

ADVANCES IN MYOTONIC DYSTROPHY TYPE 1 DRUG DISCOVERY THROUGH  
DESIGN OF NOVEL LIGANDS AND MECHANISM ESTABLISHMENT

BY

AMIN HAGHIGHAT JAHROMI

DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Biophysics and Computational Biology  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

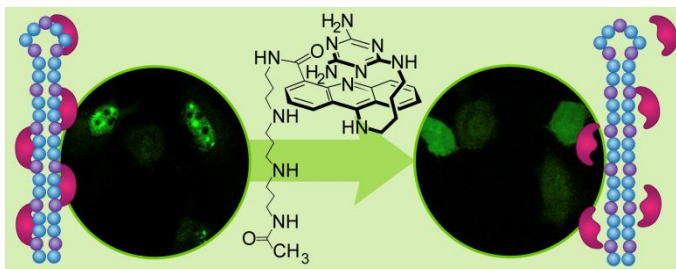
Doctoral Committee:

Professor Steven C. Zimmerman, Chair  
Professor Maria Spies, University of Iowa  
Professor Paul J. Hergenrother  
Professor Susan A. Martinis

# Abstract

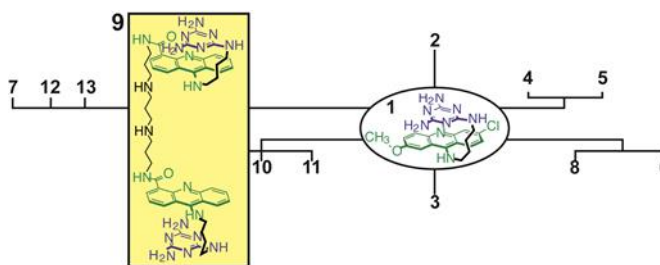
Myotonic dystrophy type 1 (DM1) is caused by an expanded CUG repeat ( $\text{CUG}^{\text{exp}}$ ) that sequesters muscleblind-like 1 protein (MBNL1), a protein that regulates alternative splicing.

$\text{CUG}^{\text{exp}}$  RNA is a validated drug target for this currently untreatable disease. Herein, we describe the development of a bioactive small molecule (Chapter 2) and a small library of dimeric ligands (Chapter



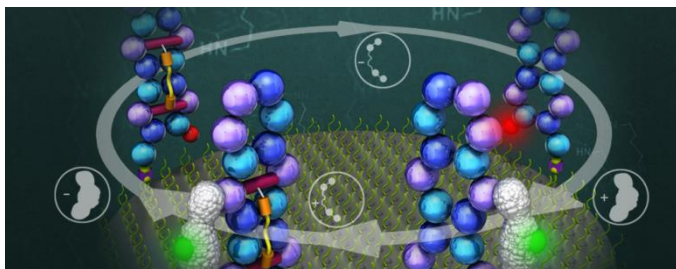
3) leading to an optimized bivalent ligand. These novel ligands target  $\text{CUG}^{\text{exp}}$  RNA and are able to inhibit the  $\text{CUG}^{\text{exp}}$ -MBNL1 interaction in cells that model DM1. In a DM1 cell model these ligands were found to disperse  $\text{CUG}^{\text{exp}}$  ribonuclear foci, release MBNL1, and partially

reverse the mis-splicing of the insulin receptor pre-mRNA. Direct evidence for ribonuclear foci dispersion by this ligand was obtained in a live DM1 cell model using time-lapse confocal



microscopy. In Chapter 4, We report a single-molecule approach to study the binding of MBNL1 to  $(\text{CUG})_{n=4,6}$  and the effect of small molecules on this interaction. MBNL1 is able to bind to the  $(\text{CUG})_n$ -inhibitor complex indicating that the inhibition is not a straight forward competitive process. A simple bivalent ligand, shows a binding to  $(\text{CUG})_n$  almost 50-fold

more tightly than the corresponding monomeric ligand and is more effective in destabilizing MBNL1- $(\text{CUG})_4$ . The single-molecule method and the analysis framework might be extended to the study of



other biomolecular interactions. Chapter 5 includes a preliminary effort to solve the mystery of  $\text{CUG}^{\text{exp}}$  unfolding/folding upon interaction with MBNL1. To approach this unanswered yet key structural question about how MBNL1 binds  $\text{CUG}^{\text{exp}}$ , preliminary bulk FRET (Fluorescence Resonance Energy Transfer) studies, as well as single-molecule FRET studies are described.

بِسْمِ الرَّحْمَنِ الرَّحِيمِ

To whom the universe is waiting for,...

# Acknowledgements

First and foremost I owe my deepest gratitude to my research advisor, Dr. Steven C. Zimmerman, for the opportunity to obtain my degree under his guidance. He is a wonderful person, both scientifically and personally, and I learned a lot from him on a daily basis. I will always appreciate his mentorship and hope to take some of his nice personality with me. Besides being a great chemist, Steve is a wonderful person; this combination makes him the best advisor.

I thank Dr. Maria Spies for the fruitful collaboration we had for the past couple of years and for being a great teacher for me. I learned a lot of single molecule biophysics from her and enjoyed working in her lab both here at the U. of I. and for a few days in her new lab at the University of Iowa. She was the instructor of BIOP 401 in my first semester when I got here almost a month late, and I am grateful for her support back then. Dr. Susan A. Martinis was the first professor Narjes and I met on our arrival here and I am very grateful for her support and encouragement we got from her. I thank Dr. Paul J. Hergenrother for the great collaboration we had and everything I learned from him in our monthly meetings. All of you were more than just a committee member to me and I am grateful to you all.

I would like to thank Dr. Anne M. Baranger for the collaboration we had while she was here as well as Dr. Eric Oldfield for his mentorship during my first 2 years here. I am thankful to our missed beloved department head, Bob Clegg. I had the privilege to know him and am grateful for his continues support.

I thank my master's advisor, Dr. Mehdi Adib at University of Tehran, who introduced me to the fascinating world of research. I appreciate everything he taught me.

I can never thank my mother and father enough. They were my first teachers and I owe everything I have to them. My father was the first university graduate in the family who got his degree in Law from University of Tehran, the most prestigious university in Iran. Impressively, he followed his passion of teaching Persian literature, and always taught me to follow my passion. My mom was my best elementary school teacher who patiently taught me everything.

Last but not least, I thank my amazing wife, Narjes, for being so lovely. I have a wonderful life with her, full of joy and happiness. She is the reason I followed this path and I always enjoyed her feedback and suggestions in my career. Words can't describe how much I appreciate her for being the best wife, one can ever have.



# Table of Contents

|   |           |
|---|-----------|
| Chapter 1. Introduction .....   | 1         |
| 1.1 RNA as a Drug Target .....  | 1         |
| 1.2 Non-coding RNA .....  | 2         |
| 1.3 Role of RNA in Myotonic Dystrophy and Other Diseases.....   | 4         |
| 1.4 Pathogenesis of Myotonic Dystrophy .....  | 6         |
| 1.5 Therapeutic Avenues for Myotonic Dystrophy .....  | 7         |
| 1.6 Myotonic Dystrophy Drug Discovery Program.....  | 9         |
| 1.7 References .....  | 10        |
| <br>  |           |
| Chapter 2. A Novel CUG <sup>exp</sup> ·MBNL1 Inhibitor with Therapeutic<br>Potential for Myotonic Dystrophy Type 1..... | 18        |
| <b>2.1 Abstract .....</b>   | <b>18</b> |
| <br>  |           |
| 2.2 Introduction .....  | 18        |
| 2.3 Results and Discussion.....   | 19        |
| 2.4 Conclusion.....   | 35        |
| 2.5 Materials and Methods.....  | 36        |
| 2.6 Experimental Synthetic Procedures .....   | 40        |
| 2.7 Acknowledgements.....   | 43        |
| 2.8 References .....  | 44        |
| 2.9 NMR Spectra .....   | 48        |
| <br>  |           |
| Chapter 3. The Bivalent Ligand Approach Leads to a Bioactive<br>Inhibitor of MBNL1·CUG <sup>exp</sup> Complex.....      | 62        |
| <b>3.1 Abstract .....</b>   | <b>62</b> |
| <br>  |           |
| 3.2 Introduction .....  | 62        |
| 3.3 Results and Discussion .....  | 64        |

|   |            |
|---|------------|
| 3.4 Conclusion.....   | 94         |
| 3.5 Materials and Methods.....  | 94         |
| 3.6 Experimental Synthetic Procedures.....  | 99         |
| 3.7 Acknowledgements.....   | 111        |
| 3.8 References.....   | 112        |
| 3.9 NMR Spectra.....  | 116        |
| <br>  |            |
| Chapter 4. Single Molecule Study of the CUG Repeat·MBNL1<br>Interaction and its Inhibition by Small Molecules ..... | 169        |
| <b>4.1 Abstract .....</b>   | <b>169</b> |
| <br>  |            |
| 4.2 Introduction .....  | 169        |
| 4.3 Results .....   | 171        |
| 4.4 Discussion .....  | 194        |
| 4.5 Conclusion.....   | 196        |
| 4.6 Materials and Methods.....  | 198        |
| 4.7 Experimental Synthetic Procedures.....  | 204        |
| 4.8 Acknowledgements.....   | 209        |
| 4.9 References .....  | 210        |
| 4.10 NMR Spectra .....  | 214        |
| <br>  |            |
| Chapter 5: CUG <sup>exp</sup> Conformation Study .....  | 235        |
| 5.1 Study of the Effect of MBNL1 on CUG <sup>exp</sup> Conformation .....   | 235        |
| 5.2 Acknowledgements.....   | 238        |
| 5.3 References .....  | 239        |

# Chapter 1.

## Introduction

### 1.1 RNA as a Drug Target

Nucleic acids and proteins form the basic macromolecules of every living organism. As the central dogma of biology describes, DNA contains the genetic information which through the RNA intermediate may be translated into proteins. Traditionally RNA has been considered merely as a passive carrier of genetic information from DNA (genetic information depository) to protein, the functional macromolecules of the cell.<sup>1</sup> As a result, drug discovery efforts have been largely focused on targeting proteins, in almost every disease.<sup>2</sup>

For example, in genetic diseases it has been traditionally assumed that a mutated gene causes the expression of defective proteins, which then affects proper cellular function. Therefore, drug development programs have been historically focused on targeting the defective proteins in the corresponding disease.

As Frith *et al.* explain, considering that over 50% of the genomic DNA of complex organisms is transcribed into RNA, whereas under 2% of this DNA actually encodes proteins,<sup>3</sup> we expect to find RNA at many key regulatory steps of a cell's life.<sup>4</sup> The continuing functional and structural discoveries of RNA biochemistry is also an opportunity for the development of RNA-based therapeutics.<sup>5</sup>

A classic example where RNA serves as a drug target is with ribosome-targeting antibiotics. In fact, to date the most common RNA drug targets have been ribosomal RNA and HIV RNA.<sup>6-8</sup> Proteins are also capable of selective RNA complexation. However, they generally are not able to serve as therapeutic agents for two main reasons: poor pharmacological properties and structural complexity.<sup>9, 10</sup>

Small molecules can bind specific RNA motifs and control biochemical processes in cells and for this reason have been called the missing link in the central dogma by Schreiber *et al.*<sup>11</sup> Unlike proteins, small molecules often show good physicochemical and pharmacological properties as potential therapeutic agents.

Small molecules that specifically bind RNA targets are less common than those that bind DNA targets,<sup>12</sup> due to physicochemical differences between DNA and RNA. For

example, as described by Carlson *et al.* and Weeks *et al.* the 2'-OH group presents additional polar functionality at each residue of RNA. RNA is a versatile molecule, which usually exists in the single stranded form. However, it can fold into different structures and conformations such as hairpin-loops, duplexes, internal-loops, mismatches, and bulges. The most common form of duplex RNA, the A-form contains a shallow "minor groove" and a narrow and deep "major groove" which makes it less accessible for selective binding.<sup>13,14</sup> Because of the polyphosphate backbone and the electronegative groups on the bases protruding into the grooves, RNA (and also DNA) exists as a negatively charged anionic polymer. Therefore, most RNA ligands are therefore positively charged molecules. As a result some of these molecules tend to be relatively nonspecific in binding RNA sequences. It is because of the positive charge of aminoglycosides that they bind with a low specificity to various RNA sequences, such as UU mismatch.<sup>15-17</sup>

## 1.2 Non-coding RNA

Thanks to recent structural and functional discoveries of RNA, it is known that RNA plays multiple essential roles in cells. For instance, RNA has a control over the fundamental process of genetic expression, by a variety of non-coding RNAs (ncRNAs).<sup>18, 19</sup> Genetic expression involves the following four steps:

- Step 1: DNA replication
- Step 2: Transcription of DNA into a precursor messenger RNA (pre-mRNA)
- Step 3: Splicing of this pre-mRNA to a mature mRNA
- Step 4: Translation of the mRNA sequence into a protein sequence

Riboswitch is an example of ncRNA that is able to bind specific metabolites and affect gene expression in steps 2 and 4.<sup>20</sup> MicroRNAs are small ncRNAs that have been shown to be heavily involved in gene regulation by binding complementary sequences in messenger RNAs and degrading them via the RNA interference (RNAi) pathway in step 3.<sup>21</sup> In Step 4, RNAs can undergo self-catalysis or cleavage and also exogenously catalyze fundamental cellular processes as seen in ribosomal RNA and in spliceosomes.<sup>22, 23</sup> In summary, ncRNAs are actively involved in transcription regulation (microRNA, siRNA and riboswitches), alternative splicing (snRNA), and catalyzing biochemical reactions similar to proteins (e.g., ribozymes).

