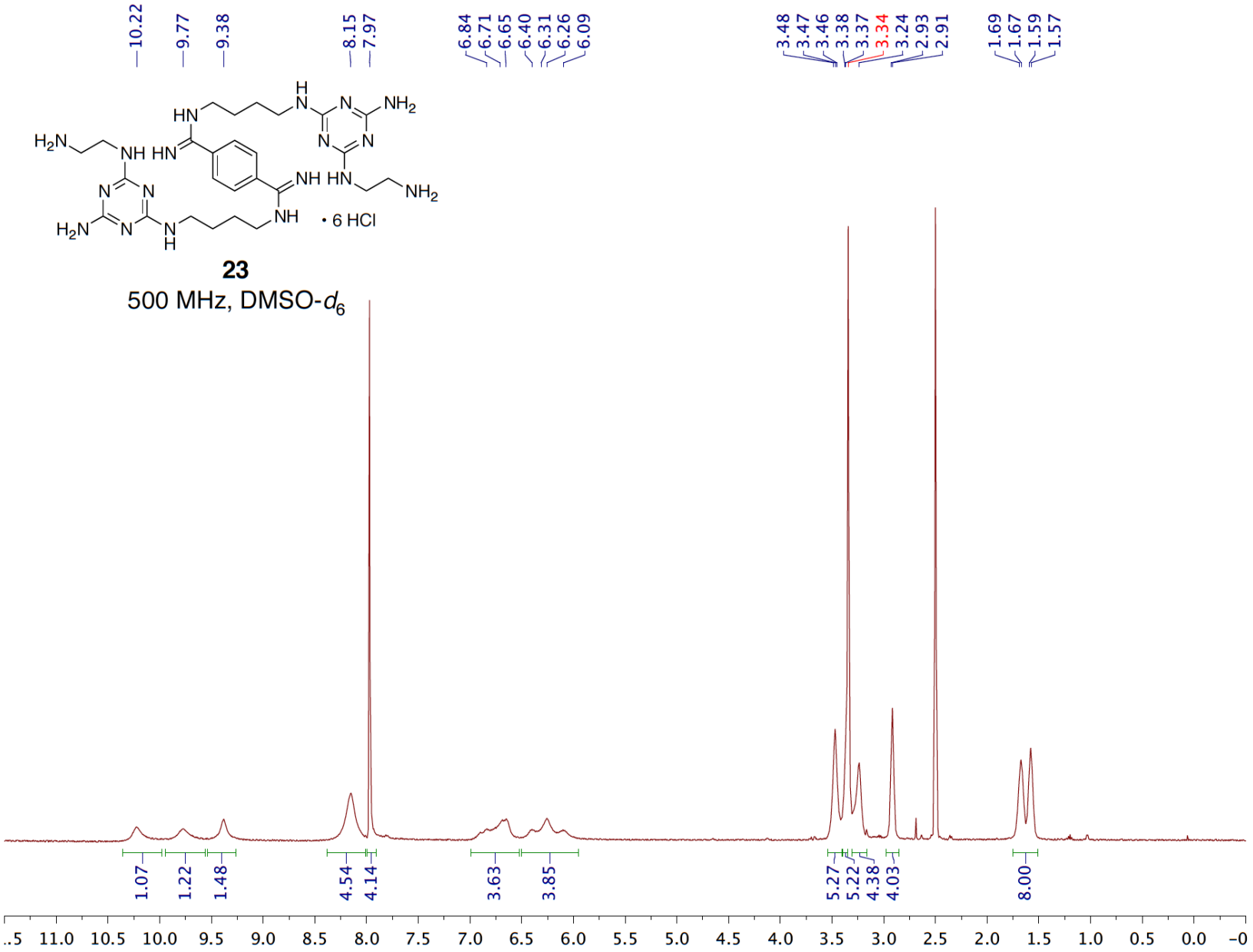

19c
500 MHz, DMSO-d_6
4 HCl

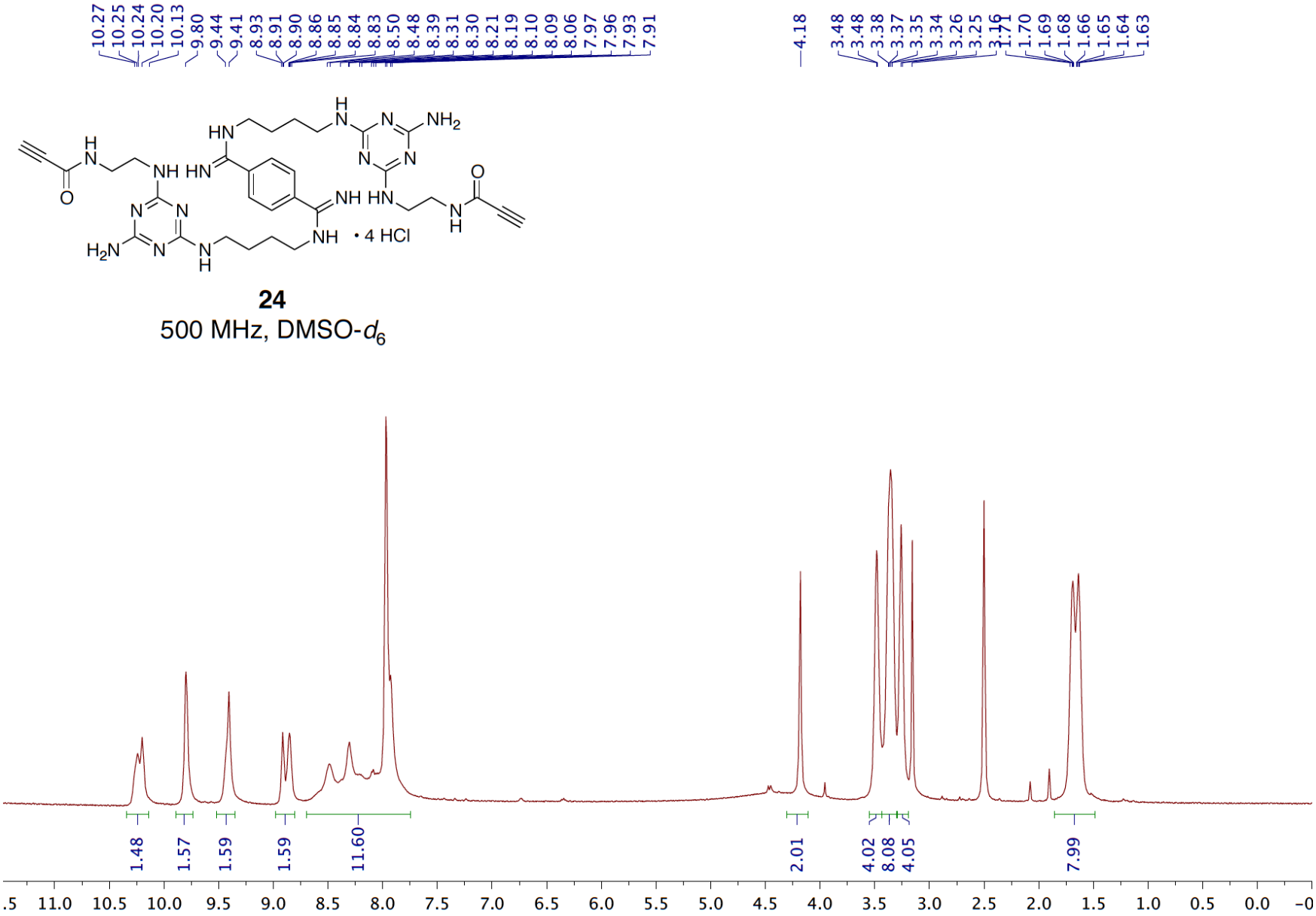

10.30
9.83
9.50
9.46
8.58
8.52
8.40
8.39
8.37
8.06
8.05
8.03
7.99
7.93
7.86
4.57
4.56
4.52
4.42
4.40
4.39
3.50
3.49
3.48
3.34
3.33
2.08
2.06
1.72
1.70
1.69
1.68
1.67
1.66
1.65
1.63
1.62
1.62
1.60
1.59
• 9 HCl
20a
500 MHz, DMSO-d6
3.18
3.31
3.38
25.15
8.02
10.46
1.97
16.00


161.61
132.84
128.70
128.68
47.08
42.46
29.52
25.98
24.61
• 9 HCl
20a
125 MHz, DMSO-d_6
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0


10.30
9.83
9.50
9.47
8.51
8.40
8.39
8.38
8.09
8.07
8.07
8.06
8.05
8.05
8.03
7.99
7.98
7.96
7.93
7.87
4.40
4.39
4.38
3.51
3.50
3.49
3.34
3.34
3.33
3.31
3.30
2.07
2.06
2.04
1.71
1.70
1.68
1.67
1.66
1.65
1.63
1.62
1.60
1.59
• 9 HCl
20b
500 MHz, DMSO-d_6
3.03
3.18
3.13
25.33
2.92
10.30
10.36
2.03
1.97
16.00

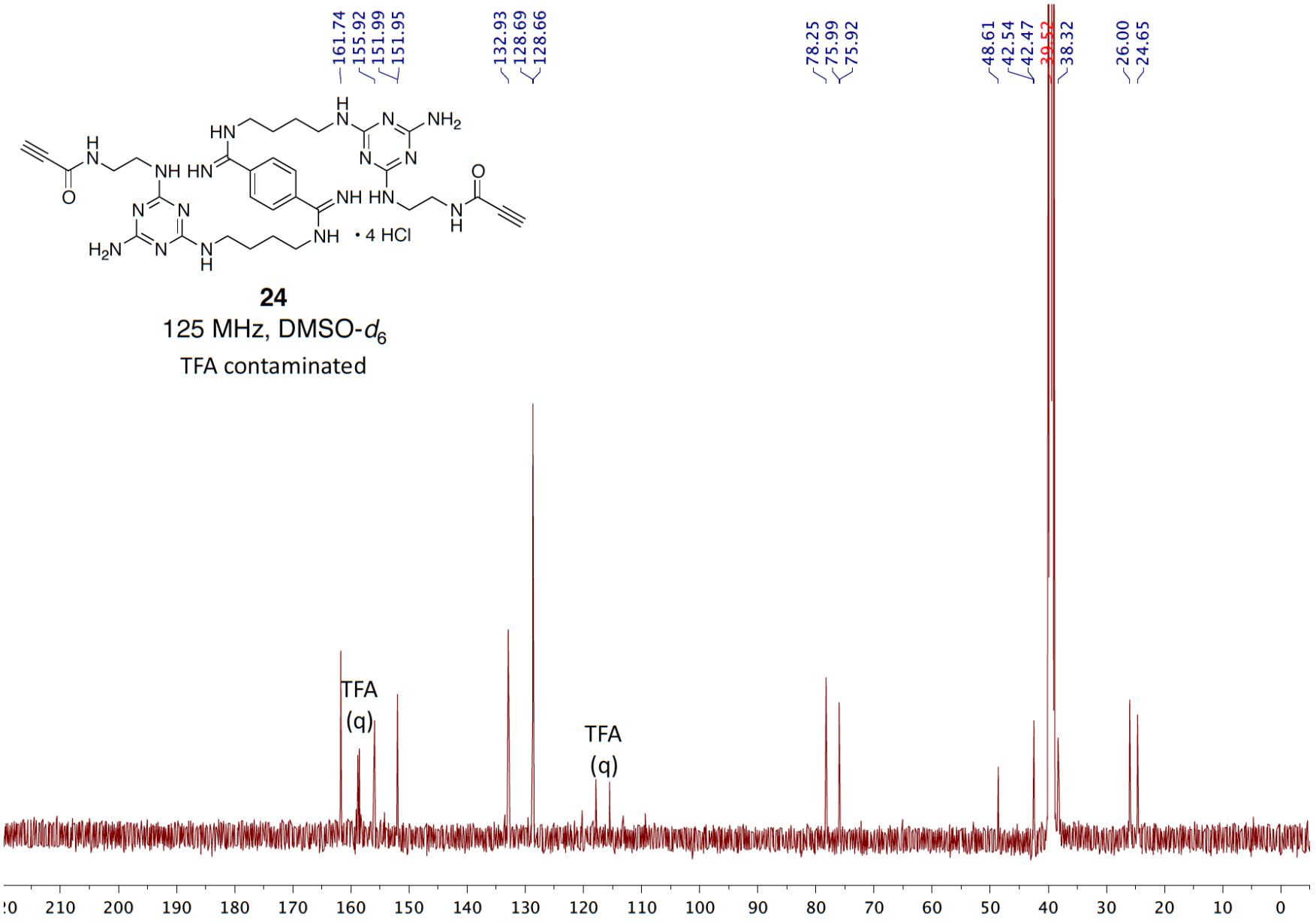

161.56
155.70
132.79
128.63
122.90
47.01
42.43
25.93
24.56
• 9 HCl
20b
125 MHz, DMSO-d_6
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