Mutations in certain regions of ncRNAs are linked to several human diseases. Some of these mutations affect the splicing of the transcript produced from the same gene,

which is called “in *cis*” i.e., RNA may indirectly play a role in diseases through the translation of a mutated gene, which may then lead to loss of protein function. On the other hand, some diseases result from mutations in the splicing machinery and the regulatory proteins that affect splicing of other transcripts, which is called “in *trans*” i.e., an RNA produced from a mutant gene may nonspecifically inhibit the function of a regulatory protein, such as the sequestration of splicing factors in RNA repeat disorders, and cause a disease. The discovery of these disease-causing mutations in RNAs provides a range of therapeutic targets,<sup>24, 25</sup>

Some of pathogenic in *trans* mutations, including some forms of cancers, several developmental and neurological diseases like autism, myotonic dystrophy (DM),<sup>26</sup> fragile-X syndrome, frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), and Alzheimer’s disease (AD),<sup>27</sup> are shown in table 1.1.<sup>28</sup>

**Table 1.1** *Trans*-Acting mutations affecting RNA-dependent functions that cause disease

| <b>Disease</b>                              | <b>Gene/Mutation</b>                | <b>Function</b>               |
|---|-------------------------------------|-------------------------------|
| Prader Willi syndrome                       | <i>SNORD116</i>                     | ribosome biogenesis           |
| Spinal muscular atrophy (SMA)               | <i>SMN2</i>                         | Splicing                      |
| Dyskeratosis congenita (X-linked)           | <i>DKC1</i>                         | telomerase/translation        |
| Dyskeratosis congenita (autosomal dominant) | <i>TERC</i>                         | telomerase                    |
| Dyskeratosis congenita (autosomal dominant) | <i>TERT</i>                         | telomerase                    |
| Diamond-Blackfan anemia                     | <i>RPS19, RPS24</i>                 | ribosome biogenesis           |
| Shwachman-Diamond syndrome                  | <i>SBDS</i>                         | ribosome biogenesis           |
| Treacher-Collins syndrome                   | <i>TCOF1</i>                        | ribosome biogenesis           |
| Prostate cancer                             | <i>SNHG5</i>                        | ribosome biogenesis           |
| Myotonic dystrophy, type 1 (DM1)            | <i>DMPK</i> (RNA gain of function)  | protein kinase                |
| Myotonic dystrophy, type 2 (DM2)            | <i>ZNF9</i> (RNA gain of function)  | RNA binding                   |
| Spinocerebellar ataxia 8 (SCA8)             | <i>ATXN8</i> (RNA gain of function) | unknown/noncoding RNA         |
| Huntington's disease-like 2 (HDL2)          | <i>JPH3</i> (RNA gain of function)  | ion channel function          |
| Fragile X-associated tremor ataxia syndrome | <i>FMR1</i> (RNA gain of function)  | translation/mRNA localization |
| Fragile X syndrome                          | <i>FMR1</i>                         | translation/mRNA localization |
| X-linked mental retardation                 | <i>UPF3B</i>                        | translation/nonsense decay    |
| Oculopharyngeal muscular dystrophy (OPMD)   | <i>PABPN1</i>                       | 3' end formation              |

**Table 1.1 (cont.)**

|   |                                     |                               |
|---|-------------------------------------|-------------------------------|
| Human pigmentary genodermatosis                 | <i>DSRAD</i>                        | Editing                       |
| Retinitis pigmentosa                            | <i>PRPF31</i>                       | Splicing                      |
| Retinitis pigmentosa                            | <i>PRPF8</i>                        | Splicing                      |
| Retinitis pigmentosa                            | <i>HPRP3</i>                        | Splicing                      |
| Retinitis pigmentosa                            | <i>PAP1</i>                         | Splicing                      |
| Cartilage-hair hypoplasia (recessive)           | <i>RMRP</i>                         | Splicing                      |
| Autism  | <i>7q22-q33 locus breakpoint</i>    | noncoding RNA                 |
| Beckwith-Wiedemann syndrome (BWS)               | <i>H19</i>                          | noncoding RNA                 |
| Charcot-Marie-Tooth (CMT) Disease               | <i>GRS</i>                          | Translation                   |
| Charcot-Marie-Tooth (CMT) Disease               | <i>YRS</i>                          | Translation                   |
| Amyotrophic lateral sclerosis (ALS)             | <i>TARDBP</i>                       | Splicing, transcription       |
| Leukoencephalopathy                             | <i>EIF2B1</i>                       | Translation                   |
| Wolcott-Rallison syndrome                       | <i>EIF2AK3</i>                      | translation (protease)        |
| Mitochondrial myopathy and sideroblastic anemia | <i>PUS1</i>                         | Translation                   |
| Hypertrophic cardiomyopathy                     | <i>TSMF</i>                         | translation (mitochondrial)   |
| Hereditary spastic paraplegia                   | <i>SPG7</i>                         | ribosome biogenesis           |
| Leukoencephalopathy                             | <i>DARS2</i>                        | translation (mitochondrial)   |
| Susceptibility to diabetes mellitus             | <i>LARS2</i>                        | translation (mitochondrial)   |
| Deafness  | <i>MTRNR1</i>                       | ribosome biogenesis           |
| MELAS syndrome, deafness                        | <i>MTRNR2</i>                       | ribosome biogenesis           |
| Cancer  | <i>SFRS1</i>                        | Splicing, translation, export |
| Cancer  | <i>RBM5</i>                         | Splicing                      |
| Multiple disorders                              | <i>mitochondrial tRNA mutations</i> | translation (mitochondrial)   |
| Cancer  | <i>miR-17-92 cluster</i>            | RNA interference              |
| Cancer  | <i>miR-372, miR-373</i>             | RNA interference              |

a. Table obtained from Cooper *et al.*<sup>28</sup>

### 1.3 Role of RNA in Myotonic Dystrophy and Other Diseases

Myotonic dystrophy (DM) is an RNA-mediated disease caused by a non-coding CTG repeat expansion in the DMPK gene, chromosome 19q.13.3 (DM type 1, abbreviated DM1) or a CCTG expansion in the ZNF9 gene, chromosome 3q. (DM type 2, abbreviated DM2).<sup>29-31</sup> These repeats are among the so called microsatellites or simple sequence repeats (SSRs).<sup>32</sup> They are tandemly repeated tracts of DNA composed of 1–6 base pair (bp) long units that can be found either in translated or untranslated 3'- or 5'-regions of the genome. Tetranucleotide repeats are the most abundant type in exons

whereas trinucleotide repeats are found most often in introns. These repeats are a result of misalignment and slippage of complimentary strands during DNA replication and repair.<sup>33,34</sup> This slipped sequence repeat can undergo expansion through a complex mechanism.<sup>35</sup> Trinucleotide repeat expansion disorders (TRED) can be divided in the following two general categories<sup>36</sup>:

- Polyglutamine (PolyQ) disorders (Table 1.2)
- Non-polyQ disorders (Table 1.3)

**Table 1.2.** Polyglutamine TRED diseases.<sup>37</sup>

| Disease  | Gene  | Number of repeats |
|--|---|-------------------|
| DRPLA (Dentatorubropallidoluysian atrophy)             | <i>ATN1 or DRPLA</i>                          | 49 – 88           |
| HD (Huntington's disease)                              | <i>HTT (Huntington)</i>                       | >35               |
| SBMA (Spinobulbar muscular atrophy or Kennedy disease) | <i>Androgen receptor on the X chromosome.</i> | 38 – 62           |
| SCA1 (Spinocerebellar ataxia Type 1)                   | <i>ATXN1</i>                                  | 49 – 88           |
| SCA2 (Spinocerebellar ataxia Type 2)                   | <i>ATXN2</i>                                  | 33 - 77           |
| SCA3 (Spinocerebellar ataxia Type 3)                   | <i>ATXN3</i>                                  | 55 - 86           |
| SCA6 (Spinocerebellar ataxia Type 6)                   | <i>CACNA1A</i>                                | 21 - 30           |
| SCA7 (Spinocerebellar ataxia Type 7)                   | <i>ATXN7</i>                                  | 38 - 120          |
| SCA17 (Spinocerebellar ataxia Type 17)                 | <i>TBP</i>                                    | 47 - 63           |

PolyQ diseases originate from a CAG repeat in the coding region that translates into Polyglutamine in the protein. Non-polyQ disorders are caused by an expansion in the non-coding region. These disorders can be further divided into two categories:

- Loss-of-function pathogenesis such as Freidreich ataxia and fragile-X syndrome
- Gain-of-function pathogenesis such as fragile X-associated tremor ataxia syndrome (FXTAS), spinocerebellar ataxia 3 (SCA3) and myotonic dystrophy (DM1 and DM2).

Myotonic dystrophy type 1 (DM 1), also known as Steinert's disease, is the most common adult muscular dystrophy with a prevalence of 1:8000.<sup>38, 39</sup> In general, DM is an autosomal dominant inherited disease, with multi-systemic disorders such as myotonia, myopathy, cardiac arrhythmias, insulin resistance and progressive muscle degeneration.<sup>38</sup> Myopathy is described as muscle weakness whereas myotonia accounts for delayed muscle relaxation, often described as stiffness. DM2 has similar presentation to DM 1 but with milder symptoms, in general.

**Table 1.3.** Non-Polyglutamine TRED diseases.<sup>37</sup>

| Disease            | Gene                              | Codon | Number of repeats |
|--------------------|-----------------------------------|-------|-------------------|
| Fragile X syndrome | <i>FMR1</i> , on the X-chromosome | CGG   | 230+              |

|   |                                   |     |           |
|---|-----------------------------------|-----|-----------|
| FXTAS (Fragile X-associated tremor/ataxia syndrome) | <i>FMR1</i> , on the X-chromosome | CGG | 50-200    |
| POI (Premature ovarian insufficiency)               | <i>FMR1</i> , on the X-chromosome | CGG | 50-200    |
| DM1 (Myotonic dystrophy type 1)                     | <i>DMPK</i>                       | CTG | 80- >2500 |
| HDL2 (Huntington's disease-like 2)                  | <i>JPH3</i>                       | CTG | 66-78     |
| SCA 8 (Spinocerebellar ataxia Type 8)               | <i>SCA8</i>                       | CTG | 110 – 250 |
| SCA 12 (Spinocerebellar ataxia Type 12)             |                                   | CAG | 45-96     |

## 1.4 Pathogenesis of Myotonic Dystrophy

In eukaryotic cells, pre-mRNA needs to be converted to mature mRNA in a process called splicing, in which non-coding introns are removed and remaining exons are ligated to each other. This process is carried out by spliceosomes which recognizes and cuts out introns.<sup>40, 41</sup> Therefore, during this process, various mRNAs can be made from a single pre-mRNA transcript, depending on which introns are excluded and which exons are included.<sup>42</sup> This process is called alternative splicing, a common event in human genome. Different studies have found that between 74% to 95% of human genes undergo alternative splicing.<sup>43, 44</sup> DM is among the diseases that are caused from alternative splicing misregulation.<sup>45</sup>

The pathogenesis of DM involves misregulation of alternative splicing by the sequestration of regulatory factors such as the MBNL (muscleblind like) protein family and CUG-BP1.<sup>46</sup> The MBNL family of proteins, especially MBNL1, is sequestered by mutant transcripts containing C(C)UG repeat RNAs which form discrete ribonuclear foci and these, in turn, lead to the loss of MBNL1 function.<sup>29</sup> CUG-BP1 binds to CUG and CCUG,<sup>47</sup> but doesn't colocalize with them in the nucleus.<sup>48</sup> Other proteins, such as hnRNP H and the transcription factor Sp1 are also suggested to be involved in the DM pathogenesis,<sup>49</sup> however their role is unclear at this time.<sup>50, 51</sup> The current model for DM pathogenesis is a gain-of-function of CUG-BP1 and loss-of-function of MBNL1.<sup>46</sup> As a result, various pre-mRNAs whose splicing is regulated by MBNL1 would be mis-spliced. For example, mis-splicing of chloride channel gives rise to myotonia symptoms and mis-splicing of insulin receptor causes insulin resistance in DM patients.<sup>52, 53</sup> Because MBNL1 is an RNA binding protein that regulates alternative splicing of a specific subset of pre-mRNAs during postnatal development, loss of this splicing factor, due to its sequestration, results in missplicing of MBNL1 target pre-mRNAs and perturbation of developmental signals. The following five findings support the role of MBNL1 in DM



pathogenesis: 1) MBNL1 binds to CUG<sup>exp</sup> with a high affinity; 2) it is colocalized with CUG<sup>exp</sup> in ribonuclear foci; 3) a mouse knockout model of MBNL1 causes characteristics of DM;<sup>54</sup> 4) MBNL1 binds to introns of pre-mRNAs that are mis-spliced in DM;<sup>55</sup> 5) overexpression of MBNL1 in a DM1 mouse model corrects the related pre-mRNA splicing defects.<sup>56</sup> These findings support the DM1 pathogenesis hypothesis and validate CUG<sup>exp</sup> as a drug target. In myotonic dystrophy, intervention at the RNA level might be the most advantageous. Elimination of DM symptoms by turning off the transcription of CUG<sup>exp</sup> in a mouse model, has confirmed the reversibility of DM1 pathogenesis.<sup>57</sup>

## 1.5 Therapeutic Avenues for Myotonic Dystrophy

Current therapeutics for DM patients is limited to palliative therapy with minimal clinical effects. Such palliative treatments include: selenium, vitamin E, baclofen, nifedipine, creatine monohydrate and testosterone, DHEA-S (dihydroepiandrosterone sulfate) and bioflavonoids.<sup>58</sup> They mainly relieve muscular degeneration, or relax delayed muscular contraction.<sup>59</sup> Possible drug targets against myotonic dystrophy were described by Foff et al. as the follows:<sup>60</sup>

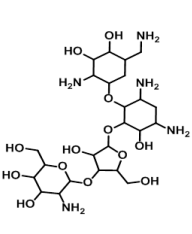

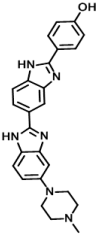
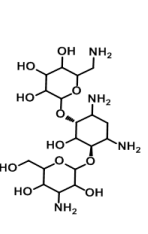
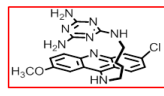
- **Targeting DNA repeat expansion and instability.** Although this is an unexplored field small molecules that target the repeat sequence itself and interfere with the expansion can be promising ligands.<sup>61</sup> For example, a naphthyridine-azaquinolone ligand selectively binds to CAG repeats and induces nucleotide flipping; it has been suggested as a diagnostic tool for repeat expansion.<sup>62</sup>
- **Targeting the toxic RNA.** This approach mainly uses ribozymes (**ribonucleic acid enzyme**). Also called an RNA enzyme or catalytic RNA), ribozymes are RNA molecules that catalyze chemical reactions. Thus, they can cleave other RNAs and ultimately lead to degradation of their targets.<sup>63</sup> Despite some progresses in this approach,<sup>64, 65</sup> selectivity is still a big challenge.<sup>66</sup>
- **Targeting the Structural CUG<sup>exp</sup> Hairpin.** As described by Foff *et al.*,<sup>67</sup> CUG<sup>exp</sup> long hairpin structures have been shown to be substrates for Dicer, a ribonuclease whose main function in the RNA interference pathway is to recognize double-stranded RNA duplexes and induce cleavage of these structures into shorter, 21-nt duplexes. These duplexes are then incorporated into the RNA-induced silencing complex (RISC) and, as such, function to guide

substrate selection for degradation.<sup>68</sup> A synthetic oligoribonucleotide (CUG)<sub>7</sub> in DM1 fibroblasts led to a selective reduction in CUG<sup>exp</sup> level,<sup>69</sup> supporting the hypothesis that using siRNAs (small interfering RNAs) to target the CUG<sup>exp</sup> in DM1 would tag it for destruction via the Dicer pathway. Delivery and selectivity remain the big challenges in this approach.