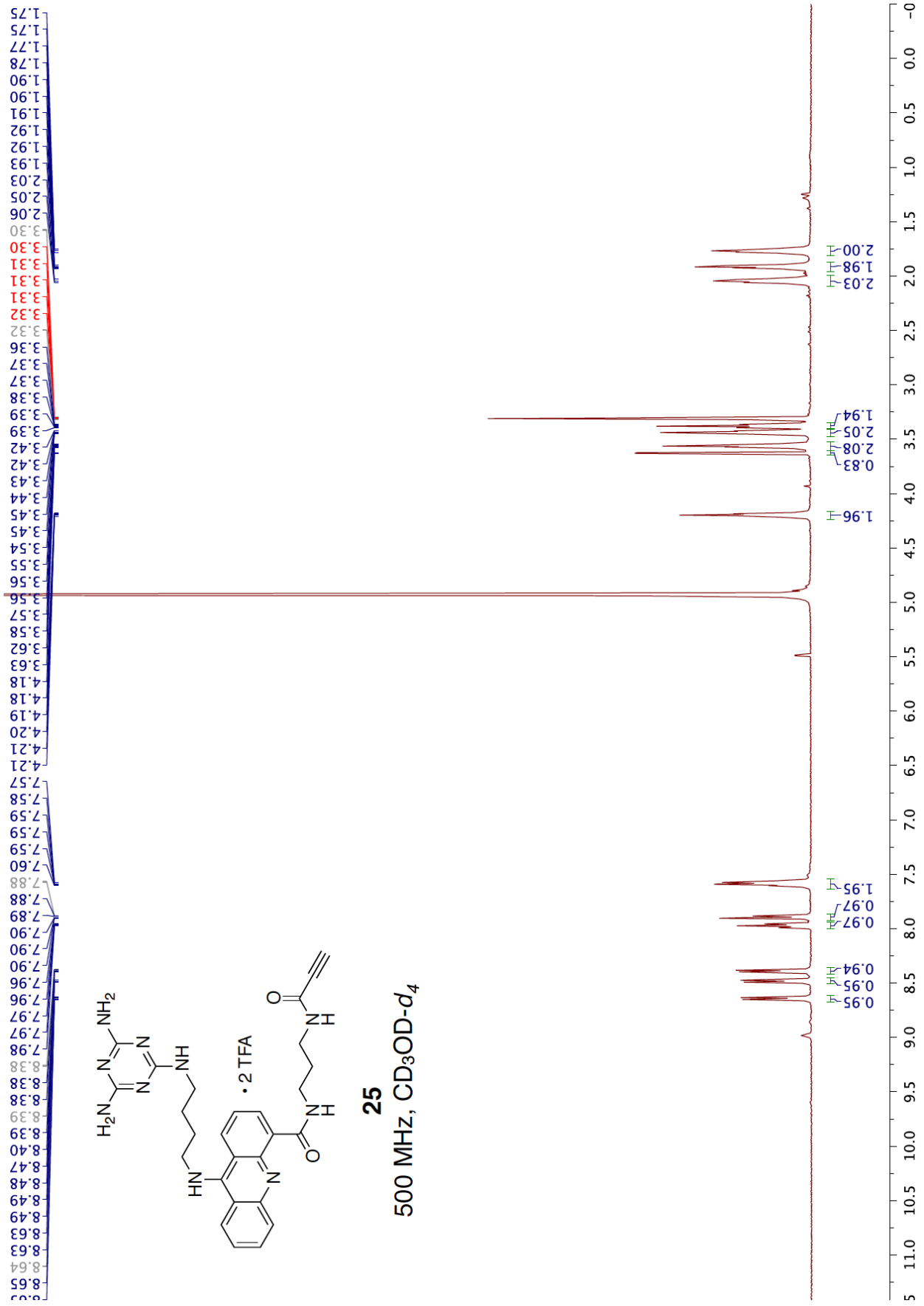

10.22
9.77
9.38
8.15
7.97
6.84
6.71
6.65
6.40
6.31
6.26
6.09
3.48
3.47
3.46
3.38
3.37
3.34
3.24
2.93
2.91
1.69
1.67
1.59
1.57
23
500 MHz, DMSO-d6
1.07
1.22
1.48
4.54
4.14
3.63
3.85
5.27
5.22
4.38
4.03
8.00

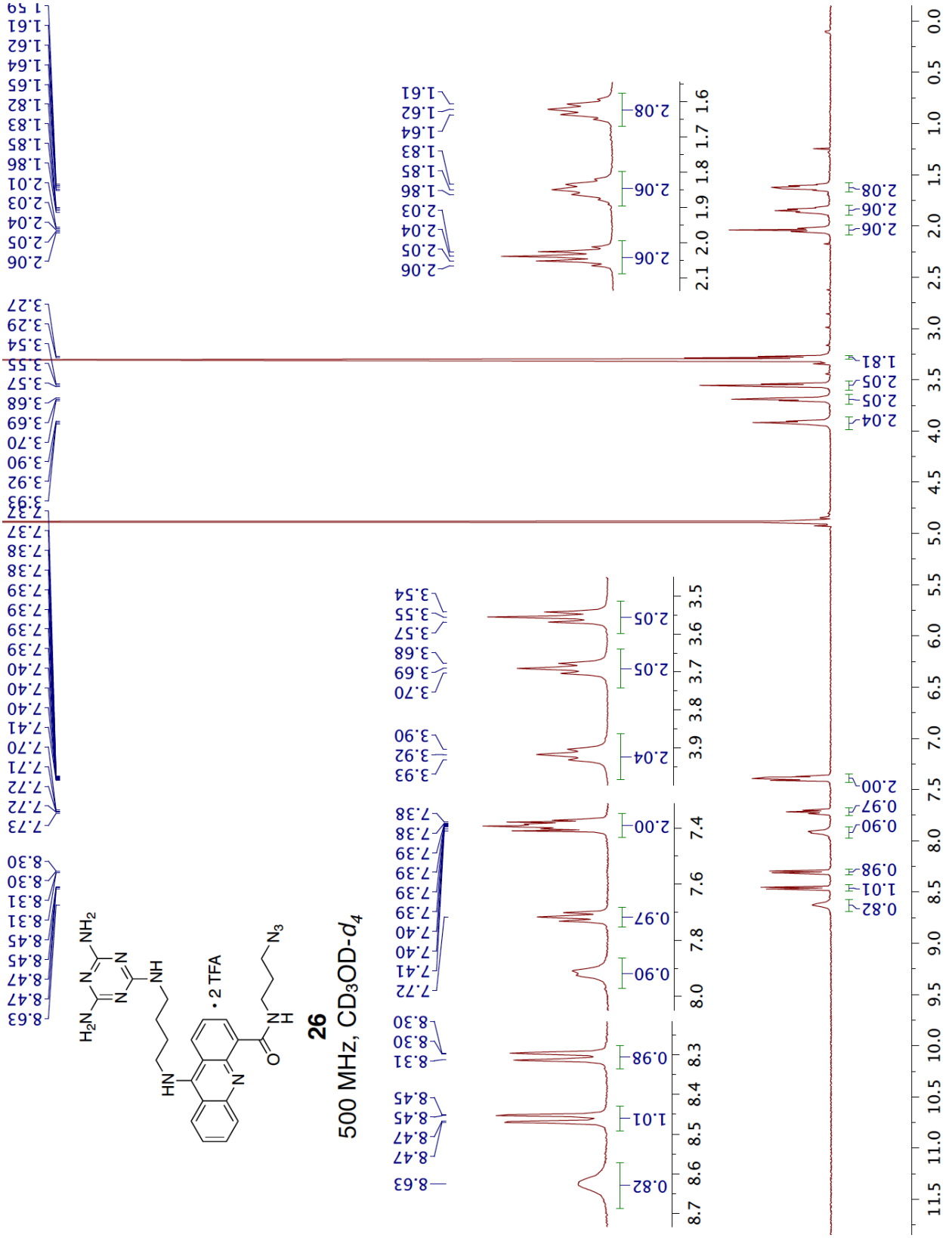


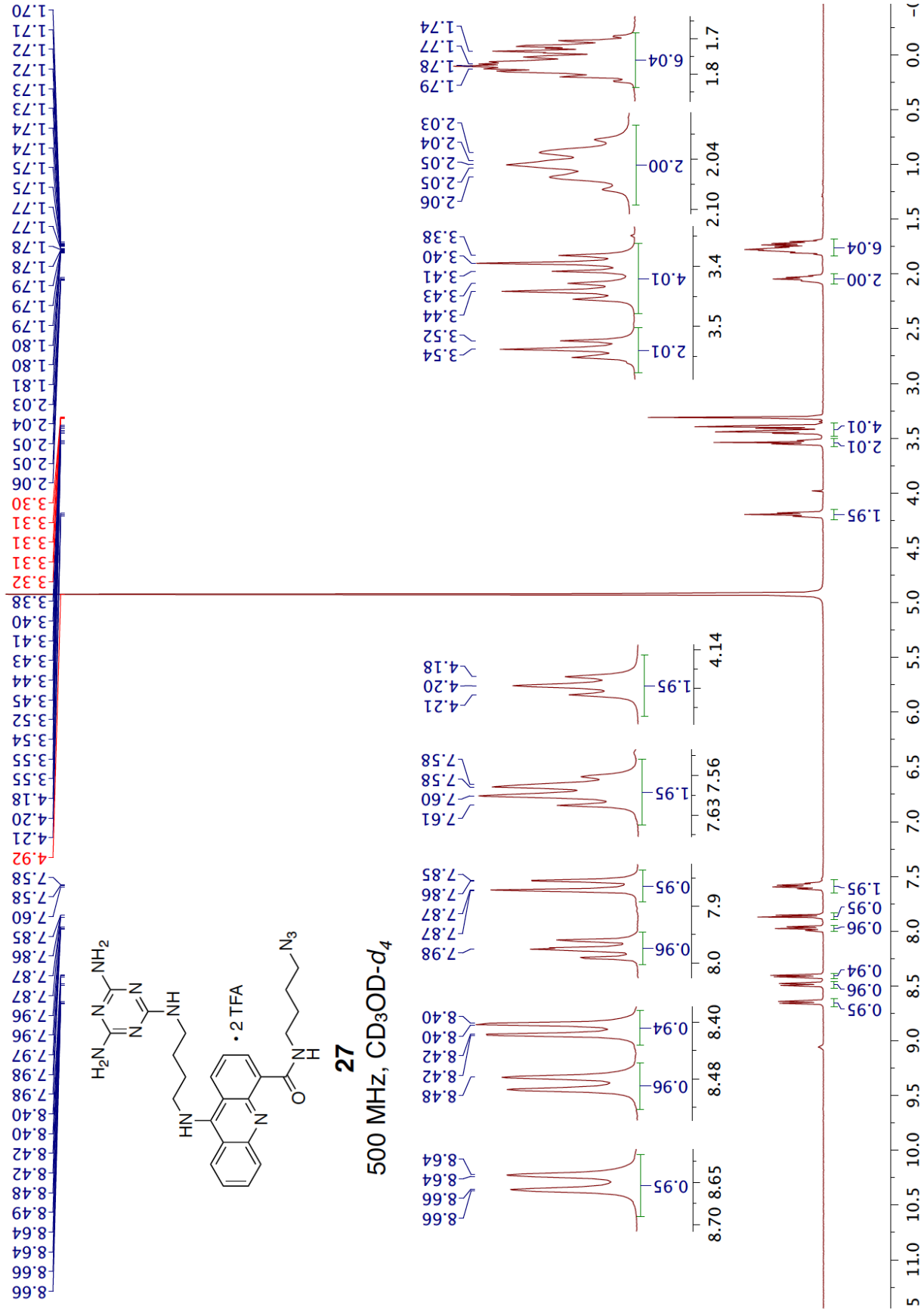
24

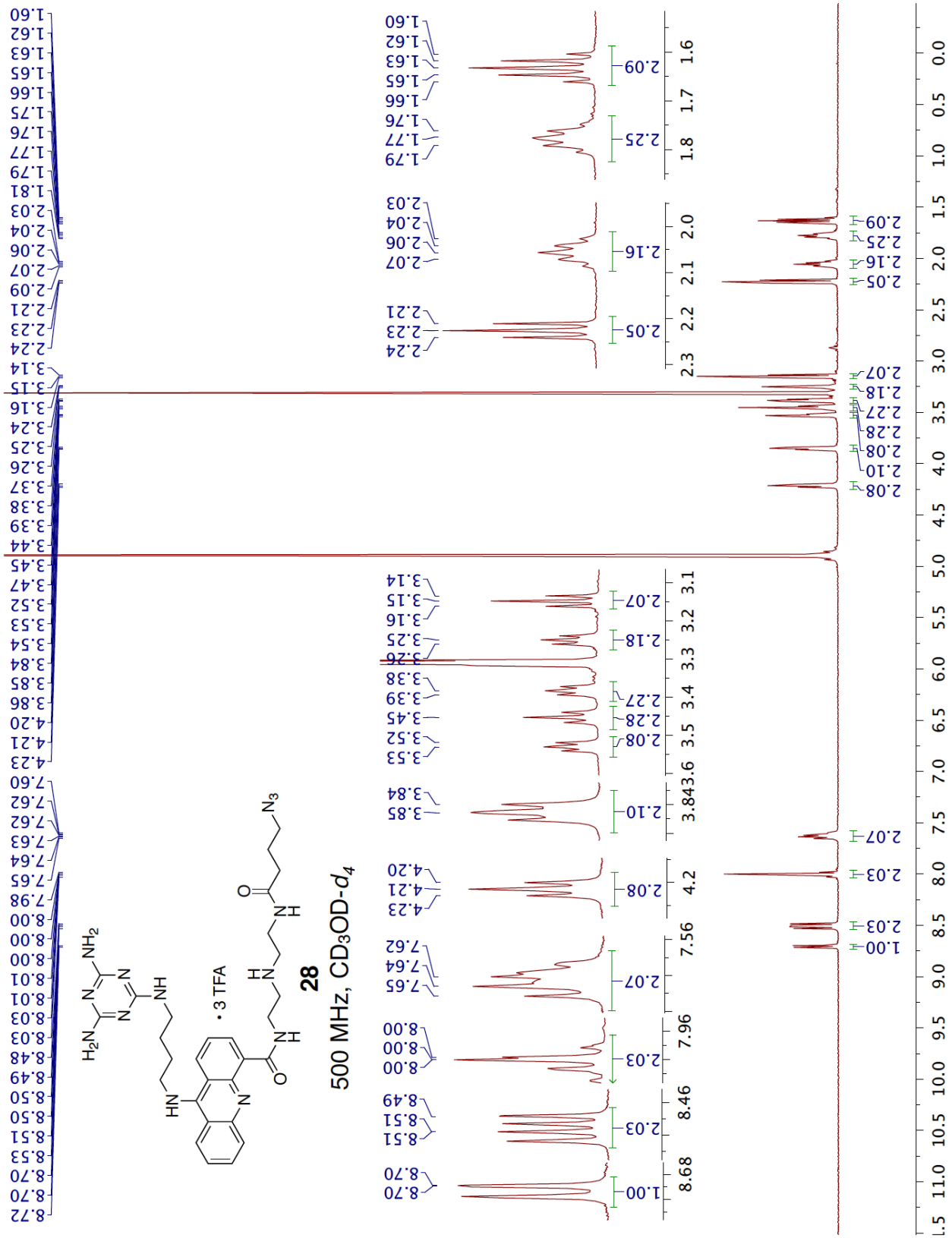
500 MHz, DMSO-d_6

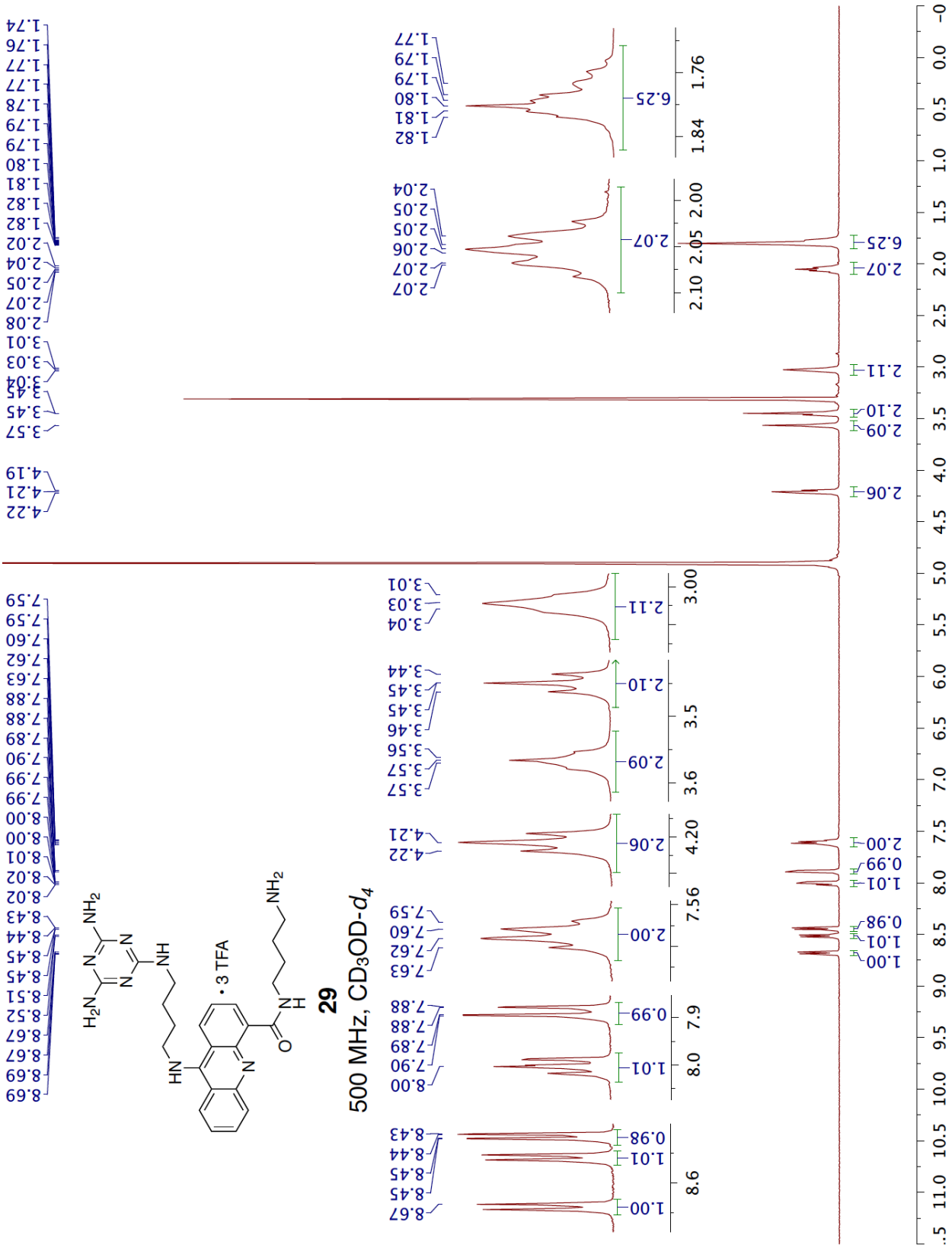

161.74
155.92
151.99
151.95
132.93
128.69
128.66
78.25
75.99
75.92
48.61
42.54
42.47
39.52
38.32
26.00
24.65
• 4 HCl
24
125 MHz, DMSO-d_6
TFA contaminated
TFA
(q)
TFA
(q)
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