- **Targeting CUG<sup>exp</sup>·MBNL1 binding.** The disruption of the CUG<sup>exp</sup>·MBNL1 complex can be achieved by antisense oligonucleotides (AON) or small molecules. Two forms of AON are currently applied in DM1. One is 2'-O-methyl-modified AONs (MOE), and the other is the phosphorodiamidate morpholino antisense molecules (PMO or morpholino). Morpholinos differ from other antisense strategies in that they provide inhibition via steric hindrance, rather than by any targeted degradation.<sup>70</sup> 2'-O-Methyl-phosphorothioate-modified (CAG)<sub>7</sub> AON,<sup>71</sup> and AON containing (CAG)<sub>25</sub>,<sup>72</sup> have the effect of disruption of CUG<sup>exp</sup>·MBNL1 interaction and a significant degradation of the CUG<sup>exp</sup> itself.

There are several small molecules at various stages of development as potential therapeutic agents for DM1. Kanamycin A was found as a high-affinity binder of pyrimidine-rich loops similar to CUG<sup>exp</sup>.<sup>73, 74</sup> Bis-benzimidazole Hoechst 33258, binds to the DM1 motif as well.<sup>75</sup> Pentamidine,<sup>76</sup> benzo[g]quinolone-based heterocycles,<sup>77</sup> a Hoechst derivative (H1),<sup>78</sup> a modularly assembled Hoechst 33258,<sup>79, 80</sup> and an acridine-melamine conjugate, reported by our laboratory,<sup>81</sup> are examples of bioactive CUG repeat binders.

**Table 1.4.** Small molecule inhibitors of CUG<sup>exp</sup>·MBNL1 interaction. IC<sub>50</sub> values are against different length of CUG repeats and cannot be directly compared.

| Ligand                |  |  |  |  |  |
|-----------------------|---|---|--|---|---|
| Name                  | Kanamycin A   | Pentamidine   | Hoechst 33258  | Neomycin B  | Melamine-Acridine conjugate   |
| IC <sub>50</sub> (μM) | 190   | 58  | 110  | 280   | 52  |

- **Altering Levels of RNA Binding Proteins.** Increasing levels of MBNL1 or decreasing levels of CUGBP1 are two main options in this approach. Because MBNL1 is an RNA binding protein that regulates alternative splicing of a specific subset of pre-mRNAs during postnatal development, loss of this splicing factor, due to its sequestration, results in missplicing of MBNL1 target pre-mRNAs and perturbation of developmental signals. Treatment of a DM1 mouse model with a protein kinase C (PKC) inhibitor, Ro-31-8220, significantly improved mortality (from 80% mock treated to 20% treated). Concomitantly, CUGBP1 was not hyperphosphorylated, and protein levels were not elevated. Cardiac conduction and contractility were also improved in the treated mice.<sup>82</sup> This finding shows that prevention of PKC mediated phosphorylation of CUGBP1 is effective. However, therapeutic impact is presumably low, given the complexity of the cellular process.
- **Alteration of Downstream Splicing Targets.** As described by Foff *et al.*,<sup>67</sup> targeting this far down the pathway gives a limited return on the investment. In other words, anti-sense oligonucleotide designed to reverse the insulin receptor missplicing may fix the insulin resistance, but the rest of signs and symptoms would still be present. This may ultimately be a viable therapeutic approach, but only if a method for targeting each downstream process is developed. The fact that there are so many missplicing involved in DM disease, makes the success of this approach unlikely.<sup>49</sup>
- **Symptom Management.** The use of mexelitine, which modulates sodium channels and thus lessens myotonia, and central nervous system stimulants to address fatigue are two such examples that are routinely applied in practice.<sup>83</sup> It has been shown that injection of exogenous insulin-like growth factor 1 (IGF-1) could improve muscle strength and function in adult DM1 patients, presumably owing to its anabolic effect on muscle.<sup>84, 85</sup>

## 1.6 Myotonic Dystrophy Drug Discovery Program

As it was explained in detail in the previous sections, the CTG<sup>exp</sup> produces a toxic RNA transcript (CUG<sup>exp</sup>) that does not exit the nucleus but associates with proteins. One of these proteins, MBNL1, is an important regulator of alternative splicing.<sup>86</sup> Sequestration of MBNL1 in nuclear foci leads to multiple misspliced pre-mRNAs, incorrect protein levels, and ultimately the signs and symptoms of the disease.<sup>87</sup>

Knowing that CUG<sup>exp</sup> is the causative agent of DM1, almost all of the DM drug discovery efforts have been focused on targeting CUG<sup>exp</sup>.

Antisense oligonucleotides and small interfering RNAs control gene expression by triggering the degradation of mRNA by recruiting RNase H or the RNA-induced silencing complex, respectively.<sup>88</sup> In a mouse model of DM1 a morpholino antisense oligonucleotide (ASO),<sup>72</sup> a 2'-O-(2-methoxyethyl) ASO<sup>89</sup> and a D-amino acid hexapeptide (ABP1)<sup>90</sup> reversed this process, rescued the missplicing and reversed the phenotype in mice thereby validating CUG<sup>exp</sup> as a therapeutic target. Because no drugs are currently available to treat DM1, there is intense interest in finding small molecules that may function in a manner similar to the morpholino antisense oligonucleotides, but avoid the limitations inherent in the antisense therapeutic approach.<sup>60</sup> Pentamidine,<sup>76</sup> benzo[g]quinolone heterocycle derivatives,<sup>77</sup> a Hoechst derivative (H1),<sup>78</sup> and modularly assembled Hoechst 33258<sup>79</sup> are examples of bioactive CUG<sup>exp</sup> binders that are being developed in other laboratories and have been found to restore MBNL1 function in DM1 cell and animal models.<sup>91</sup> Recently, a CUG<sup>exp</sup> binder small molecule was appended with a moiety that generates hydroxyl radicals upon photolysis. This hydroxyl-mediated CUG<sup>exp</sup> cleavage improved DM1-associated defects in cell culture.<sup>92</sup>

## 1.7 References

- (1) Vicens, Q. (2009) RNA's coming of age as a drug target, *J. Incl. Phenom. Macro.* 65, 171-188.
- (2) Gao, Z., Li, H., Zhang, H., Liu, X., Kang, L., Luo, X., Zhu, W., Chen, K., Wang, X., and Jiang, H. (2008) PDTD: a web-accessible protein database for drug target identification, *BMC Bioinformatics* 9, 104.
- (3) Frith, M. C., Pheasant, M., and Mattick, J. S. (2005) The amazing complexity of the human transcriptome, *European journal of human genetics : EJHG* 13, 894-897.
- (4) Sharp, P. A. (2009) The centrality of RNA, *Cell* 136, 577-580.
- (5) Vicens, Q., and Westhof, E. (2003) RNA as a drug target: the case of aminoglycosides, *ChemBioChem* 4, 1018-1023.
- (6) Wilson, W. D., and Li, K. (2000) Targeting RNA with small molecules, *Curr. Med. Chem.* 7, 73-98.
- (7) Pearson, N. D., and Prescott, C. D. (1997) RNA as a drug target, *Chem. Biol.* 4, 409-414.

- (8) Tor, Y. (2003) Targeting RNA with small molecules, *ChemBioChem* 4, 998-1007.
- (9) Burd, C. G., and Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins, *Science* 265, 615-621.
- (10) Huang, Y. F., Chiu, L. Y., Huang, C. C., and Huang, C. K. (2010) Predicting RNA-binding residues from evolutionary information and sequence conservation, *BMC Genomics* 11 Suppl 4, S2.
- (11) Schreiber, S. L. (2005) Small molecules: the missing link in the central dogma, *Nat. Chem. Biol.* 1, 64-66.
- (12) Thomas, J. R., and Hergenrother, P. J. (2008) Targeting RNA with small molecules, *Chem. Rev.* 108, 1171-1224.
- (13) Carlson, C. B., Stephens, O. M., and Beal, P. A. (2003) Recognition of double-stranded RNA by proteins and small molecules, *Biopolymers* 70, 86-102.
- (14) Weeks, K. M., and Crothers, D. M. (1993) Major groove accessibility of RNA, *Science* 261, 1574-1577.
- (15) Tran, T., and Disney, M. D. (2011) Molecular recognition of 6'-N-5-hexynoate kanamycin A and RNA 1x1 internal loops containing CA mismatches, *Biochemistry* 50, 962-969.
- (16) Aminova, O., Paul, D. J., Childs-Disney, J. L., and Disney, M. D. (2008) Two-dimensional combinatorial screening identifies specific 6'-acylated kanamycin A- and 6'-acylated neamine-RNA hairpin interactions, *Biochemistry* 47, 12670-12679.
- (17) Disney, M. D., and Childs-Disney, J. L. (2007) Using selection to identify and chemical microarray to study the RNA internal loops recognized by 6'-N-acylated kanamycin A, *ChemBioChem* 8, 649-656.
- (18) Mattick, J. S. (2001) Non-coding RNAs: the architects of eukaryotic complexity, *EMBO reports* 2, 986-991.
- (19) Mattick, J. S., and Makunin, I. V. (2006) Non-coding RNA, *Hum. Mol. Genet.* 15 Spec No 1, R17-29.
- (20) Meister, G., and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA, *Nature* 431, 343-349.
- (21) Ambros, V. (2004) The functions of animal microRNAs, *Nature* 431, 350-355.
- (22) Williamson, J. R. (2008) Biophysical studies of bacterial ribosome assembly, *Curr. Opin. Struct. Biol.* 18, 299-304.
- (23) Moore, P. B. (2009) The ribosome returned, *J. Biol.* 8, 8.

- (24) Pearson, N. (1997) RNA as a drug target, *Chemistry* 4, 409-414.
- (25) Xavier, K. A., Eder, P. S., and Giordano, T. (2000) RNA as a drug target: methods for biophysical characterization and screening, *Trends Biotechnol.* 18, 349-356.
- (26) Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., and Thornton, C. a. (2009) Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA., *Science (New York, N.Y.)* 325, 336-339.
- (27) Ule, J. (2008) Ribonucleoprotein complexes in neurologic diseases, *Curr. Opin. Neurobiol.* 18, 516-523.
- (28) Cooper, T. A., Wan, L., and Dreyfuss, G. (2009) RNA and disease, *Cell* 136, 777-793.
- (29) Fardaei, M., Rogers, M. T., Thorpe, H. M., Larkin, K., Hamshere, M. G., Harper, P. S., and Brook, J. D. (2002) Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells, *Hum. Mol. Genet.* 11, 805-814.
- (30) 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., and et al. (1992) 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* 69, 385.
- (31) Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., and Ranum, L. P. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9, *Science* 293, 864-867.
- (32) Toth, G., Gaspari, Z., and Jurka, J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis, *Genome Res.* 10, 967-981.
- (33) Pearson, C. E., Nichol Edamura, K., and Cleary, J. D. (2005) Repeat instability: mechanisms of dynamic mutations, *Nat. Rev. Genet.* 6, 729-742.
- (34) Awad, R., Ahmed, F., Bourbonnais-Spear, N., Mullally, M., Ta, C. A., Tang, A., Merali, Z., Maquin, P., Caal, F., Cal, V., Poveda, L., Vindas, P. S., Trudeau, V. L., and Arnason, J. T. (2009) Ethnopharmacology of Q'eqchi' Maya antiepileptic and anxiolytic plants: effects on the GABAergic system, *J. Ethnopharmacol.* 125, 257-264.
- (35) Mirkin, S. M. (2007) Expandable DNA repeats and human disease., *Nature* 447, 932-940.

- (36) Orr, H. T., and Zoghbi, H. Y. (2007) Trinucleotide repeat disorders, *Annu. Rev. Neurosci.* 30, 575-621.
- (37) Cooper, T. a., Wan, L., and Dreyfuss, G. (2009) RNA and disease., *Cell* 136, 777-793.
- (38) Machuca-Tzili, L., Brook, D., and Hilton-Jones, D. (2005) Clinical and molecular aspects of the myotonic dystrophies: a review, *Muscle Nerve* 32, 1-18.
- (39) Cooper, T. A. (2006) A reversal of misfortune for myotonic dystrophy?, *New Engl. J. Med.* 355, 1825-1827.
- (40) Wahl, M. C., Will, C. L., and Luhrmann, R. (2009) The spliceosome: design principles of a dynamic RNP machine, *Cell* 136, 701-718.
- (41) Staley, J. P., and Guthrie, C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things, *Cell* 92, 315-326.
- (42) Smith, C. W., and Valcarcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control, *Trends Biochem. Sci* 25, 381-388.
- (43) Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R., and Shoemaker, D. D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays, *Science* 302, 2141-2144.
- (44) Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing, *Nat. Genet.* 40, 1413-1415.
- (45) Garcia-Blanco, M. A., Baraniak, A. P., and Lasda, E. L. (2004) Alternative splicing in disease and therapy, *Nat. Biotechnol.* 22, 535-546.
- (46) Ranum, L. P., and Cooper, T. A. (2006) RNA-mediated neuromuscular disorders, *Annu. Rev. Neurosci.* 29, 259-277.
- (47) Timchenko, N. A., Welm, A. L., Lu, X., and Timchenko, L. T. (1999) CUG repeat binding protein (CUGBP1) interacts with the 5' region of C/EBPbeta mRNA and regulates translation of C/EBPbeta isoforms, *Nucleic Acids Res.* 27, 4517-4525.
- (48) Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T., and Thornton, C. A. (2004) Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons, *Hum. Mol. Genet.* 13, 3079-3088.
- (49) Mahadevan, M. S. (2012) Myotonic dystrophy: is a narrow focus obscuring the rest of the field?, *Current opinion in neurology* 25, 609-613.