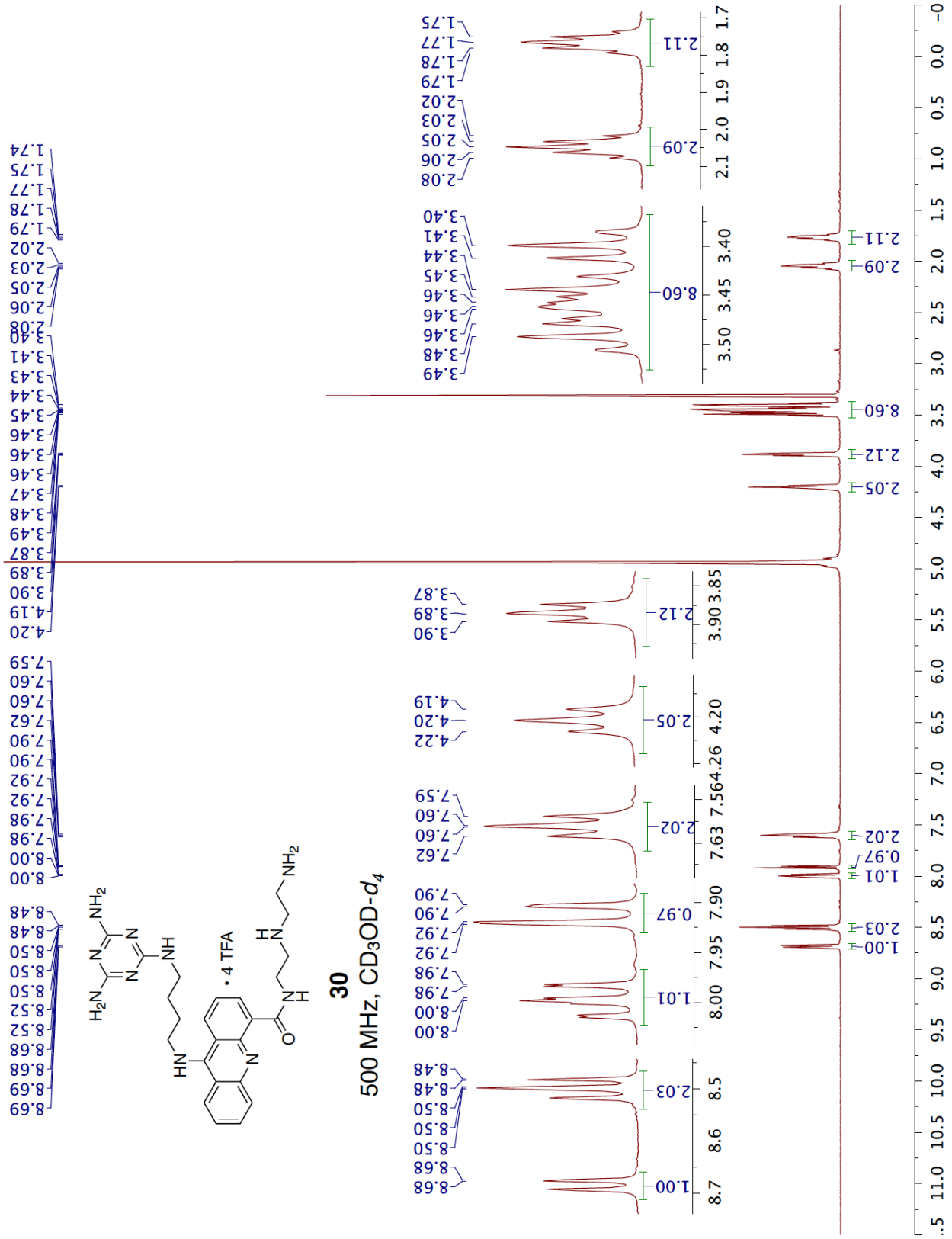

25
500 MHz, CD3OD-d4
· 2 TFA

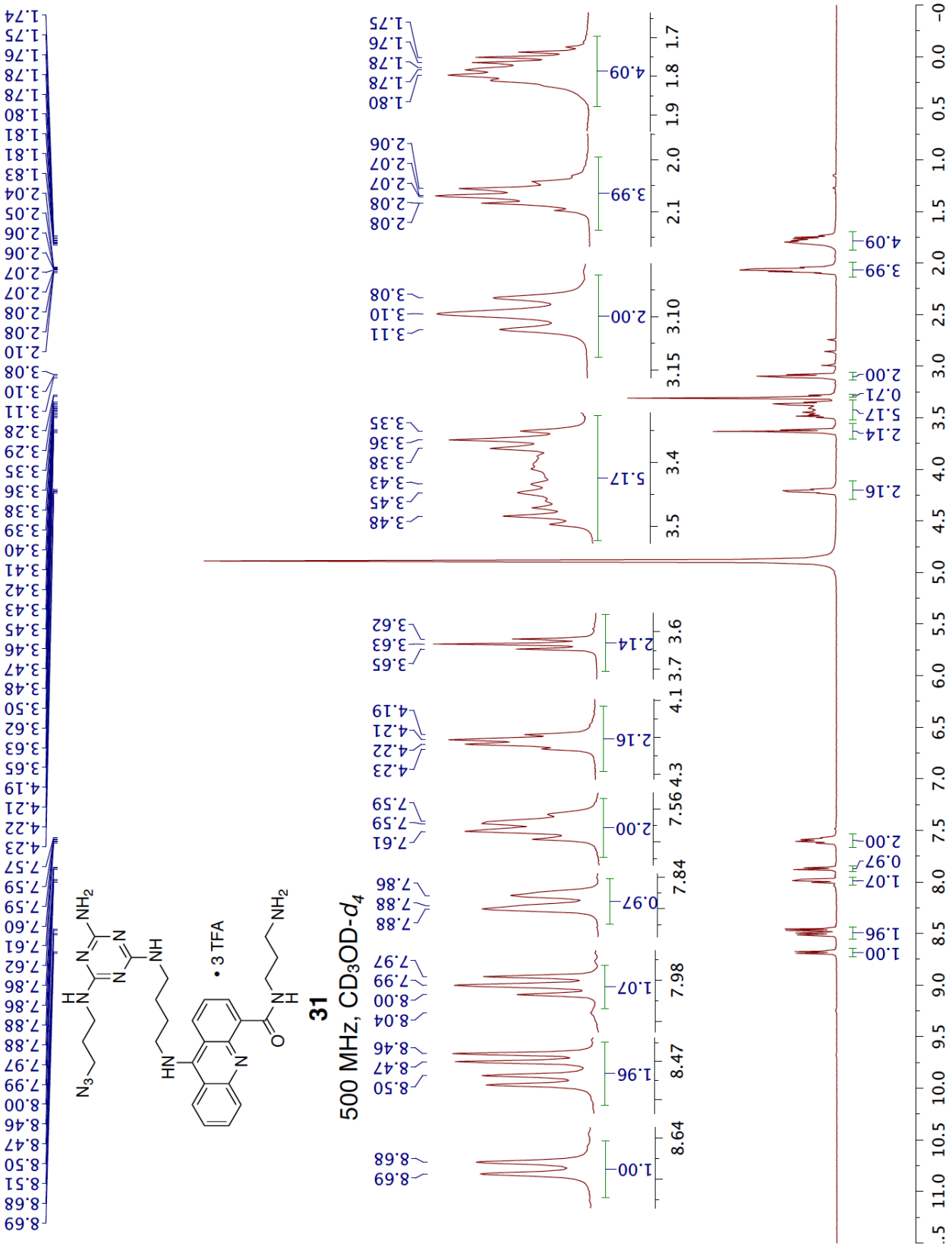

26
500 MHz, CD3OD-d4
• 2 TFA

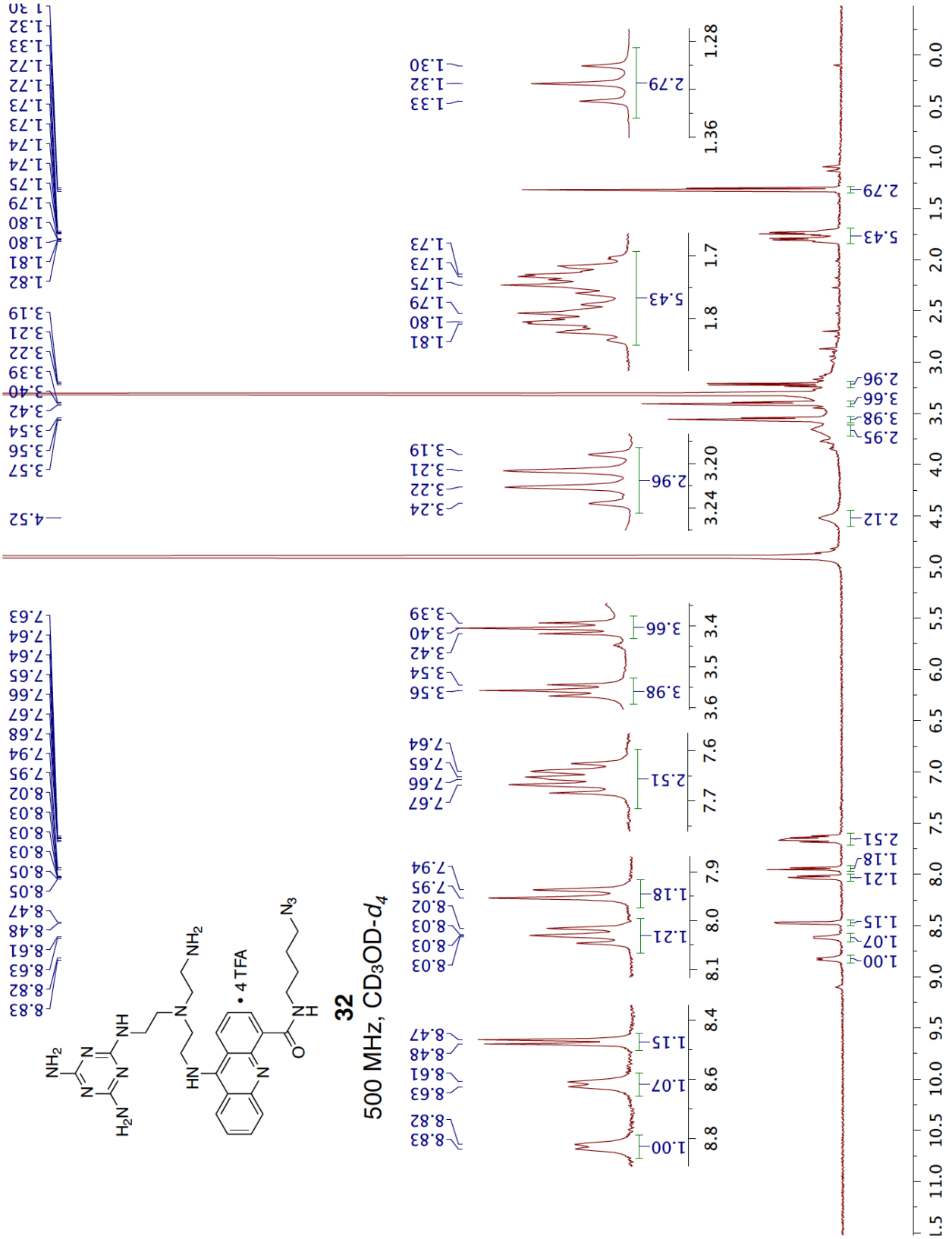

27
500 MHz, CD3OD-d4
• 2 TFA

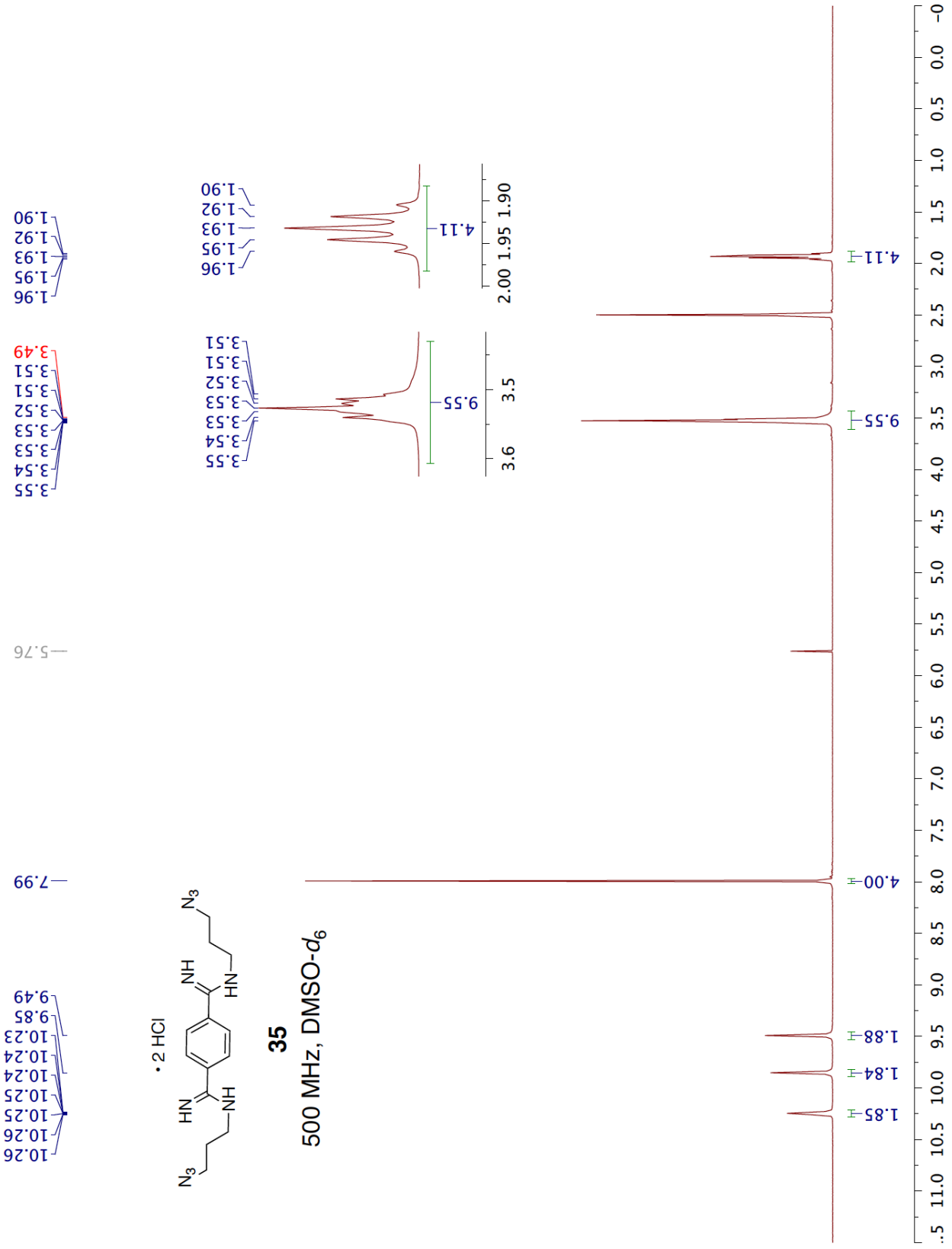

28
500 MHz, CD3OD-d4
• 3 TFA

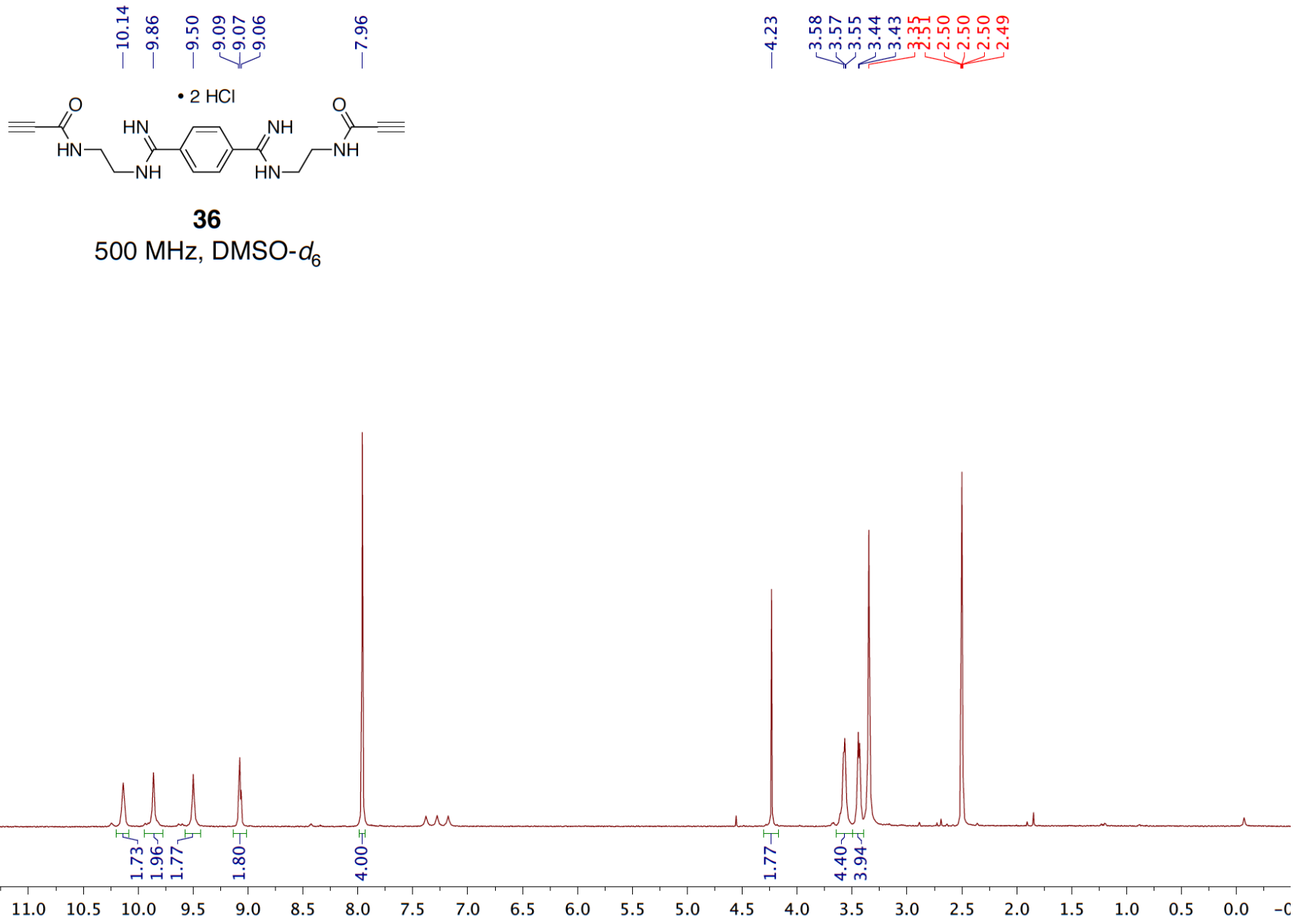

29
500 MHz, CD3OD-d4
• 3 TFA


30
500 MHz, CD3OD-d4
• 4 TFA


31
500 MHz, CD3OD-d4
• 3 TFA