- (50) Kim, D.-H. (2005) HnRNP H inhibits nuclear export of mRNA containing expanded CUG repeats and a distal branch point sequence, *Nucleic Acids Res.* 33, 3866-3874.
- (51) Ebralidze, A., Wang, Y., Petkova, V., Ebralidse, K., and Junghans, R. P. (2004) RNA leaching of transcription factors disrupts transcription in myotonic dystrophy, *Science* 303, 383-387.
- (52) Savkur, R. S., Philips, A. V., and Cooper, T. A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy, *Nat. Genet.* 29, 40-47.
- (53) Mankodi, A., Takahashi, M. P., Jiang, H., Beck, C. L., Bowers, W. J., Moxley, R. T., Cannon, S. C., and Thornton, C. A. (2002) Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy, *Mol. Cell* 10, 35-44.
- (54) Kanadia, R. N., Johnstone, K. A., Mankodi, A., Lungu, C., Thornton, C. A., Esson, D., Timmers, A. M., Hauswirth, W. W., and Swanson, M. S. (2003) A muscleblind knockout model for myotonic dystrophy, *Science* 302, 1978-1980.
- (55) Ho, T. H., Savkur, R. S., Poulos, M. G., Mancini, M. A., Swanson, M. S., and Cooper, T. A. (2005) Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy, *J. Cell Sci.* 118, 2923-2933.
- (56) Kanadia, R. N., Shin, J., Yuan, Y., Beattie, S. G., Wheeler, T. M., Thornton, C. A., and Swanson, M. S. (2006) Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A.* 103, 11748-11753.
- (57) Mahadevan, M. S., Yadava, R. S., Yu, Q., Balijepalli, S., Frenzel-McCardell, C. D., Bourne, T. D., and Phillips, L. H. (2006) Reversible model of RNA toxicity and cardiac conduction defects in myotonic dystrophy, *Nat. Genet.* 38, 1066-1070.
- (58) Meola, G., and Sansone, V. (2004) Treatment in myotonia and periodic paralysis, *Revue neurologique* 160, S55-69.
- (59) Di Prospero, N. A., and Fischbeck, K. H. (2005) Therapeutics development for triplet repeat expansion diseases, *Nat. Rev. Genet.* 6, 756-765.
- (60) Foff, E. P., and Mahadevan, M. S. (2011) Therapeutics development in myotonic dystrophy type 1, *Muscle Nerve* 44, 160-169.



- (61) Gomes-Pereira, M., Cooper, T. A., and Gourdon, G. (2011) Myotonic dystrophy mouse models: towards rational therapy development, *Trends Mol. Med.* 17, 506-517.
- (62) Nakatani, K., Hagihara, S., Goto, Y., Kobori, A., Hagihara, M., Hayashi, G., Kyo, M., Nomura, M., Mishima, M., and Kojima, C. (2005) Small-molecule ligand induces nucleotide flipping in (CAG)<sub>n</sub> trinucleotide repeats, *Nat. Chem. Biol.* 1, 39-43.
- (63) Sioud, M. (2004) Ribozyme- and siRNA-mediated mRNA degradation: a general introduction, *Methods Mol. Biol.* 252, 1-8.
- (64) Phylactou, L. A., Darrah, C., and Wood, M. J. (1998) Ribozyme-mediated trans-splicing of a trinucleotide repeat, *Nat. Genet.* 18, 378-381.
- (65) Phylactou, L. A. (2004) Repair of myotonic dystrophy protein kinase (DMPK) transcripts by trans-splicing ribozymes, *Methods Mol. Biol.* 252, 373-383.
- (66) Langlois, M. (2003) Hammerhead ribozyme-mediated destruction of nuclear foci in myotonic dystrophy myoblasts, *Mol. Ther.* 7, 670-680.
- (67) Foff, E. P., and Mahadevan, M. S. (2011) Therapeutics development in myotonic dystrophy type 1., *Muscle Nerve* 44, 160-169.
- (68) Soifer, H. S., Rossi, J. J., and Saetrom, P. (2007) MicroRNAs in disease and potential therapeutic applications, *Mol. Ther.* 15, 2070-2079.
- (69) Krol, J., Fiszler, A., Mykowska, A., Sobczak, K., de Mezer, M., and Krzyzosiak, W. J. (2007) Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets, *Mol. Cell* 25, 575-586.
- (70) Moulton, H. M., and Moulton, J. D. (2010) Morpholinos and their peptide conjugates: therapeutic promise and challenge for Duchenne muscular dystrophy, *Biochim. Biophys. Acta* 1798, 2296-2303.
- (71) Mulders, S. A., van den Broek, W. J., Wheeler, T. M., Croes, H. J., van Kuik-Romeijn, P., de Kimpe, S. J., Furling, D., Platenburg, G. J., Gourdon, G., Thornton, C. A., Wieringa, B., and Wansink, D. G. (2009) Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A.* 106, 13915-13920.
- (72) Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., and Thornton, C. A. (2009) Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA, *Science* 325, 336-339.
- (73) Disney, M. D., Labuda, L. P., Paul, D. J., Poplawski, S. G., Pushechnikov, A., Tran, T., Velagapudi, S. P., Wu, M., and Childs-Disney, J. L. (2008) Two-dimensional

- combinatorial screening identifies specific aminoglycoside-RNA internal loop partners., *J. Am. Chem. Soc.* 130, 11185-11194.
- (74) Tran, T., and Disney, M. D. (2010) Two-dimensional combinatorial screening of a bacterial rRNA A-site-like motif library: defining privileged asymmetric internal loops that bind aminoglycosides, *Biochemistry* 49, 1833-1842.
- (75) Cho, J., and Rando, R. R. (2000) Specific binding of Hoechst 33258 to site 1 thymidylate synthase mRNA, *Nucleic Acids Res.* 28, 2158-2163.
- (76) Warf, M. B., Nakamori, M., Matthys, C. M., Thornton, C. A., and Berglund, J. A. (2009) Pentamidine reverses the splicing defects associated with myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A.* 106, 18551-18556.
- (77) Ofori, L. O., Hoskins, J., Nakamori, M., Thornton, C. A., and Miller, B. L. (2012) From dynamic combinatorial 'hit' to lead: in vitro and in vivo activity of compounds targeting the pathogenic RNAs that cause myotonic dystrophy, *Nucleic Acids Res.* 40, 6380-6390.
- (78) Parkesh, R., Childs-Disney, J. L., Nakamori, M., Kumar, A., Wang, E., Wang, T., Hoskins, J., Tran, T., Housman, D., Thornton, C. A., and Disney, M. D. (2012) Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching, *J. Am. Chem. Soc.* 134, 4731-4742.
- (79) Pushechnikov, A., Lee, M. M., Childs-Disney, J. L., Sobczak, K., French, J. M., Thornton, C. A., and Disney, M. D. (2009) Rational design of ligands targeting triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3, *J. Am. Chem. Soc.* 131, 9767-9779.
- (80) Childs-Disney, J. L., Hoskins, J., Rzuczek, S. G., Thornton, C. A., and Disney, M. D. (2012) Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potentially bioactive, *ACS Chem. Biol.* 7, 856-862.
- (81) Arambula, J. F., Ramisetty, S. R., Baranger, A. M., and Zimmerman, S. C. (2009) A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding, *Proc. Natl. Acad. Sci. U. S. A.* 106, 16068-16073.
- (82) Wang, G. S., Kuyumcu-Martinez, M. N., Sarma, S., Mathur, N., Wehrens, X. H., and Cooper, T. A. (2009) PKC inhibition ameliorates the cardiac phenotype in a mouse model of myotonic dystrophy type 1, *J. Clin. Invest.* 119, 3797-3806.

- (83) Logigian, E. L., Martens, W. B., Moxley, R. T. t., McDermott, M. P., Dilek, N., Wiegner, A. W., Pearson, A. T., Barbieri, C. A., Annis, C. L., Thornton, C. A., and Moxley, R. T., 3rd. (2010) Mexiletine is an effective antimyotonia treatment in myotonic dystrophy type 1, *Neurology* 74, 1441-1448.
- (84) Heatwole, C. R., Eichinger, K. J., Friedman, D. I., Hilbert, J. E., Jackson, C. E., Logigian, E. L., Martens, W. B., McDermott, M. P., Pandya, S. K., Quinn, C., Smirnow, A. M., Thornton, C. A., and Moxley, R. T., 3rd. (2011) Open-label trial of recombinant human insulin-like growth factor 1/recombinant human insulin-like growth factor binding protein 3 in myotonic dystrophy type 1, *Arch. Neurol.* 68, 37-44.
- (85) Vlachopapadopoulou, E., Zachwieja, J. J., Gertner, J. M., Manzione, D., Bier, D. M., Matthews, D. E., and Slonim, A. E. (1995) Metabolic and clinical response to recombinant human insulin-like growth factor I in myotonic dystrophy--a clinical research center study, *J. Clin. Endocrinol. Metab.* 80, 3715-3723.
- (86) O'Rourke, J. R., and Swanson, M. S. (2009) Mechanisms of RNA-mediated disease, *J. Biol. Chem.* 284, 7419-7423.
- (87) Echeverria, G. V., and Cooper, T. A. (2012) RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity, *Brain Res.* 1462, 100-111.
- (88) Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature* 411, 494-498.
- (89) Wheeler, T. M., Leger, A. J., Pandey, S. K., MacLeod, A. R., Nakamori, M., Cheng, S. H., Wentworth, B. M., Bennett, C. F., and Thornton, C. A. (2012) Targeting nuclear RNA for in vivo correction of myotonic dystrophy, *Nature* 488, 111-115.
- (90) Garcia-Lopez, A., Llamusi, B., Orzaez, M., Perez-Paya, E., and Artero, R. D. (2011) In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models, *Proc. Natl. Acad. Sci. U. S. A.* 108, 11866-11871.
- (91) Guan, L., and Disney, M. D. (2012) Recent advances in developing small molecules targeting RNA, *ACS Chem. Biol.* 7, 73-86.
- (92) Guan, L., and Disney, M. D. (2013) Small-molecule-mediated cleavage of RNA in living cells, *Angew. Chem. Int. Ed. Engl.* 52, 1462-1465.

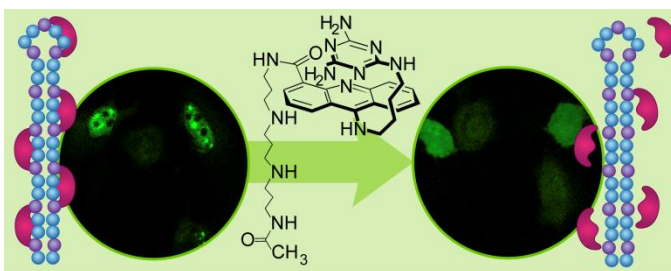
# Chapter 2.<sup>1</sup>

## A Novel CUG<sup>exp</sup>·MBNL1 Inhibitor with Therapeutic Potential for Myotonic Dystrophy Type 1

### 2.1 Abstract

Myotonic dystrophy type 1 (DM1) is caused by an expanded CUG repeat (CUG<sup>exp</sup>) that sequesters muscleblind-like 1 protein (MBNL1), a protein that regulates alternative splicing. CUG<sup>exp</sup> RNA is a validated drug target for this currently untreatable disease.

Herein, we develop a bioactive small molecule (1) that targets CUG<sup>exp</sup> RNA and is able to inhibit the CUG<sup>exp</sup>·MBNL1 interaction in cells that model DM1. The core of this small



molecule is based on ligand 2, which was previously reported to be active in an in vitro assay. A polyamine-derivative side chain was conjugated to this core to make it aqueous-soluble and cell penetrable. In a DM1 cell model this conjugate was found to disperse CUG<sup>exp</sup> ribonuclear foci, release MBNL1, and partially reverse the mis-splicing of the insulin receptor pre-mRNA. Direct evidence for ribonuclear foci dispersion by this ligand was obtained in a live DM1 cell model using time-lapse confocal microscopy.

### 2.2 Introduction

RNA is an important, yet underutilized, drug target. To date, the most common RNA drug targets have been ribosomal RNA and HIV RNA.<sup>1-3</sup> With recent structural and functional discoveries, non-coding RNA is gradually becoming an attractive drug target<sup>4-6</sup> and much is now known about designing ligands to interact with RNA.<sup>7-9</sup> Myotonic dystrophy (dystrophia myotonica, DM) is among the pathologies where RNA stands as the most appropriate target for drug discovery.<sup>10</sup> DM is the most common adult muscular

<sup>1</sup> Adapted with permission from ACS Chemical Biology. DOI: 10.1021/cb400046u

dystrophy with a prevalence of 1:8,000 to 1:20,000 worldwide.<sup>11</sup> Currently there is no treatment for DM, only palliative therapy.<sup>12</sup>

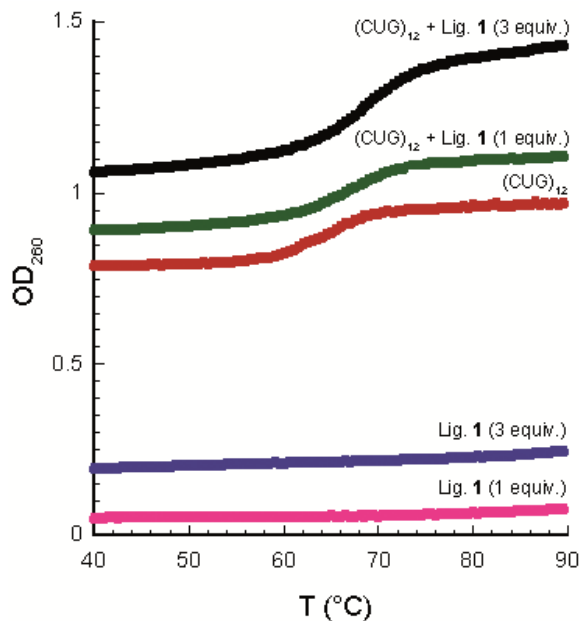
Myotonic dystrophy type 1 (DM1), originates from the progressive expansion of CTG repeats in the 3'- untranslated region of the *DMPK* gene. Thus, expanded CUG repeat transcripts (CUG<sup>exp</sup>) are the known causative agent of DM1.<sup>13, 14</sup> The CUG<sup>exp</sup> RNA manifests its toxicity through a gain-of-function mechanism involving the sequestration of all three paralogs of human MBNL including MBNL1, a key regulatory protein of alternative splicing.<sup>15-17</sup> The MBNL1·CUG<sup>exp</sup> aggregate forms ribonuclear foci, a hallmark of DM1 cells.<sup>18</sup> In a mouse model of DM1, a morpholino antisense oligonucleotide (ASO),<sup>19</sup> 2'-O-(2-methoxyethyl) ASO,<sup>20</sup> and D-amino acid hexapeptide, each targeting CUG<sup>exp</sup>, rescued the mis-splicing and reversed the phenotype.<sup>21</sup> These studies validated CUG<sup>exp</sup> as a drug target and greatly increased interest in finding small molecules that function similarly. Pentamidine,<sup>22</sup> benzo[g]quinolone-based heterocycles,<sup>23</sup> a Hoechst derivative (H1),<sup>24</sup> a modularly assembled Hoechst 33258,<sup>25, 26</sup> and ligand **2**, reported by our laboratory,<sup>27</sup> are examples of bioactive CUG repeat binders at various stages of development as potential therapeutic agents for DM1.

Our previously reported approach, which led to ligand **2** as a binder of CUG, was based on the notion that selectivity was paramount and could be achieved by rational design focusing on recognition of the UU mismatch in double stranded CUG<sup>exp</sup>.<sup>26</sup> We found that the triaminotriazine ring (recognition unit) has a key role in the inhibition of (CUG)<sub>12</sub>·MBNL1 interaction as several acridine derivatives that lacked this unit showed no inhibition potency in our *in vitro* assay (Arambula, J. Ph.D. Thesis, University of Illinois, 2008). Although **2** proved to be among the most selective and effective inhibitors of the (CUG)<sub>12</sub>·MBNL1 interaction, despite its *in vitro* activity, it was not active in a cellular model of DM1. Its drugability was limited both because of its low water solubility and its inability to penetrate the cellular membrane. Herein we report further development of this small molecule into an active ligand *in vivo* through its conjugation to a cationic polyamine and the first observation using time-lapse confocal microscopy of foci dispersion in live cells that model DM1.

## 2.3 Results and Discussion

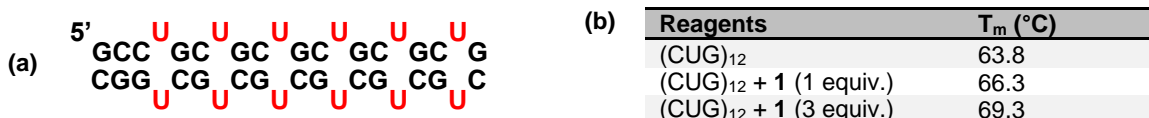
**Synthesis and Characterization.** Ligand **1** (Figure 2.1) is a conjugate of the previously reported *in vitro* active ligand **2** (Figure 2.1) and N-[3-({3-[(3-aminopropyl)amino] propyl}amino)propyl] acetamide side chain. The synthesis scheme





**Figure 2.2.** Melting curves of  $(\text{CUG})_{12}$  in the absence and presence of one and three equivalents of **1**;  $\Delta T_m$  of  $2.5 \pm 0.2$  °C and  $5.5 \pm 0.8$  °C were observed, respectively. Note: The assay was performed as described in the method section, however the flat part of the curve (10°C to 40°C) has been cut out and only the part of curve from 40°C to 90°C is shown here.

This finding indicates binding of **1** to  $(\text{CUG})_{12}$  and stabilization of the double stranded (ds)  $(\text{CUG})_{12}$  hairpin. The latter finding is important because it has been proposed that MBNL1 displays a preference for single stranded (ss) RNA.<sup>38, 39</sup> If this model is correct, any ligand that stabilizes the ds form of  $\text{CUG}^{\text{exp}}$  may prove to be a more effective inhibitor of the  $(\text{CUG})_{12}$ ·MBNL1 interaction. Thus, the observed stabilization of the ds form of  $(\text{CUG})_{12}$  was an encouraging result, although not sufficient to ensure selective and effective inhibition of  $(\text{CUG})_{12}$ ·MBNL1 interaction.



**Figure 2.3.** Ligand **1** stabilizes the ds form of  $(\text{CUG})_{12}$ . a) Schematic representation of  $(\text{CUG})_{12}$ . b)  $T_m$  of  $(\text{CUG})_{12}$  hairpin in the presence of **1** and 3 equivalents of **1** in 1X PBS buffer. Values were measured in duplicate or triplicate with repeats agreeing within 1%.

**Inhibition of  $(\text{CUG})_{12}$ ·MBNL1 Interaction by **1**.** To our knowledge, SPR has not previously been used to characterize the MBNL1-CUG interaction or its inhibition by small molecules.<sup>23, 40</sup> Because the technique is particularly well suited for quantifying the

binding of proteins to a target on the SPR chip, we developed a simple SPR-based method to directly measure MBNL1 complexation of (CUG)<sub>12</sub> in real time under equilibrium conditions and in a label-free format. Further, we were able to quantify the inhibition potency of **1** and its selectivity. The selectivity was assessed by performing the assay in the presence of a large excess of competitor tRNA.

Thus, biotinylated (CUG)<sub>12</sub> was immobilized on a streptavidin coated SPR sensor chip and incubated with different concentrations of **1** to reach a steady state response (response units, RU, see Figure 2.4a) over 150 s. The response to the binding of **1** was negligible in comparison to protein binding so the direct contribution of **1** could be ignored. Successive injections of a 0.65  $\mu$ M solution of MBNL1 containing the same concentration of **1** as in the pre-incubation, led to varying responses depending on the concentration of **1**. Because the SPR signal directly reflects the binding of MBNL1 to the biotinylated (CUG)<sub>12</sub>, the differences in the response curves are a direct result of inhibition by **1**.

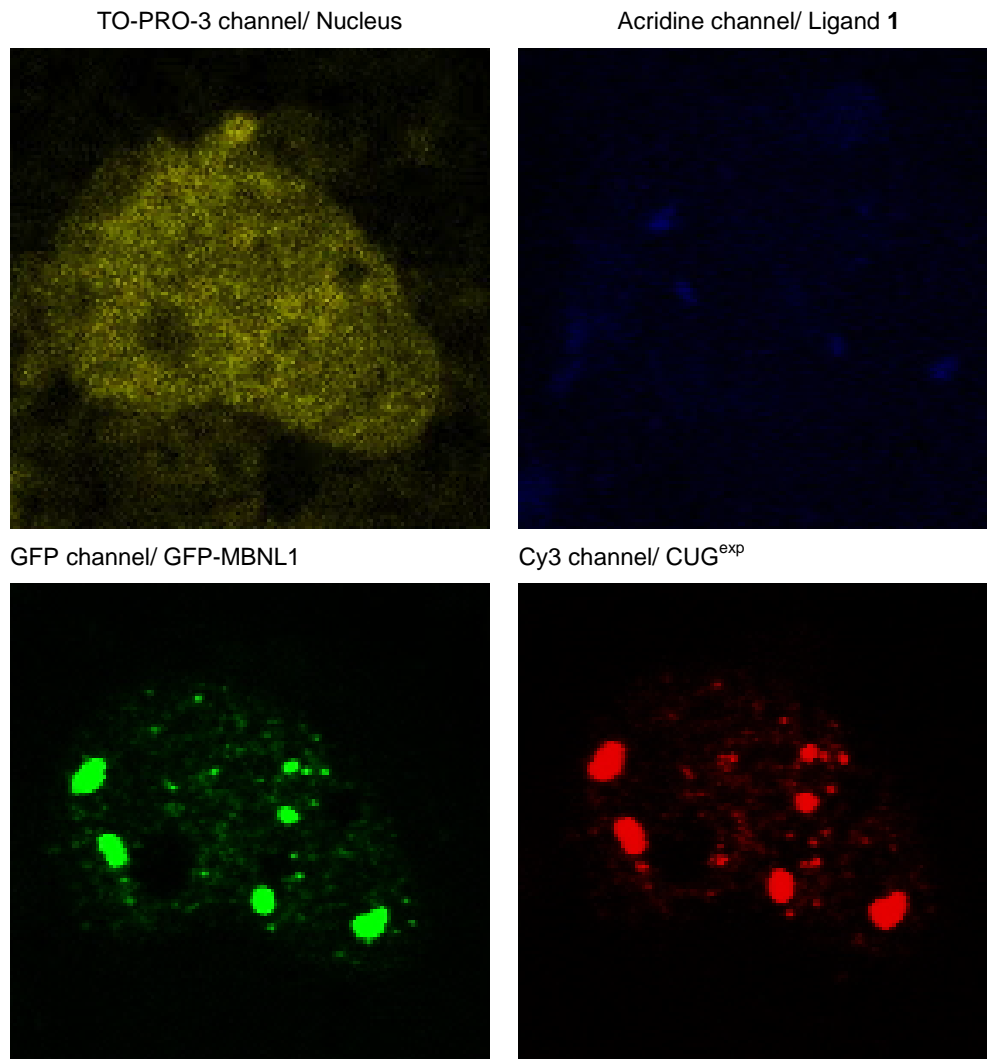
The curves recorded in the presence and absence of 580  $\mu$ M yeast tRNA were identical, indicating selective inhibition by **1**. All of the data shown herein were from runs in the presence of tRNA. The maximum RU at 150 s was recorded for each concentration of **1** and converted to the fraction of (CUG)<sub>12</sub> bound by MBNL1, all values normalized to that measured in the absence of **1**. Fitting the data points in the plot of % (CUG)<sub>12</sub> bound by MBNL1 versus increasing concentrations of **1** (Figure 2.4b) gave an apparent IC<sub>50</sub> value of  $15 \pm 2$   $\mu$ M.

These *in vitro* experiments demonstrate that **1** binds to (CUG)<sub>12</sub>, stabilizing the hairpin structure and inhibiting (CUG)<sub>12</sub>·MBNL1 interaction selectively. It is noteworthy that all of the *in vitro* experiments above were carried out with (CUG)<sub>12</sub> in 1X PBS buffer. This particular buffer was chosen because it is the closest of common buffers to physiological conditions. It is also a more challenging buffer for small molecule inhibitors because it increases the (CUG)<sub>12</sub>·MBNL1 stability, as we obtained a K<sub>D</sub> value of  $5.2 \pm 2.5$  nM for (CUG)<sub>12</sub>·MBNL1 interaction in this buffer by SPR technique whereas using EMSA we and others had reported K<sub>D</sub> values of 26 and 170 nM, respectively.<sup>27, 37</sup>

**Bioactivity in DM1 Cell Model.** It is a characteristic of DM1 cells that MBNL1 aggregates with CUG<sup>exp</sup> in nuclear foci.<sup>41</sup> To visualize the effect of **1** on these ribonuclear foci, we used confocal microscopy. This was accomplished using a model for a DM1 cell. Thus, HeLa cells were transfected with two plasmids, truncated DMPK-CUG<sub>960</sub> and



GFP-MBNL1.<sup>42</sup> As a negative control, HeLa cells were transfected with truncated DMPK-CUG<sub>0</sub> (i.e., no CUG repeat) and GFP-MBNL1 plasmids.<sup>42</sup> To detect (CUG)<sub>960</sub> foci, Cy3-(CAG)<sub>10</sub> was used as a fluorescence *in situ* hybridization (FISH) probe. TO-PRO-3 was used to stain the nucleus, this particular dye was chosen because it has no overlap with the other three fluorophores in our system, i.e., the acridine ring of **1**, the GFP of GFP-MBNL1 and the Cy3 of the FISH probe. By taking advantage of the acridine fluorescence, the penetration of **1** to the cytoplasm as well as nucleus is tracked in the ligand-treated cells.



**Figure 2.4.** A magnified CUG<sub>960</sub> transfected HeLa cell (DM1 cell model) showing ribonuclear foci.

Representative images from the confocal microscopy are shown in Figure 2.5a. The negative control cells lacking the (CUG)<sub>960</sub> sequence showed no foci but rather MBNL1 dispersed throughout the nucleus (Figure 2.5a, row 1). However, co-localized MBNL1 and (CUG)<sub>960</sub> foci were observed in the DM1 cell model (Fig 2.5a, row 2). Thus, in the untreated DM1 cells the merged GFP-MBNL1 (green fluorescence) and Cy3-(CAG)<sub>10</sub> (red fluorescence) images, showed yellow spots that correspond to MBNL1 and (CUG)<sub>960</sub> co-localization in nuclear foci (last column Figure 2.5a). Likewise, incubation of the DM1 cell model with a negative control compound, 50  $\mu$ M spermine for 48 h had no effect on the foci (Figure 2.5a, row 3).

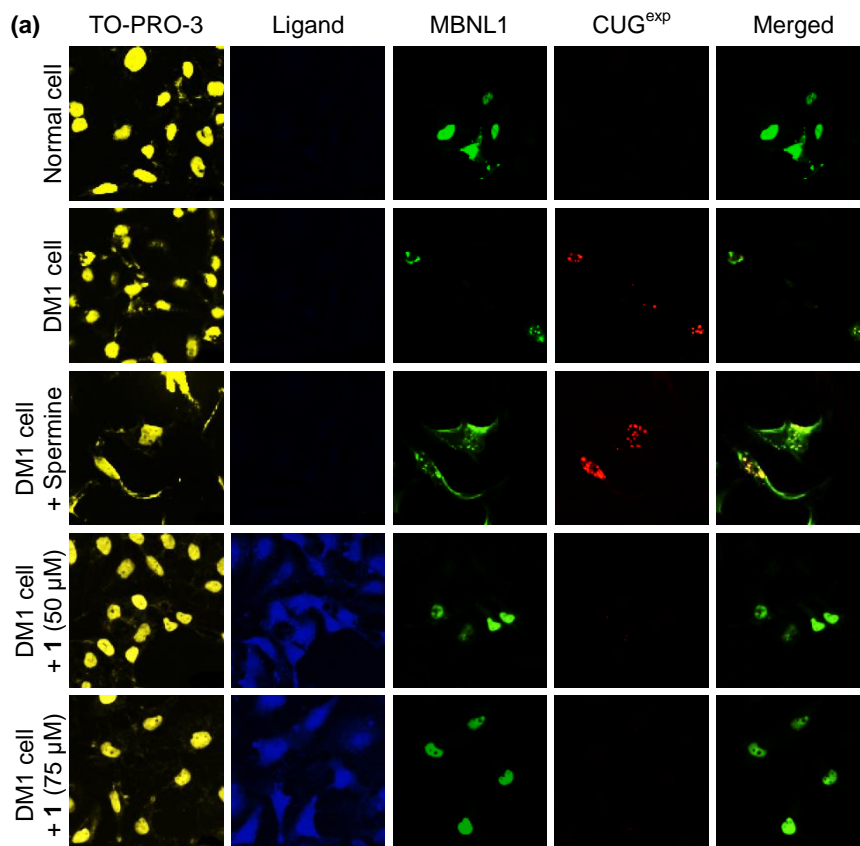
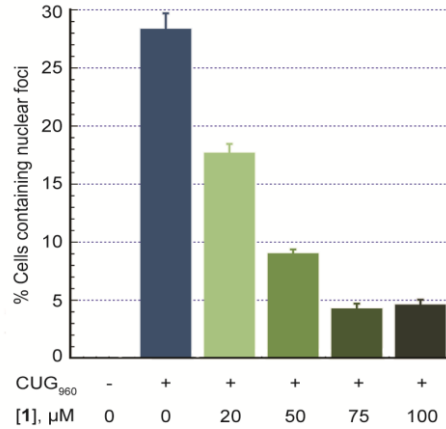