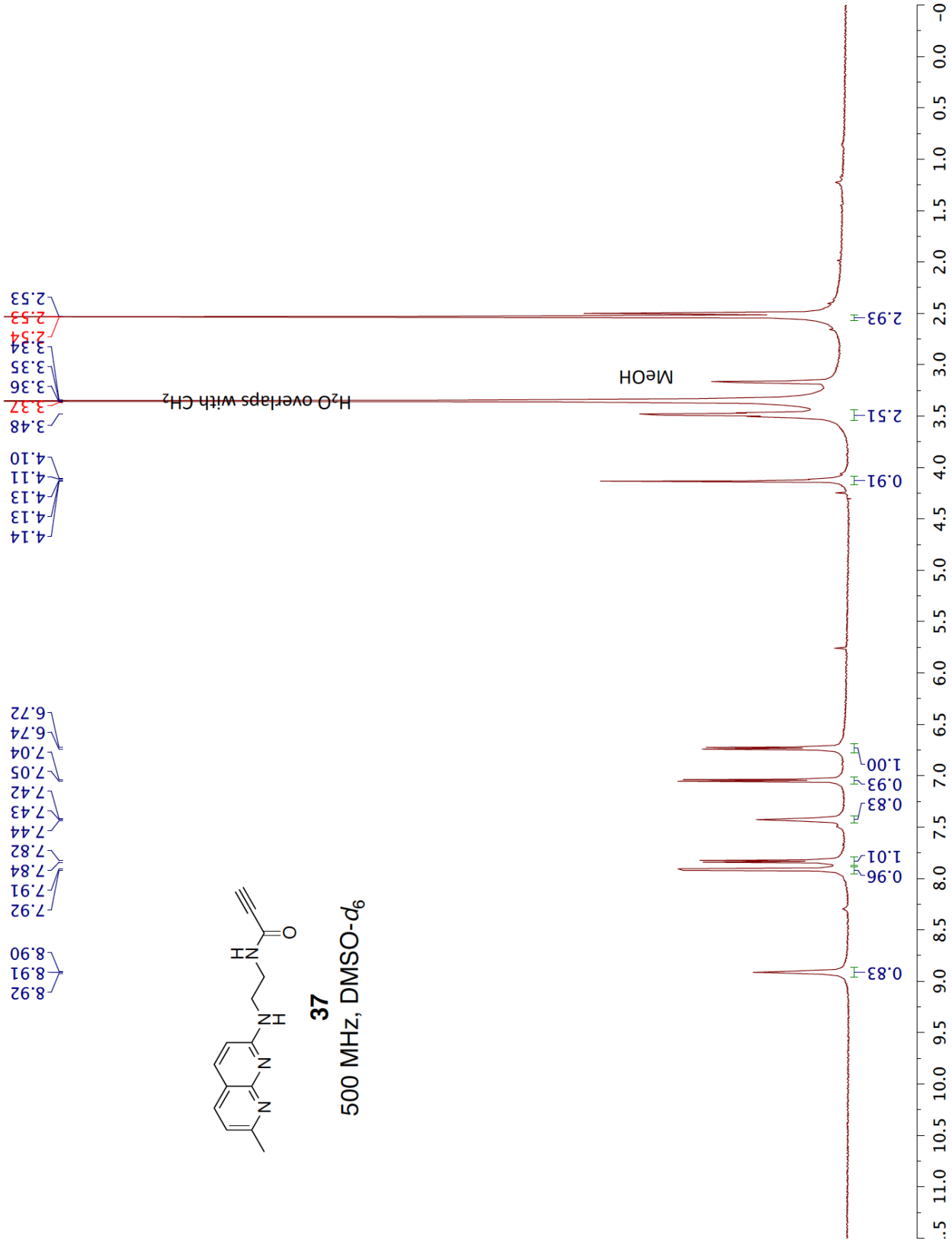

32
500 MHz, CD3OD-d4
• 4 TFA

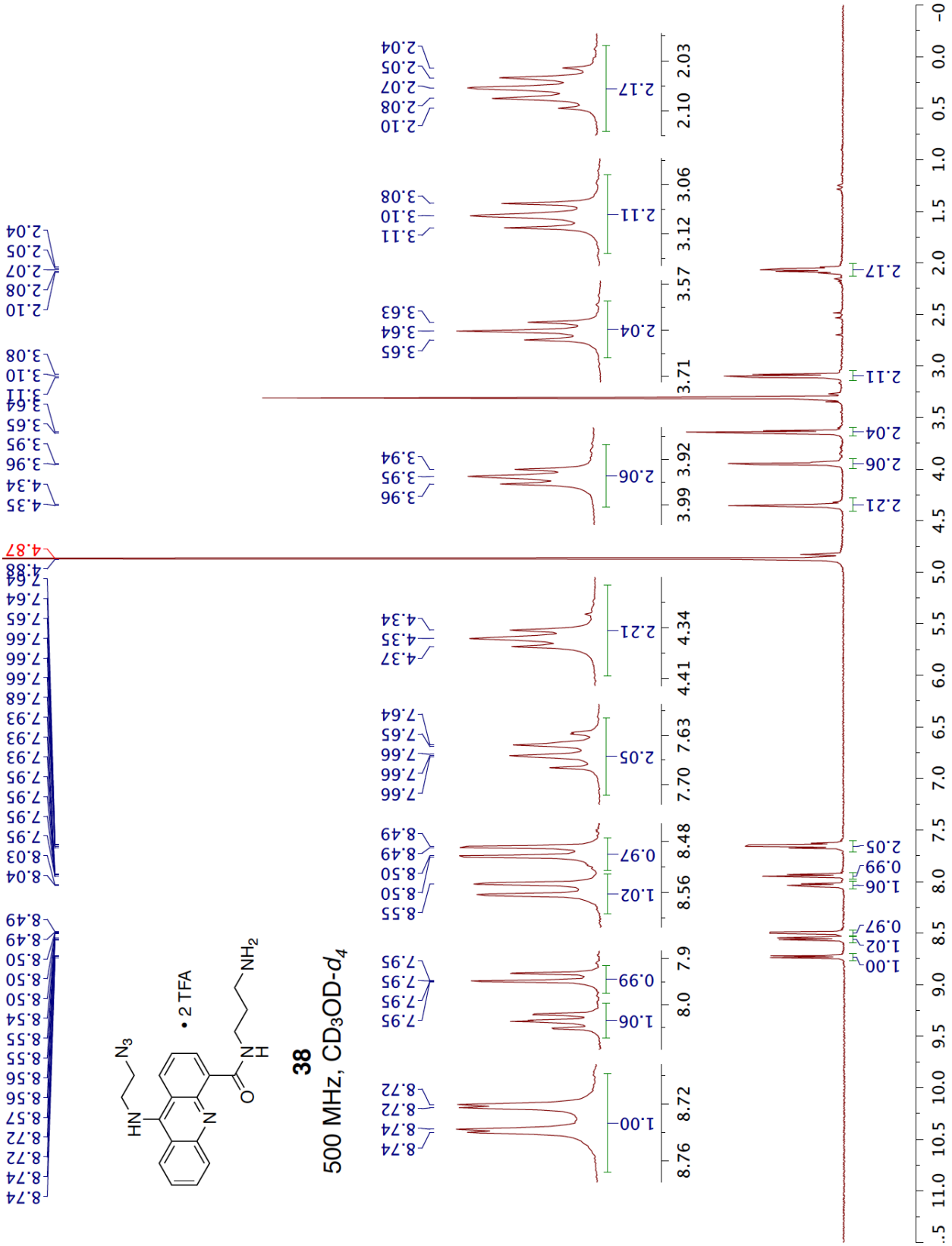

35
500 MHz, DMSO-d_6
• 2 HCl
N_3
NH
HN
NH
HN
N_3
10.26
10.26
10.25
10.25
10.25
10.24
10.24
10.23
9.85
9.49
7.99
5.76
3.55
3.54
3.53
3.53
3.52
3.53
3.53
3.51
3.51
3.49
1.96
1.95
1.93
1.92
1.90
4.00
9.55
4.11
1.88
1.84
1.85
3.55
3.54
3.53
3.53
3.52
3.53
3.53
3.51
3.51
9.55
3.6
3.5
1.96
1.95
1.93
1.92
1.90
4.11
2.00
1.95
1.90