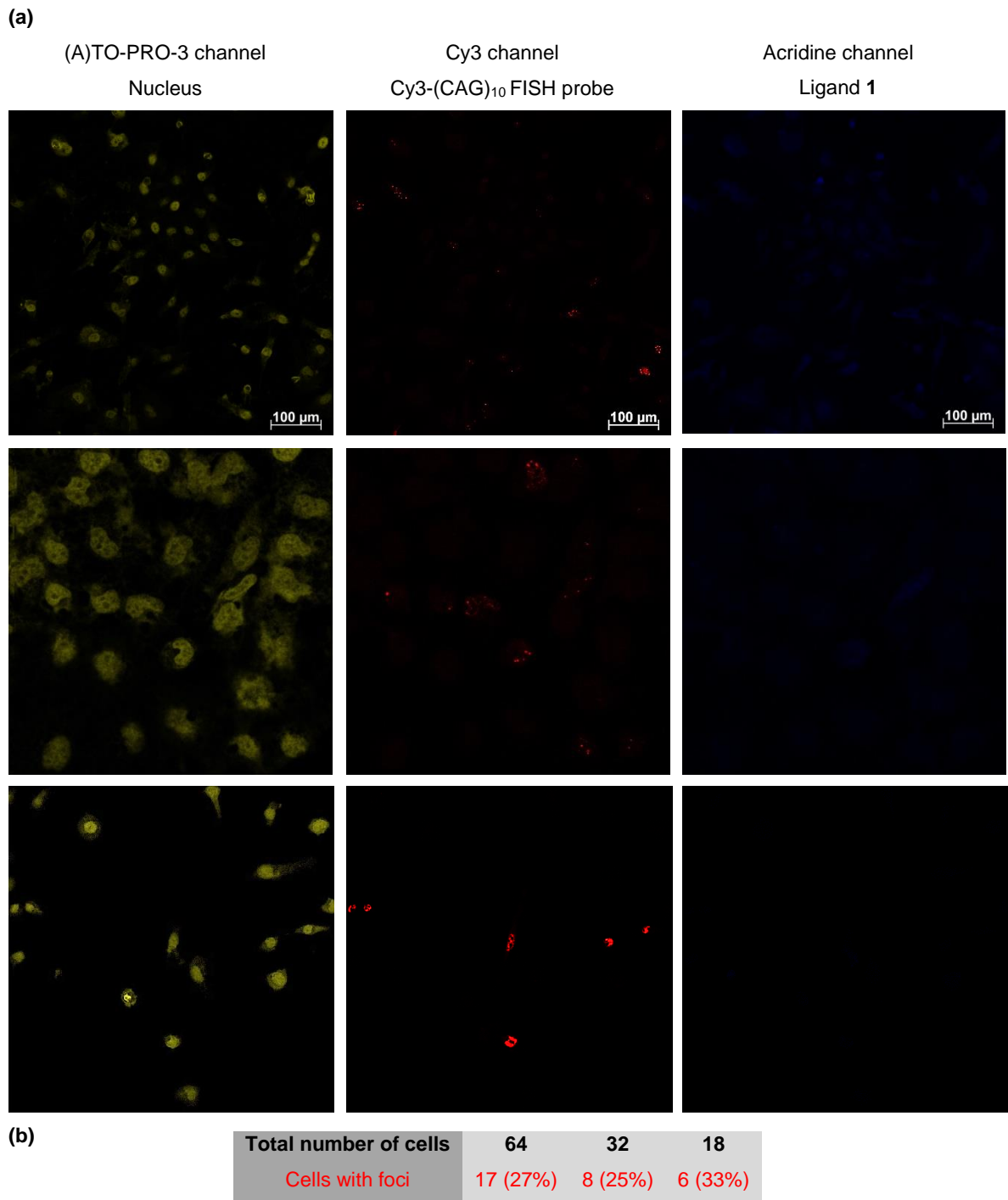
Fig 2.5 (cont. on next page)

(b)

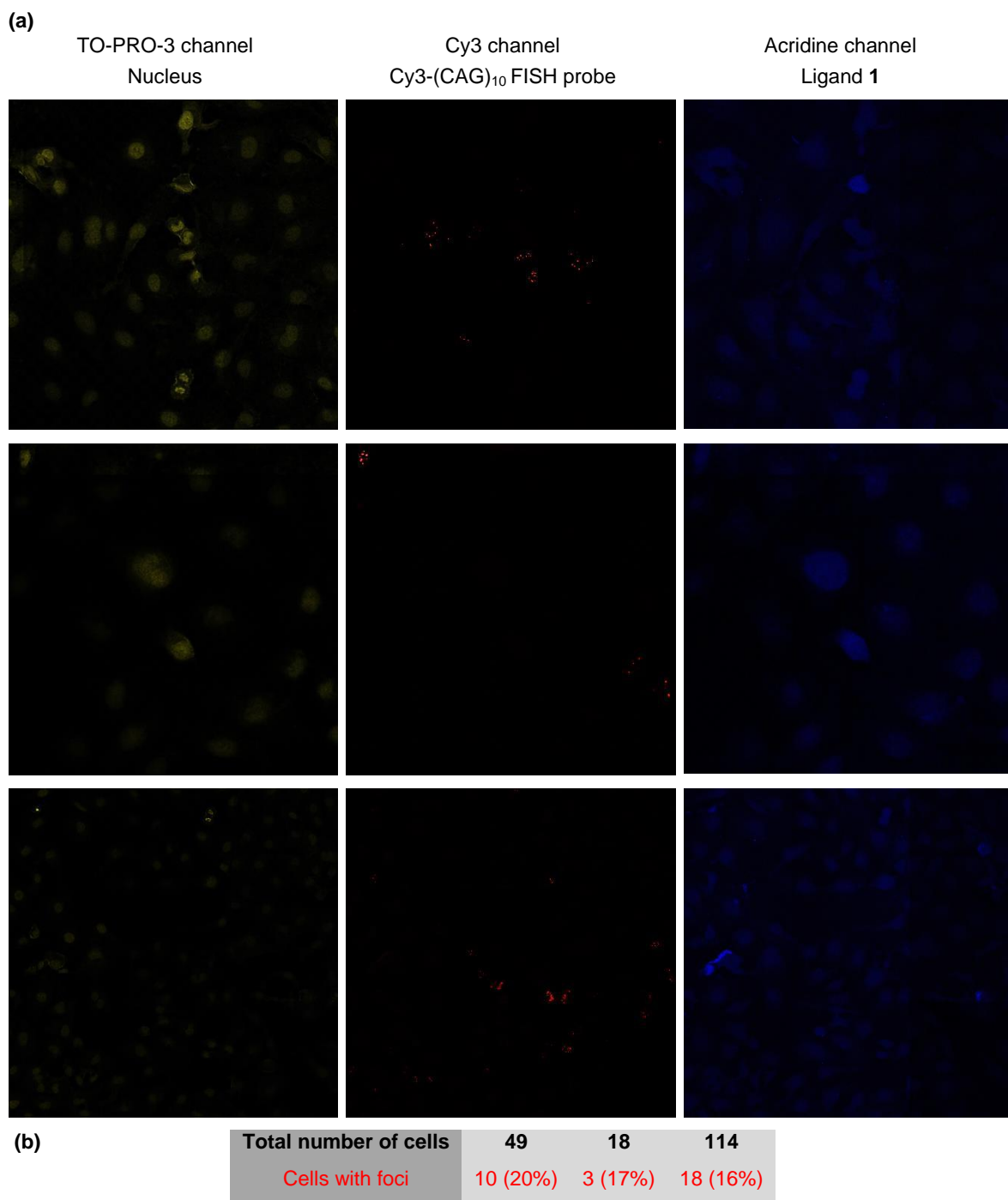


**Figure 2.5.** Ligand **1** disrupts nuclear foci in fixed DM1 cell model. a) Columns 1-4 as labeled. Column 5 is merge of columns 3 and 4. Confocal fluorescent images show MBNL1 and CUG<sup>exp</sup> foci are present in row 2 where no ligand is added as well as row 3 where spermine (50  $\mu\text{M}$ ), as a negative control compound, is added. CUG<sup>exp</sup> foci are not visible and MBNL1 is dispersed across the nucleus in negative control cells, row 1, as well as rows 4 and 5 where DM1 cell model is treated with **1** at 50 and 75  $\mu\text{M}$ , respectively, for 48 h. Each box shows a 150  $\mu\text{m} \times 150 \mu\text{m}$  area. b) Plot of CUG<sup>exp</sup> foci-containing cell fraction at various concentrations of **1**. These data are gathered from scoring over 100 cells. The error bars represent mean  $\pm$  standard error of at least three independent experiments. A magnified CUG<sub>960</sub> transfected HeLa cell (DM1 cell model) showing ribonuclear foci is shown in Figure 2.4.

However, incubation with **1**, at 50 and 75  $\mu\text{M}$  caused almost complete disappearance of the (CUG)<sub>960</sub> foci and dispersion of the MBNL1 fluorescence (Figure 2.5a, rows 4 and 5, respectively). The foci disruption is observed as a disappearance, rather than dispersion of the FISH probe, because it is an exogenous antisense nucleic acid probe only visible when concentrated by a high-localized concentration of CUG<sup>exp</sup> RNA.<sup>43</sup> Cells were classified as having or not having foci. The fraction of cells with (CUG)<sub>960</sub> foci was reduced by ca. 86% with **1** at 75  $\mu\text{M}$  with a two-tailed  $p$ -value of 0.004 (Figure 2.5b). Similar responses are seen at 50 and 100  $\mu\text{M}$  ligand.

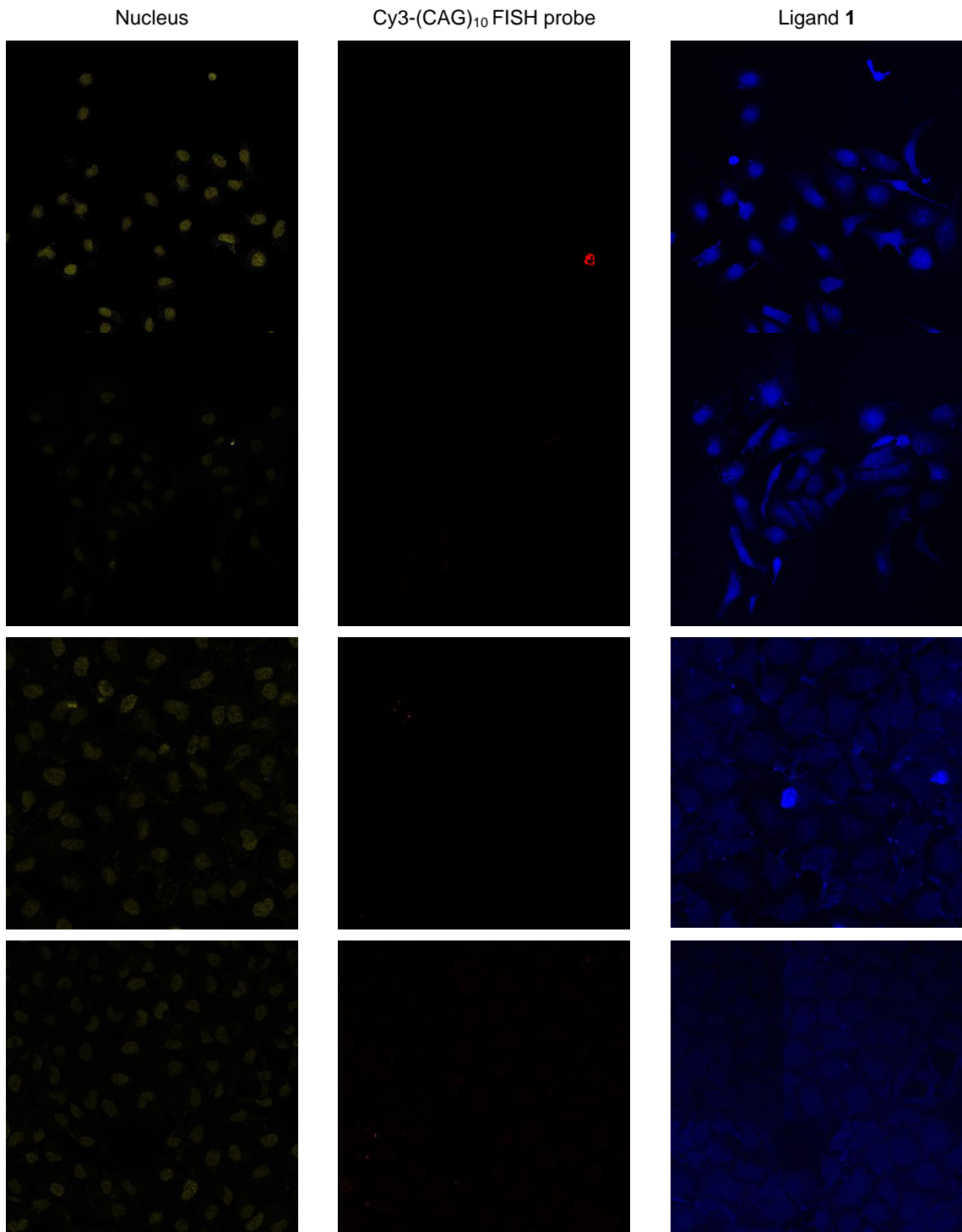


**Figure 2.6.** CUG<sub>960</sub> transfected HeLa cells (DM1 cell model) without ligand treatment. a) Three visible excitation channels (639, 555 and 408 nm) were used to detect To-Pro-3, Cy3 and acridine ring, respectively. These channels indicate the presence of nuclei, Cy3-(CAG)<sub>10</sub> FISH probe and **1**, respectively. b) 28% of DM1 model cells show CUG<sup>exp</sup> nuclear foci in DM1 cell model.



**Figure 2.7.** CUG<sub>960</sub> transfected HeLa cells (DM1 cell model) treated with [1] = 20 μM. a) Three visible excitation channels (639, 555 and 408 nm) were used to detect To-Pro-3, Cy3 and acridine ring, respectively. These channels indicate the presence of nuclei, Cy3-(CAG)<sub>10</sub> FISH probe and 1, respectively. b) 18% of DM1 model cells show CUG<sup>exp</sup> nuclear foci in the presence of [1] = 20 μM.