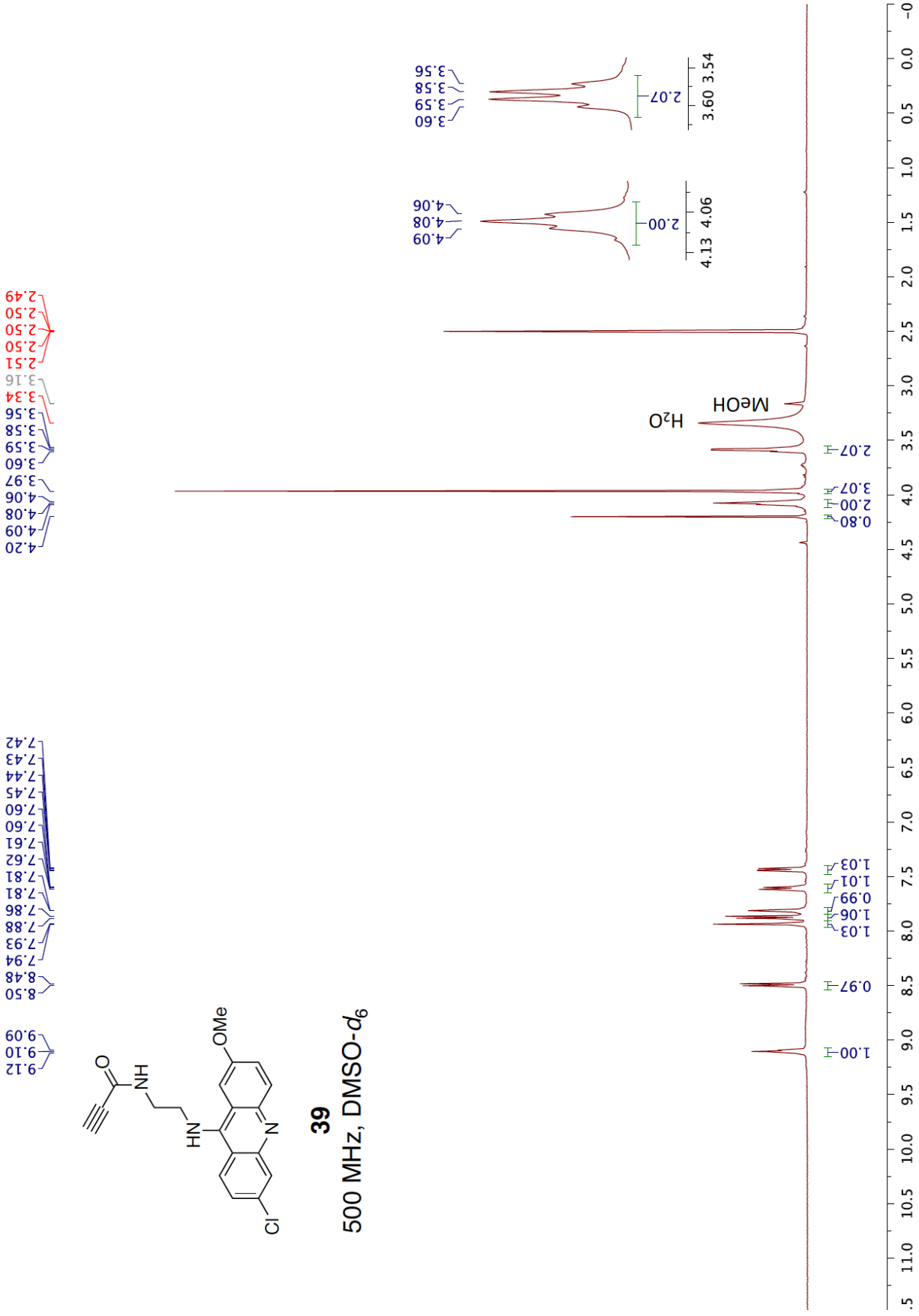


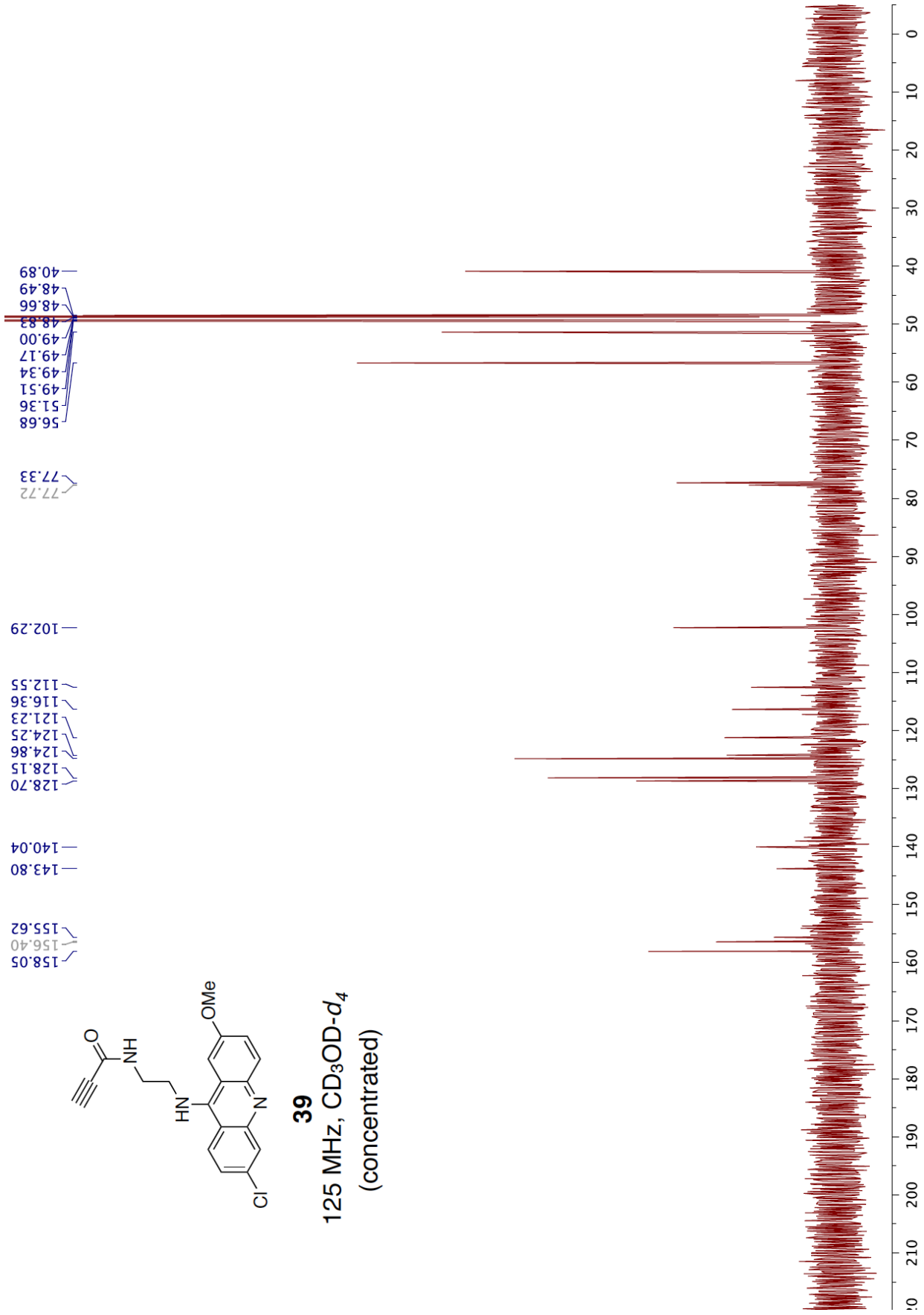
36

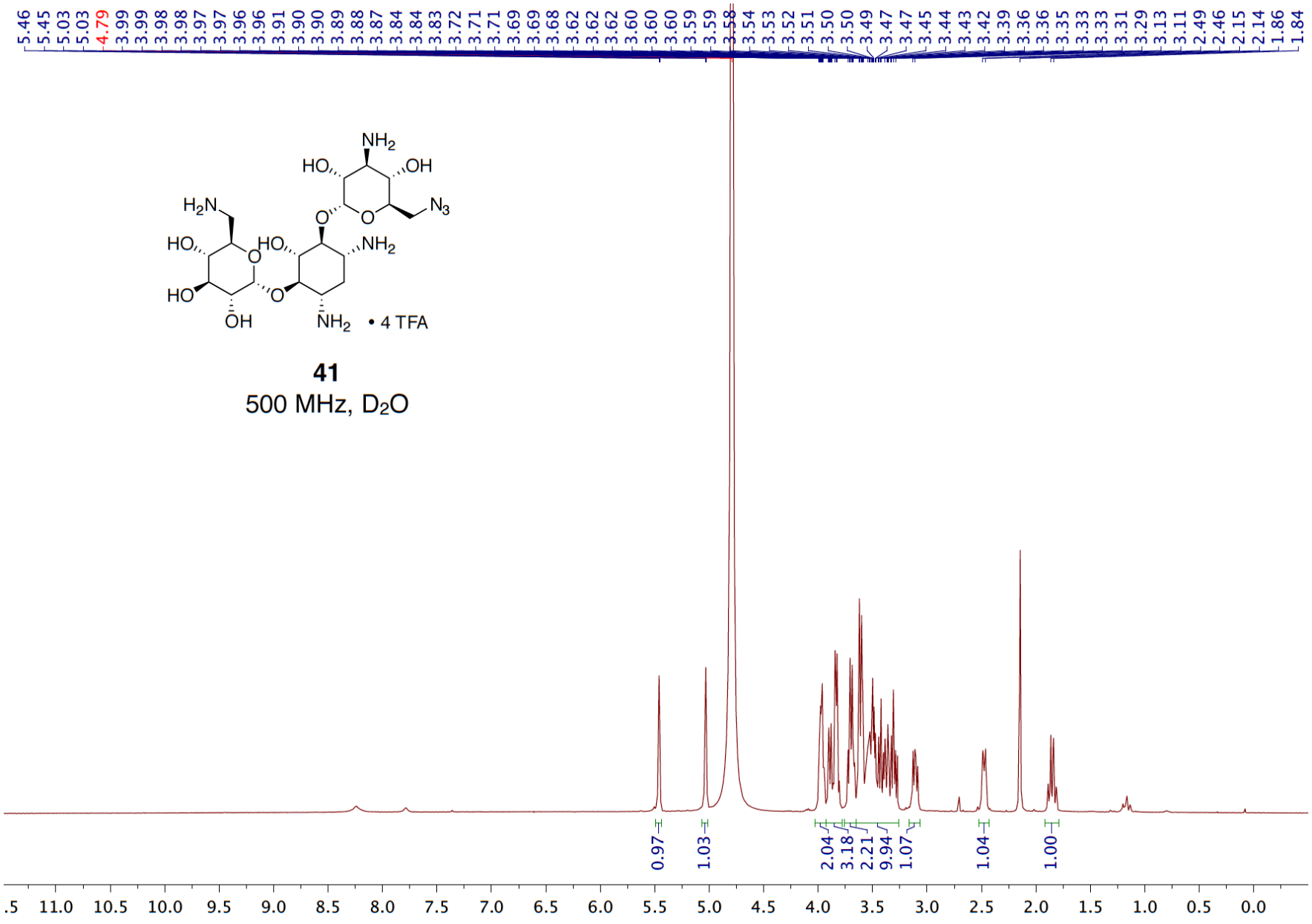
500 MHz, DMSO-d_6

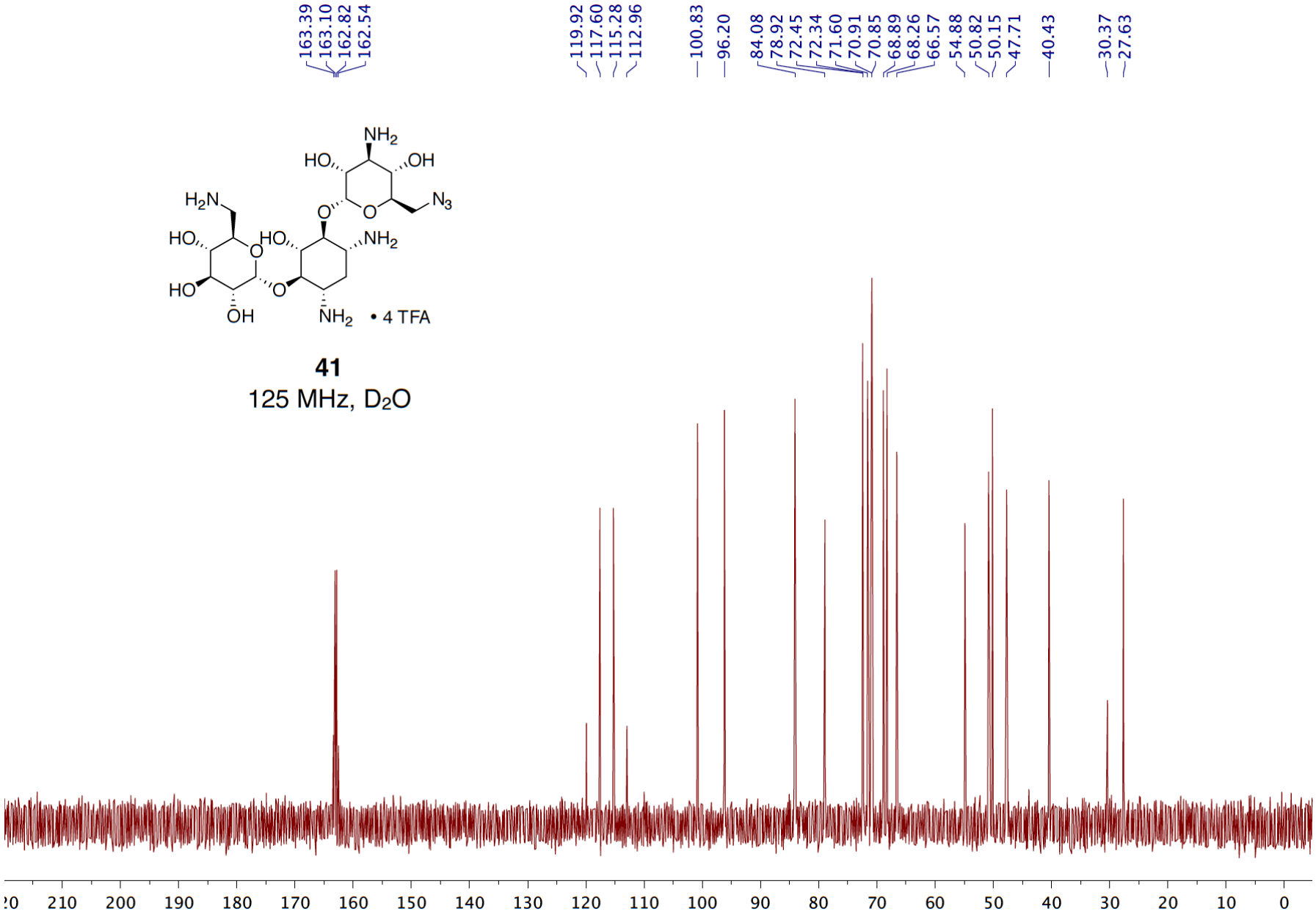

37
500 MHz, DMSO-d_6
MeOH
H_2O overlaps with CH_2

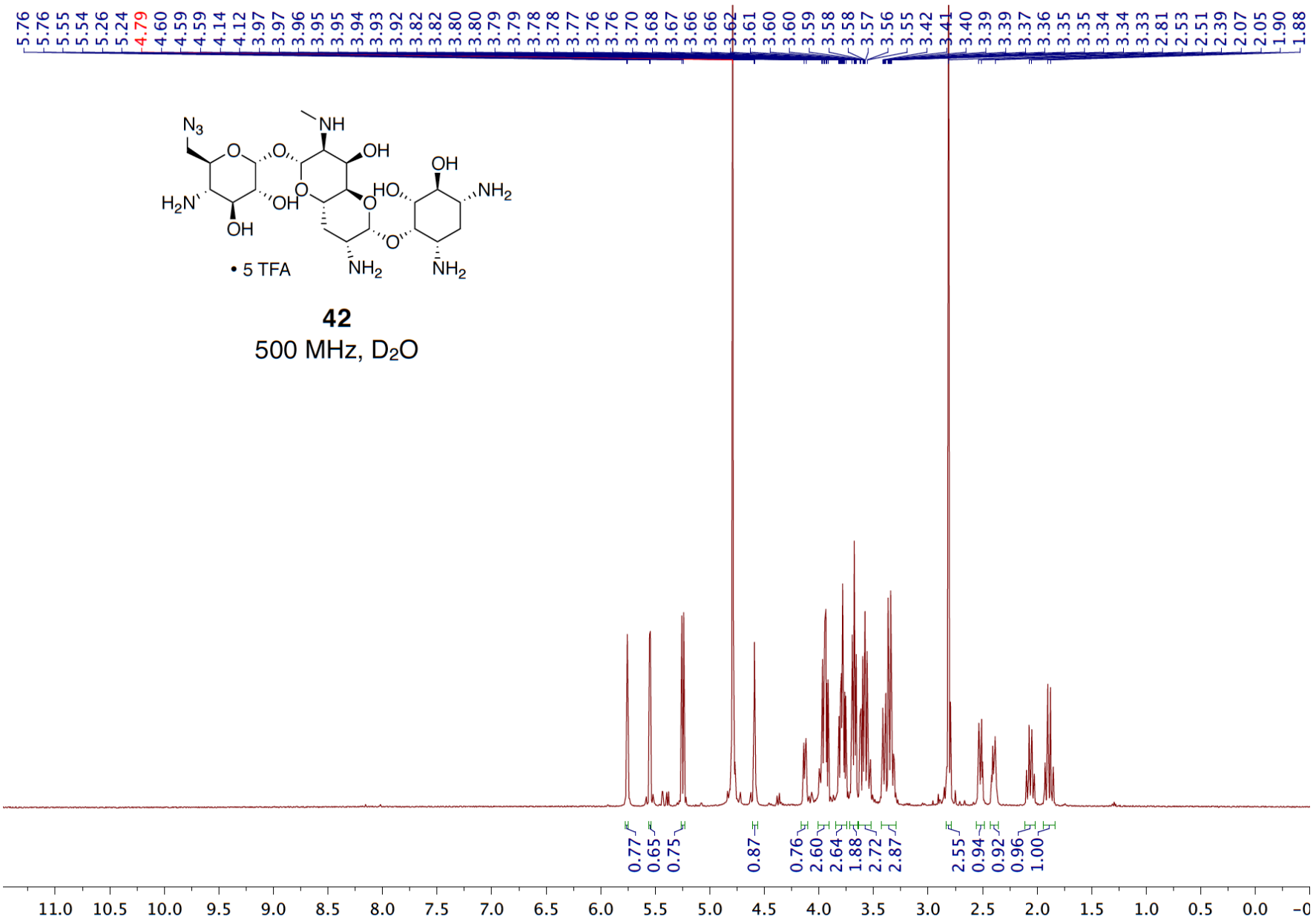

38
500 MHz, CD3OD-d4
• 2 TFA
HN
N3
N
O
NH
NH2

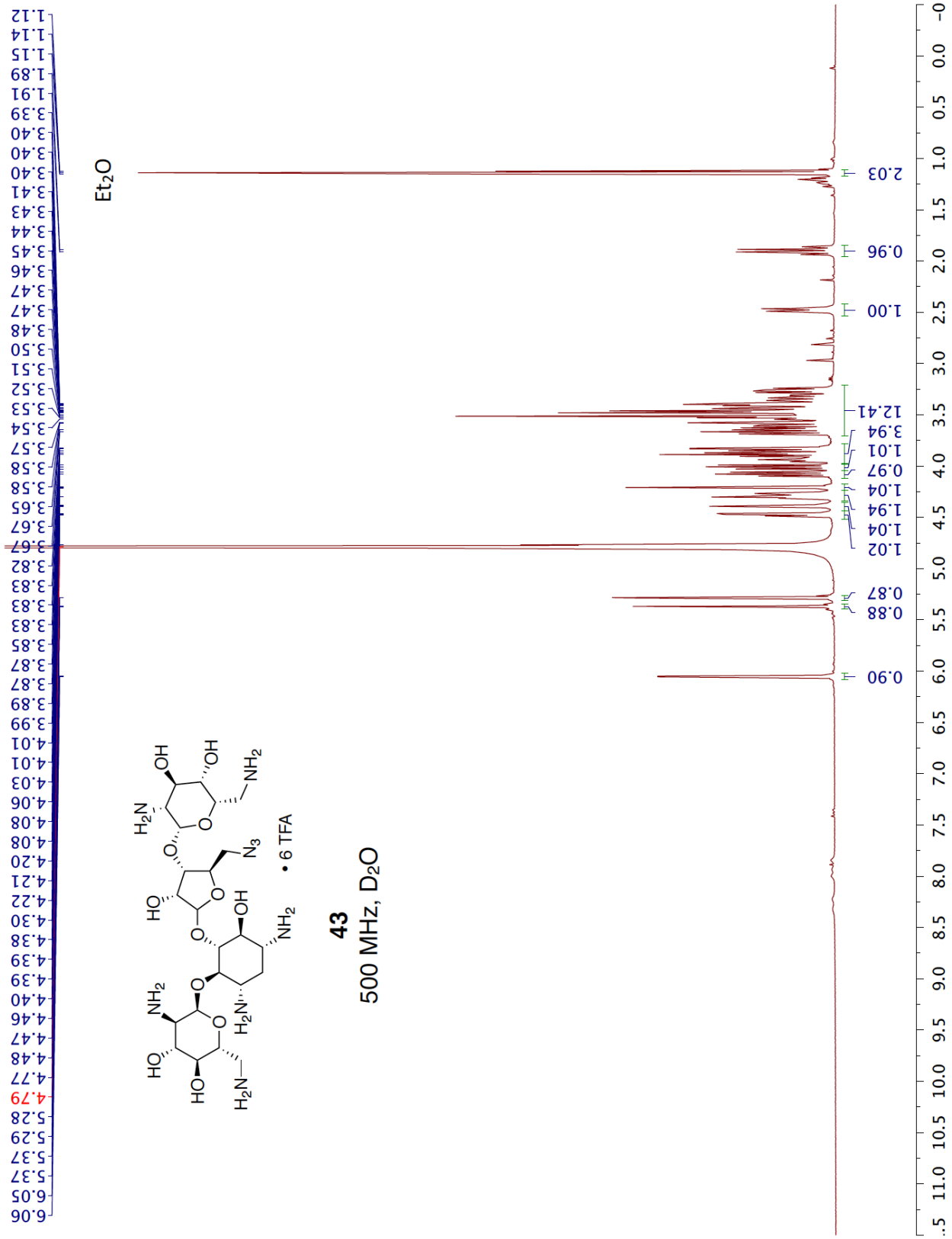

39
500 MHz, DMSO-d_6
MeOH
H_2O
Cl
OMe
NH
HN
N
O

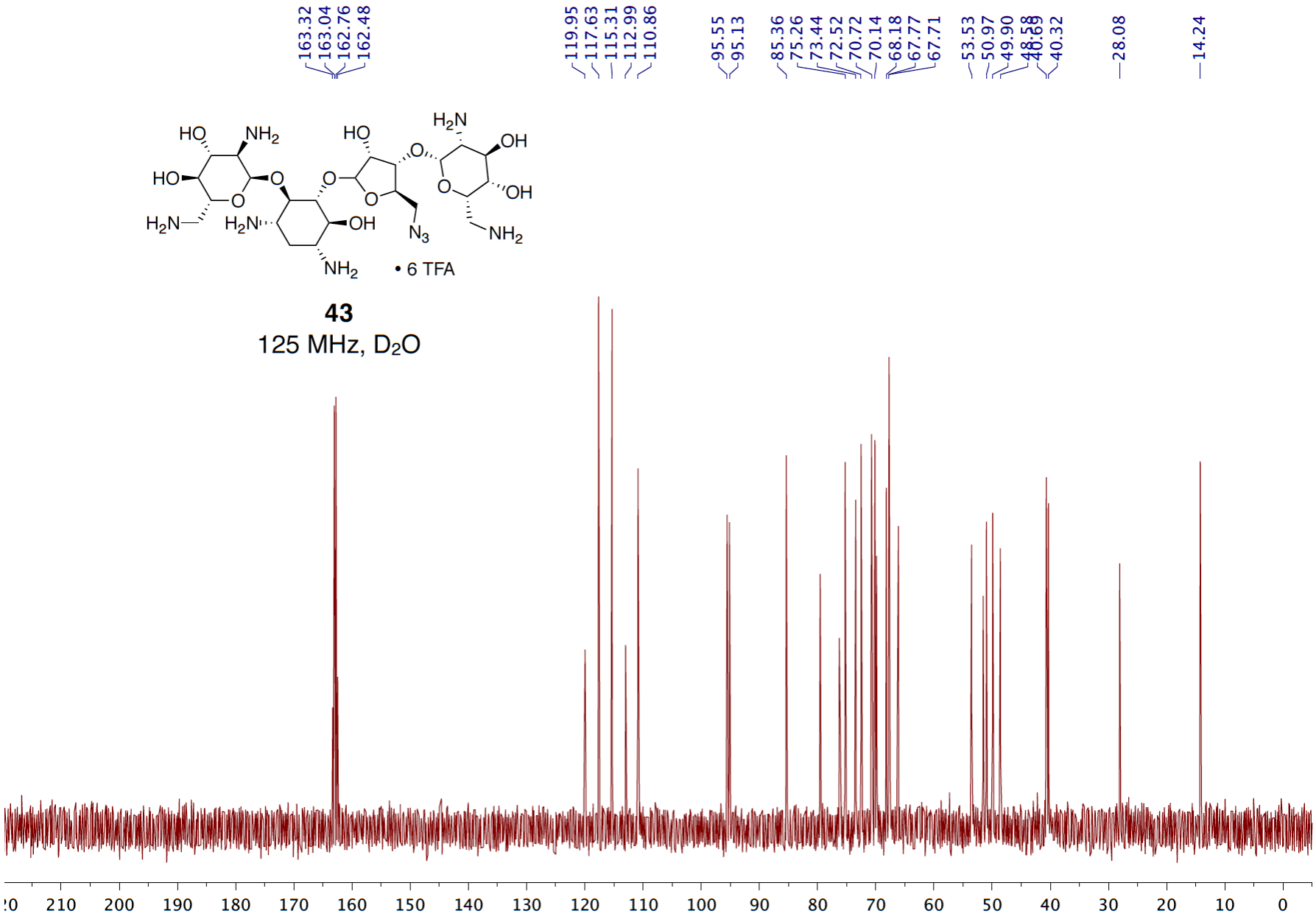

158.05
156.40
155.62
143.80
140.04
128.70
128.15
124.86
124.25
121.23
116.36
112.55
102.29
77.33
77.72
56.68
51.36
49.51
49.34
49.17
49.00
48.83
48.66
48.49
40.89
O
NH
HN
N
OMe
Cl
39
125 MHz, CD3OD-d4
(concentrated)