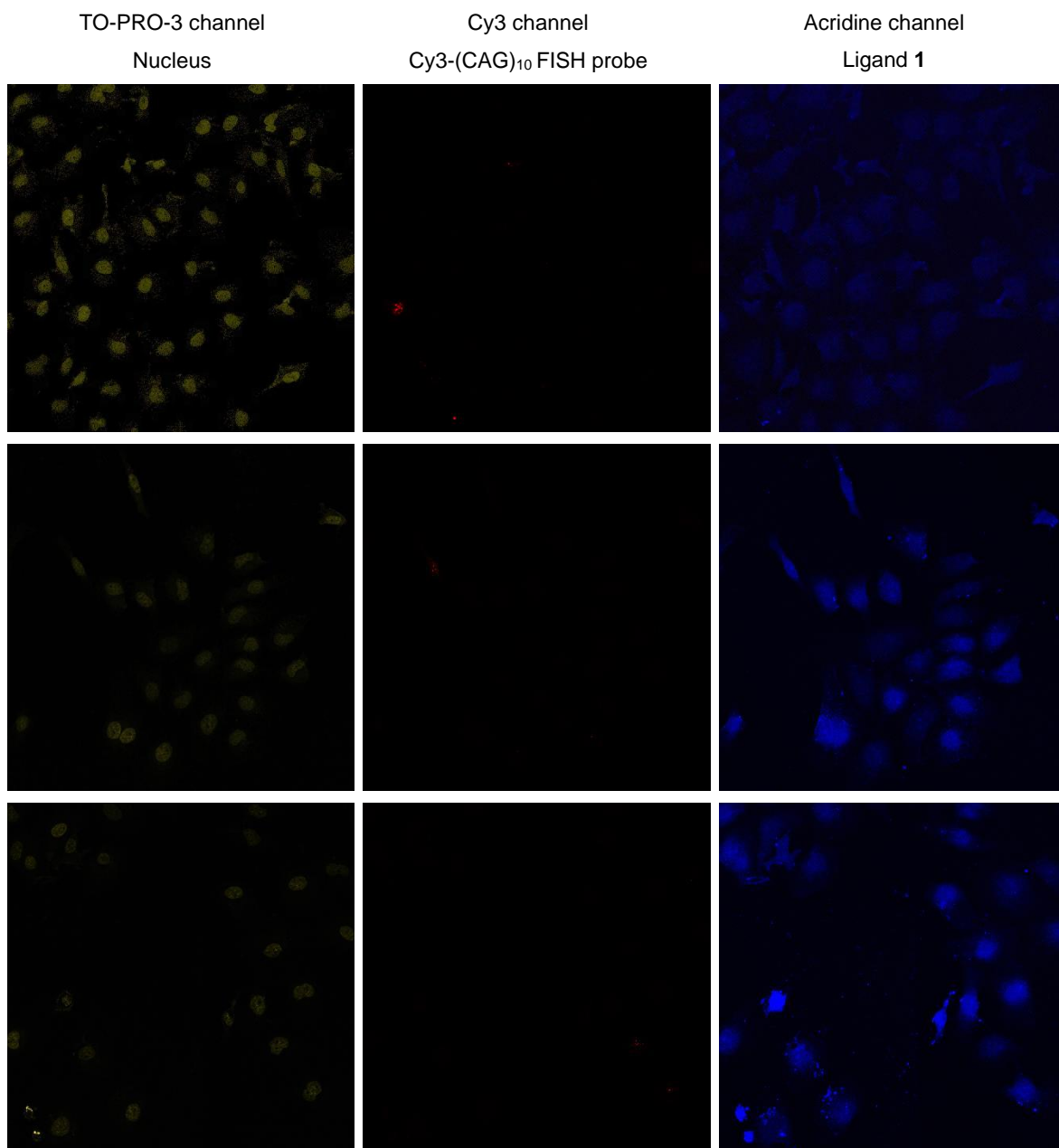




| (b) | Total number of cells | 30     | 32     | 63     | 94     |
|-----|-----------------------|--------|--------|--------|--------|
|     | Cells with foci       | 1 (3%) | 2 (6%) | 2 (3%) | 5 (5%) |

**Figure 2.9.** CUG<sub>960</sub> transfected HeLa cells (DM1 cell model) treated with [1] = 75 μM. a) 4% of DM1 model cells show CUG<sup>exp</sup> nuclear foci in the presence of [1] = 75 μM.

(a)



(b)

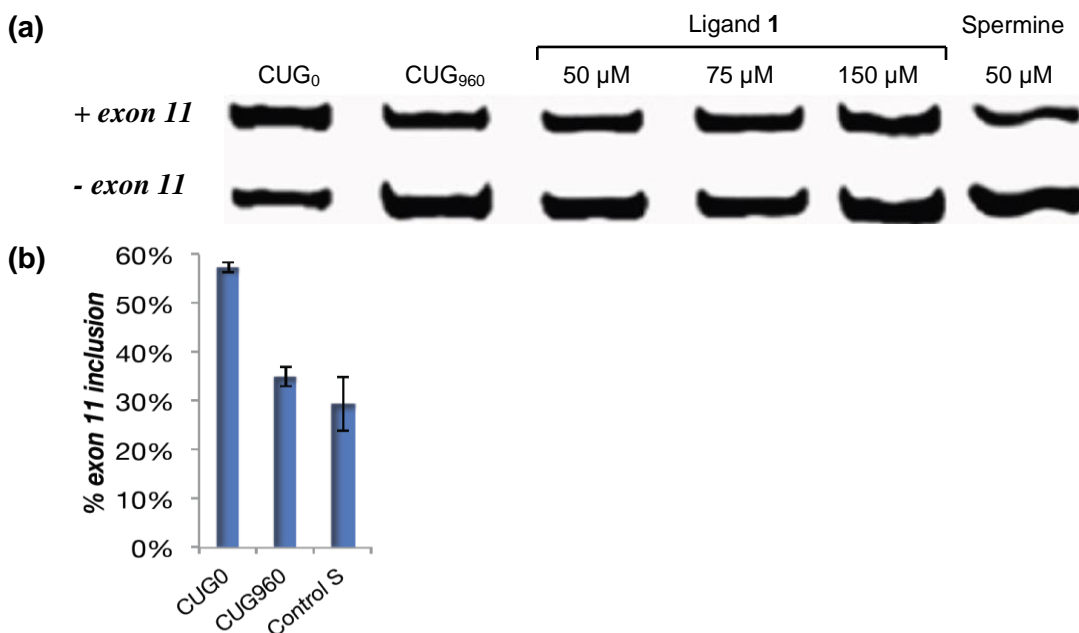
| Total number of cells | 50     | 27     | 26     |
|-----------------------|--------|--------|--------|
| Cells with foci       | 3 (6%) | 1 (4%) | 1 (4%) |

**Figure 2.10.** CUG<sub>960</sub> transfected HeLa cells (DM1 cell model) treated with [1] = 100 μM. a) Three visible excitation channels (639, 555 and 408 nm) were used to detect To-Pro-3, Cy3 and acridine ring, respectively. These channels indicate the presence of nuclei, Cy3-(CAG)<sub>10</sub> FISH probe and 1, respectively. b) 5% of DM1 model cells show CUG<sup>exp</sup> nuclear foci in the presence of [1] = 100 μM.

**Improvement of Pre-mRNA Mis-Splicing by 1.** After confirming that 1 disperses the MBNL1 foci, we sought to study the alternative splicing as a downstream measure of



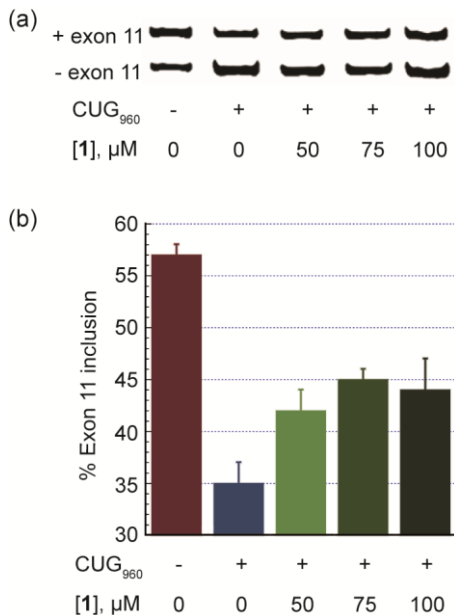
recovered MBNL1 regulatory activity. MBNL1 is known to be a key regulatory protein in alternative splicing and affecting many pre-mRNAs, including the insulin receptor (IR).<sup>44</sup> The mis-splicing of IR in DM1 cells occurs with a predominance of isoform A (with exon 11 exclusion) relative to isoform B (with exon 11 inclusion).<sup>45</sup> As described above, truncated DMPK mRNAs containing (CUG)<sub>960</sub> or no CUG repeats were expressed in HeLa cells to serve as our DM1 cell model and negative control cell, respectively. These cells were co-transfected with an IR mini-gene to study the regulation of splicing of IR by measuring the relative amounts of its two isoforms. Looking at the transcripts in the DM1 cell model showed mis-splicing of IR with  $35 \pm 2$  % isoform B (IR-B), whereas  $57 \pm 1$  % of IR-B, measured in the negative control cell, was considered the baseline exon inclusion. Treatment of the DM1 cell model with a negative control compound, 50  $\mu$ M spermine, had no effect on the IR mis-splicing (Figure 2.11).



**Figure 2.11.** Negative control ligand, spermine, at 50  $\mu$ M does not improve the mis-splicing of IR in DM1 cell model. a) In this representative gel image of IR alternative splicing, two bands corresponding to IR isoforms A (+ exon 11) and B (- exon 11), respectively, are derived by reverse transcription-polymerase chain reaction (RT-PCR). DM1 cell model is treated with **1** at 50, 75 and 150  $\mu$ M and spermine at 50  $\mu$ M. Note: This image is derived from one gel but the lanes of interest have been cut and positioned adjacent to one another. b) A plot of the corresponding data shows no change in mis-splicing with spermine at 50  $\mu$ M. The error bars represent mean  $\pm$  standard error of 4-6 independent measurements for CUG<sub>0</sub> and CUG<sub>960</sub> cells and 3 measurements for spermine-treated cells.

The splicing assay was repeated with different concentrations of **1** to see if it was capable of rescuing the mis-splicing of IR, i.e., increasing the fraction of IR-B (Figure

2.12a).<sup>46</sup> Thus, the DM1 model cells were treated with **1** at 50, 75 or 150  $\mu\text{M}$  for 48 h. A rescue of 40% for the IR splicing defect was observed at 75  $\mu\text{M}$  with  $45 \pm 1\%$  IR-B measured; two-tailed  $p$ -values of 0.002 (Figure 2.12b).



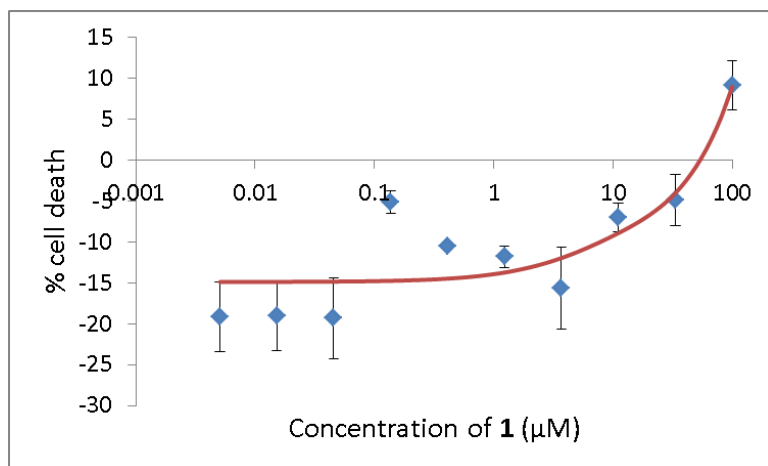
**Figure 2.12.** Ligand **1** improves mis-splicing of IR in DM1 cell model. a) A representative gel image of IR alternative splicing. Two bands corresponding to IR isoforms A (+ exon 11) and B (- exon 11), respectively, are derived by reverse transcription-polymerase chain reaction (RT-PCR). DM1 cell model is treated with **1** at 50, 75 and 150  $\mu\text{M}$ . b) A plot of the corresponding data shows 40% rescue of mis-splicing at [**1**] = 75  $\mu\text{M}$ . The error bars represent mean  $\pm$  standard error of 4-6 independent measurements.

Similar responses are seen at 50 and 150  $\mu\text{M}$  ligand. A more quantitative approach would be needed to demonstrate a dose response. It is noteworthy that cytotoxicity of **1** was evaluated by sulforhodamine B assay and less than 10% HeLa cell death was observed at the highest tested concentration (100  $\mu\text{M}$ ) after 24 h (Figure 2.13). Cytotoxicity assay was performed according to a previous method.<sup>47</sup> 98  $\mu\text{L}$  of DMEM supplemented 10% FBS per well was placed in a 96-well plate. 2  $\mu\text{L}$  of **1** with different concentrations from 0.5  $\mu\text{M}$  to 10 mM were added, five repeats for each concentration. 100  $\mu\text{L}$  of HeLa cell suspension of 10,000 cells per well were added in the 96-well plate. The cells were incubated at 37  $^{\circ}\text{C}$  for 24 h. 100  $\mu\text{L}$  of cold 10% (wt/v) trichloroacetic acid were added in each well, followed by incubation at 4  $^{\circ}\text{C}$  for 1 h. The cells were washed twice with water and then air-dried. The cells were stained with 100  $\mu\text{L}$  of 0.0057 % (wt/v) sulforhodamine B in 1% acetic acid at room temperature for 30 min. The plate was rinsed twice with 1% (v/v) acetic acid to remove unbound stain. The bound protein stain

was solubilized in 200  $\mu\text{L}$  of 10 mM Tris base, pH 10.5, and left for 30 min. The optical intensity was measured at 510 nm using a microplate reader. The cell growth inhibition was calculated using the following formula:

$$\% \text{ Cell growth inhibition} = 100\% - \frac{\text{Mean}_{\text{OD sample}} - \text{Mean}_{\text{OD dead control}}}{\text{Mean}_{\text{OD live control}} - \text{Mean}_{\text{OD dead control}}} \times 100\%$$

The data was plotted using Excel. The curves were fitted using Table Curve (Systat).



**Figure 2.13.** Ligand **1** shows less than 10% cytotoxicity in concentrations up to 100  $\mu\text{M}$  by the sulforhodamine B (SRB) assay.

**MBNL1 Foci Dispersion in Live Cells by **1**.** Although the reduction in the foci-containing fixed cells was statistically significant, that experiment represents an indirect measurement of the foci dispersion; because with fixed-cell microscopy, we are not following the same cell directly over time to observe the dispersion of ribonuclear foci upon addition of **1**. Thus, to further confirm our observation, we sought to investigate the effect of **1** in a live cell, by time-lapse confocal microscopy.