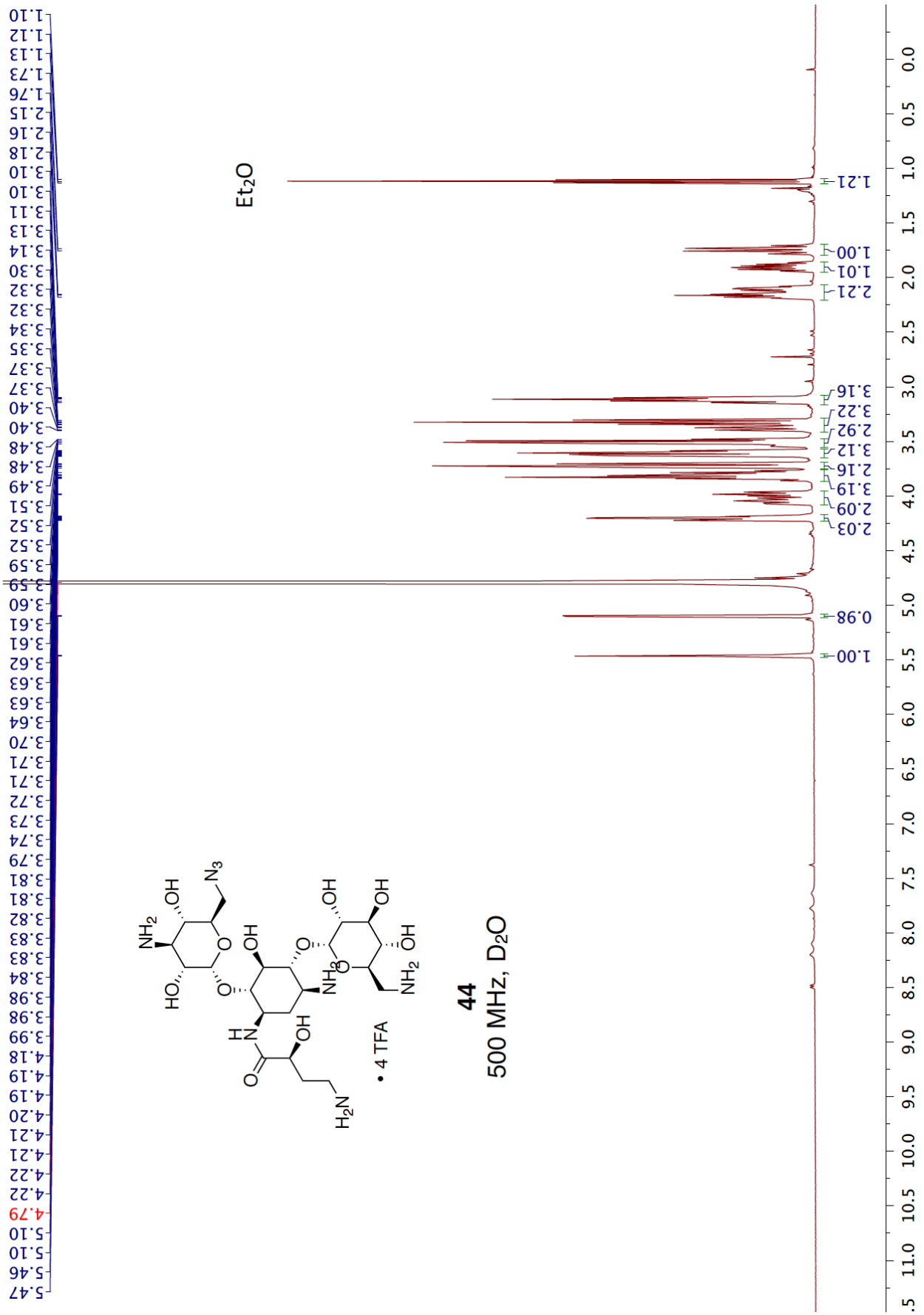

5.46 5.45 5.03 5.03 4.79 3.99 3.99 3.98 3.98 3.97 3.97 3.96 3.96 3.91 3.90 3.90 3.89 3.88 3.87 3.84 3.84 3.83 3.72 3.71 3.71 3.69 3.69 3.68 3.62 3.62 3.62 3.60 3.60 3.60 3.59 3.59 3.58 3.54 3.53 3.52 3.51 3.50 3.50 3.49 3.47 3.47 3.45 3.44 3.43 3.42 3.39 3.36 3.36 3.35 3.33 3.33 3.31 3.29 3.13 3.11 2.49 2.46 2.15 2.14 1.86 1.84
NH2
HO
OH
N3
H2N
O
HO
HO
NH2
HO
O
O
OH
NH2
• 4 TFA
41
500 MHz, D2O
0.97
1.03
2.04
3.18
2.21
9.94
1.07
1.04
1.00
.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

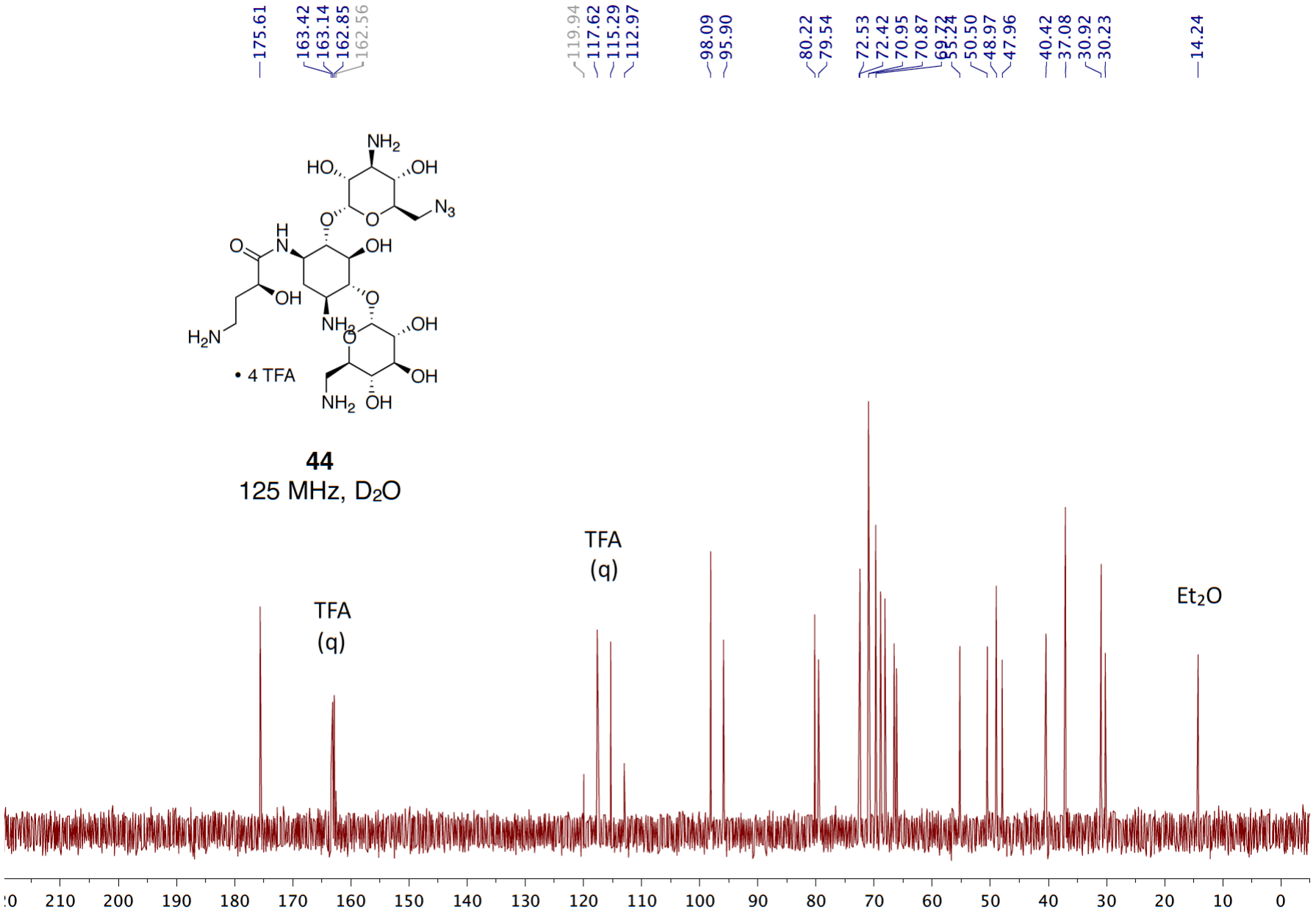

163.39
163.10
162.82
162.54
119.92
117.60
115.28
112.96
100.83
96.20
84.08
78.92
72.45
72.34
71.60
70.91
70.85
68.89
68.26
66.57
54.88
50.82
50.15
47.71
40.43
30.37
27.63
NH2
HO
OH
N3
H2N
HO
HO
NH2
HO
OH
NH2 • 4 TFA
41
125 MHz, D2O
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

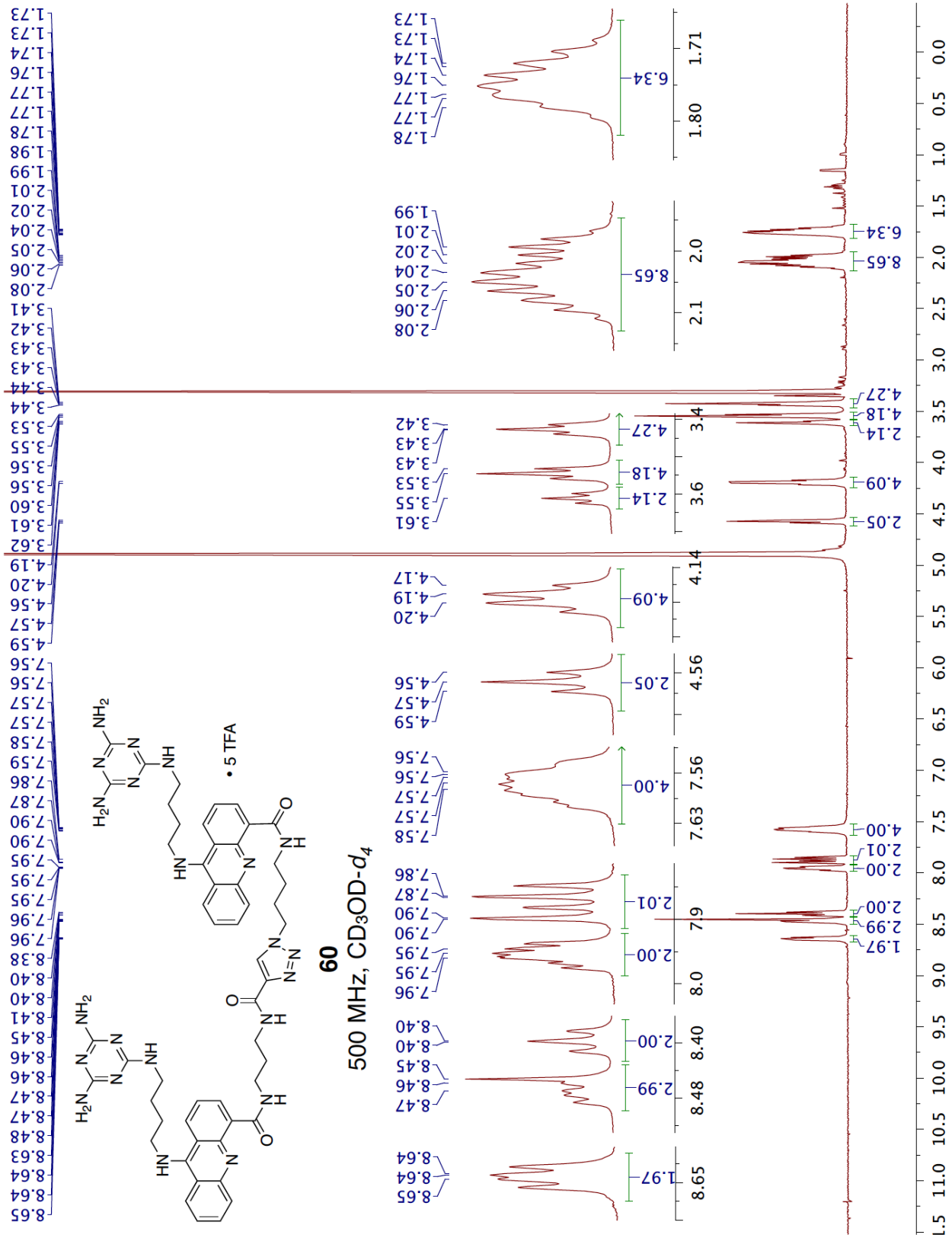

5.76 5.76 5.55 5.54 5.26 5.24 4.79 4.60 4.59 4.59 4.14 4.12 3.97 3.97 3.96 3.95 3.95 3.94 3.93 3.92 3.82 3.82 3.80 3.80 3.79 3.79 3.78 3.78 3.77 3.76 3.76 3.70 3.68 3.67 3.66 3.66 3.62 3.61 3.60 3.60 3.59 3.58 3.58 3.57 3.56 3.55 3.42 3.41 3.40 3.39 3.39 3.37 3.36 3.35 3.35 3.34 3.34 3.33 2.81 2.53 2.51 2.39 2.07 2.05 1.90 1.88
N3
NH
OH
OH
HO
NH2
H2N
OH
OH
O
O
O
O
O
• 5 TFA
NH2
NH2
42
500 MHz, D2O
0.77 0.65 0.75 0.87 0.76 2.60 2.64 1.88 2.72 2.87 2.55 0.94 0.92 0.96 1.00
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0


Et2O
43
500 MHz, D2O
• 6 TFA


163.32
163.04
162.76
162.48
119.95
117.63
115.31
112.99
110.86
95.55
95.13
85.36
75.26
73.44
72.52
70.72
70.14
68.18
67.77
67.71
53.53
50.97
49.90
40.32
28.08
14.24
HO
NH2
HO
H2N
OH
HO
O
O
O
O
O
O
H2N
OH
OH
H2N
H2N
N3
NH2
NH2
• 6 TFA
43
125 MHz, D2O
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0


Et2O
44
500 MHz, D2O
• 4 TFA


175.61
163.42
163.14
162.85
162.56
119.94
117.62
115.29
112.97
98.09
95.90
80.22
79.54
72.53
72.42
70.95
70.87
69.72
55.24
50.50
48.97
47.96
40.42
37.08
30.92
30.23
14.24
• 4 TFA
44
125 MHz, D2O
TFA
(q)
TFA
(q)
Et2O
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0


60
500 MHz, CD3OD-d4
• 5 TFA