To monitor drug uptake and foci dispersion in real time, model DM1 cells were incubated with **1** at 75  $\mu\text{M}$  and individual live cells were monitored by confocal microscopy at several time points. The first observation at time point,  $t = 0$ , was made immediately before addition of **1** and MBNL1 nuclear foci were clearly present (Figure 2.14a,  $t = 0$ ). We found it necessary to use a Petri dish with an imprinted 500  $\mu\text{m}$  grid to relocate the cell following the incubation interval.

The ability to monitor the location of **1** was made possible by the inherent fluorescence of the acridine unit. Over time, **1** penetrated the cellular and nuclear membrane and the MBNL1 foci gradually dispersed over the entire nucleus (Figure 2.14a,  $t = 2, 4$  and 7 h). The intensity of a representative MBNL1 focus (the most intense focus) at these four

time points shows this dispersion (Figure 2.14b). A negative control experiment was performed at the same time points and conditions but without incubation with **1** (Figure 2.14c). This negative control confirmed the stability of foci over time as the intensity of the GFP-MBNL1 foci was steady (Figure 2.14d). These results provide direct evidence of the ability of **1** to enter the cell and nucleus and disperse most of the MBNL1·CUG aggregates over a 2-4 h period.

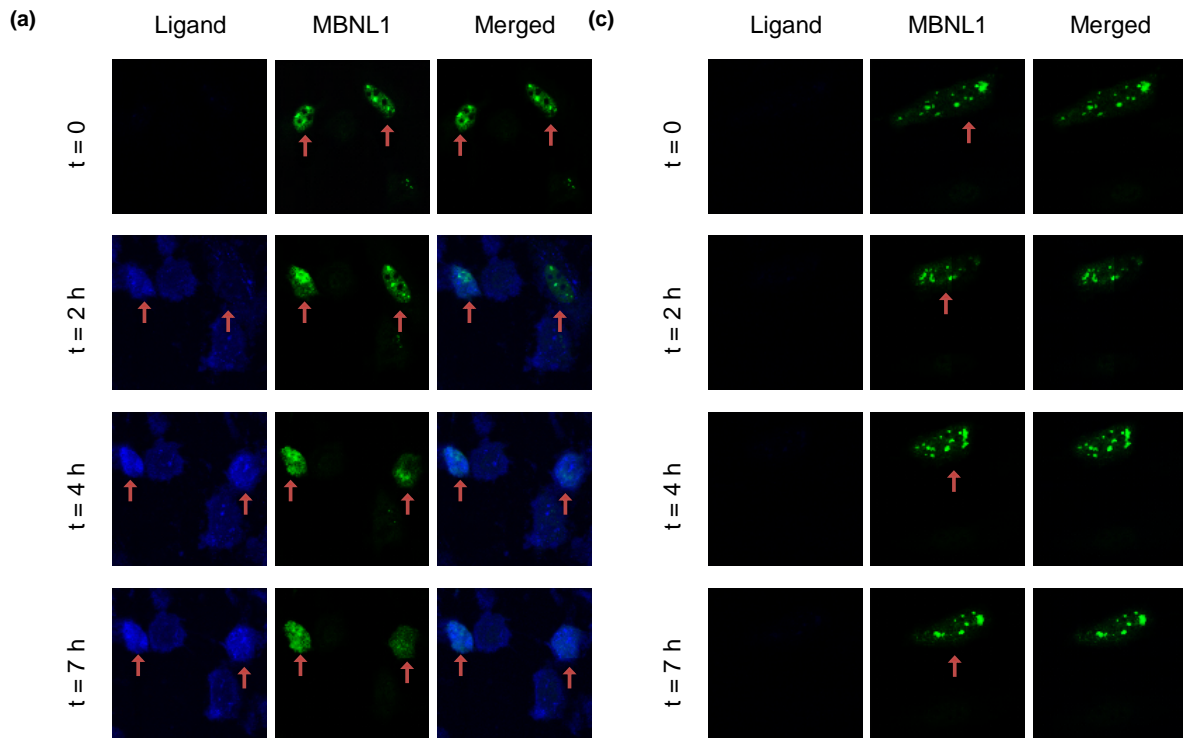
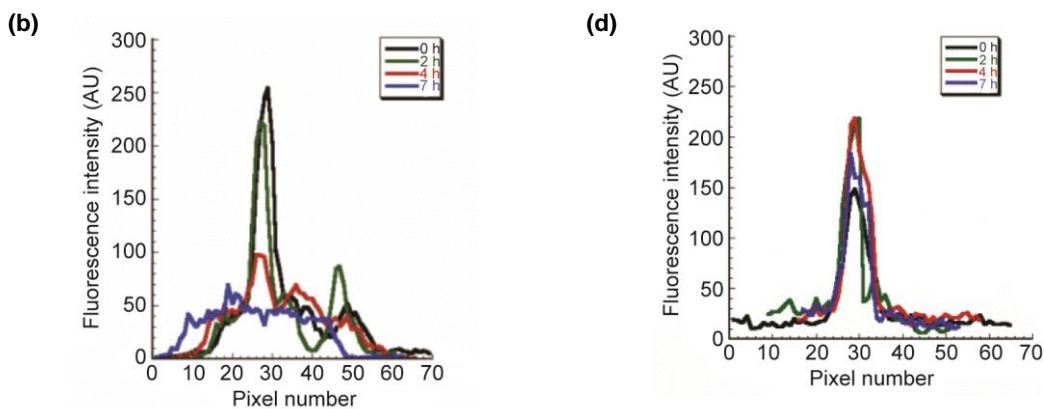


Fig 2.14 (cont. on next page)



**Figure 2.14.** Live cell microscopy demonstrates a direct evidence for MBNL1 foci dispersion with **1**. a) Live DM1 model cells are treated with **1** (75  $\mu\text{M}$ ) at  $t = 0$ , immediately after the first image is taken. Fluorescence of **1**, confirms its penetration to the nucleus. MBNL1 nuclear foci are gradually dispersing over time in two cells. Each box shows a  $100 \mu\text{m} \times 100 \mu\text{m}$  area. b) Plot of fluorescence intensity of a representative GFP-MBNL1 focus, corresponding to (a), shows dispersion over time. c) A single live cell shows stability of foci in a DM1 cell, in the absence of **1**, over the period of 7 h. Each box shows a  $100 \mu\text{m} \times 100 \mu\text{m}$  area. d) Plot of fluorescence intensity of a representative GFP-MBNL1 focus, corresponding to (c), shows no dispersion over time.

## 2.4 Conclusion

Ligand **1** was developed based on a previously reported *in vitro* active ligand.<sup>27</sup> The *in vitro* activity of **1** was assessed by optical melting and Surface Plasmon Resonance (SPR) techniques. Ligand **1** selectively binds to  $(\text{CUG})_{12}$ , stabilizes its hairpin structure and inhibits  $(\text{CUG})_{12} \cdot \text{MBNL1}$  interaction. The polyamine side-chain, provides full aqueous solubility that was absent in the initially reported ligand **2**. Ligand **1** penetrates the cellular as well as nuclear membrane. The bioactivity of **1** in model DM1 cells was studied by FISH technique using confocal microscopy and it was found to significantly reduce the number of ribonuclear foci. In splicing experiments, **1** partially rescued the IR mis-splicing. Moreover, we studied live DM1 model cells using time-lapse confocal microscopy. For the first time, we were able to observe uptake of a small molecular inhibitor of the  $(\text{CUG})_{12} \cdot \text{MBNL1}$  interaction by single live cells and further see its ability to disperse the foci over time. This approach is a powerful way to assess directly the effectiveness of small molecules targeted to CUG repeats. The positive results with compound **1** suggest that it is a good candidate for further development and therefore, toxicity and related studies are underway.



path length of the cuvettes used was 1 cm. The absorbance of 3.3  $\mu\text{M}$  (CUG)<sub>12</sub> in 1X PBS buffer in the absence and presence of 3.3 and 9.9  $\mu\text{M}$  of **1** was recorded at 260 nm with a slit width of 1 nm from 10 °C to 95 °C at a ramp rate of 0.5 °C min<sup>-1</sup>. Each profile for melting temperature analysis was generated by subtracting the absorbance of the solution of **1** in 1X PBS buffer from the (CUG)<sub>12</sub>/**1** solution. Melting temperatures were determined by fitting the melting curve using Meltwin 3.5 software.

**Surface Plasmon Resonance (SPR) Analysis.** All SPR experiments were conducted on a streptavidin coated sensor chip using a Biacore 3000 instrument. Streptavidin coated research grade sensor chips were preconditioned with three consecutive 1-min injections of 1 M NaCl/ 50 mM NaOH before the immobilization was started. 3'-biotin labeled (CUG)<sub>12</sub> was captured on flow cell 2 (Response Unit, RU, between 100-1100). Flow cell 1 was used as a reference. Inhibition analysis was carried out in PBS 1X buffer, pH = 7.4, containing 0.05% Tween-20 and 0.2 mg mL<sup>-1</sup> (7.4  $\mu\text{M}$  or 580  $\mu\text{M}$  nucleotides) bulk yeast t-RNA to confirm the specificity of inhibition. Various concentrations of **1** were passed over the immobilized RNA at a rate of 20  $\mu\text{L min}^{-1}$  for 300 s. After the initial 150 s, a solution of GST-MBNL1 protein, 650 nM, in the same buffer was flowed over the surface for 150 s. The reference-subtracted sensograms were obtained by subtracting the measured RU upon injection of PBS buffer from the sensograms. After the dissociation phase, the surface was regenerated, with a pulse of 0.5% SDS and/or 100 mM NaOH, for a few times followed by a buffer wash to reestablish baseline. For inhibition studies, the resulting sensograms were set to the baseline at t = 150 s to offset the binding of **1** to the immobilized (CUG)<sub>12</sub> surface. The peak RU at t = 150 s was recorded and converted to the percentage of (CUG)<sub>12</sub> bound by MBNL1. All values normalized to that measured in the absence of **1**. The data points were fit to a four parameter logistic curve to determine the apparent IC<sub>50</sub> using the following equation by Kaleidagraph software:

$$Y = \frac{Y_{max} - Y_{min}}{1 + \left(\frac{[1]}{IC_{50}}\right)^n} + Y_{min}$$

where Y is the percentage of (CUG)<sub>12</sub> bound by MBNL1,  $Y_{max}$  and  $Y_{min}$  are the maximum and minimum of this percentage and n is the Hill coefficient. Two or three separate SPR experiments on different sensor chips with different levels of (CUG)<sub>12</sub> immobilization were carried out to verify that the values are not affected by surface RNA density.

**FISH (Fluorescence *In Situ* Hybridization).** A total of ca. 120,000 HeLa cells were seeded in each well of a 6-well plate on coverslips. After a day, the cells were transfected with 500 ng DMPK–CUG<sub>0</sub> or DMPK–CUG<sub>960</sub> plasmid and 500 ng GFP–MBNL1 plasmid using Lipofectamine following the manufacturer’s protocol at cell confluence of 70–80%.<sup>42</sup> After 4 hours, the media was changed and **1** was added to each well at different concentrations (20, 50, 75 and 100 μM). After two days, the cells were fixed with 4% PFA then washed five times with 1X PBS. Fixed cells were permeabilized with 0.5% triton X-100 in 1X PBS at room temperature for 5 min. Cells were prewashed with 30% formamide in 2X SSC for 10 min at room temperature. Cells were probed with FISH probe (1 ng μL<sup>-1</sup> of Cy3 CAG<sub>10</sub> in 30% formamide, 2X SSC, 2 μg mL<sup>-1</sup> BSA, 66 μg mL<sup>-1</sup> yeast tRNA) for 2 h at 37 °C. Cells were then washed with 30% formamide in 2X SSC for 30 min at 37 °C, followed by washing with 1X SSC for 30 min at room temperature. The cells were washed twice with 1X PBS and then nuclei were stained with 1 μM To-Pro-3 and washed twice. Cells were mounted onto glass sides with ProLong® Gold. Slides were imaged at RT by LSM 710, AxioObserver confocal microscopy equipment using a confocal single photon technique with a plan-Apochromat 20x/0.8 M27 objective. Image analysis was performed by Axiovision interactive measurement. The following table indicates the excitation filters used in these experiments.

| Fluorophore | Component          | Excitation wavelength (nm) |
|-------------|--------------------|----------------------------|
| Acridine    | Ligand <b>1</b>    | 405                        |
| GFP         | MBNL1              | 488                        |
| Cy3         | CUG <sub>960</sub> | 555                        |
| TO-PRO-3    | Nucleus            | 639                        |

**IR Splicing Assay.** A total of ca. 120,000 cells were seeded in each well of a 6-well plate in DMEM supplemented with 10% FBS, 4.5 g L<sup>-1</sup> glucose, L-glutamine and no antibiotics. After a day, at about 70–80% confluence, the cells were transfected with 500 ng DMPK–CUG<sub>0</sub> or DMPK–CUG<sub>960</sub> plasmid and 500 ng IR mini-gene plasmid using Lipofectamine 2000 and Opti-MEM reduced serum medium following the standard protocol.<sup>42, 49</sup> The cells were incubated at 37 °C. After 4 h, the transfection medium was replaced by the DMEM medium and the cells were treated with **1** at three different concentrations (50, 75, and 150 μM).



After 48 h, cells were harvested and total RNA were isolated immediately using total RNA isolation kit and either stored at -80 °C or 1 µg of the RNA processed with reverse transcription step using iScript™ cDNA synthesis kit, the reverse transcription product was cleaned using Quick Spin kit. Approximately 70 ng of cDNA was used in PCR, 31-35 cycles (within a linear range) using PCR Master Mix kit following the standard protocol. 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. PCR products were separated on 8% PAGE gel in 1 X TBE, stained with EtBr in 15 min, de-stained with water in 15 min and observed under Molecular Imager.

The gel image was analyzed using ImageJ and the data were plotted using Kaleidagraph. The *p*-values were calculated using two-tailed Student *t* test.

**Live Cell Imaging.** A total of ca. 120,000 HeLa cells were grown in an Ibidi 35 mm Petri dish with a standard bottom, high walls and an imprinted 500 µm relocation grid. After a day, cell confluence reached to about 70–80%, cells were transfected with 500 ng DMPK–CUG<sub>960</sub> plasmid and 500 ng GFP-MBNL1 plasmid using Lipofectamine following standard protocol. After 4 h, media were changed and cells were incubated at 37 °C, 5% CO<sub>2</sub>. 24 h post-transfection, ligand **1** was added to final concentration of 75 µM. Live-cell, time-lapse images were taken before addition of **1** as well as at 2, 4 and 7 h time points at RT by a LSM 710, AxioObserver confocal microscopy equipment using a confocal single photon technique with a plan-Apochromat 20x/0.8 M27 objective. Image analysis was performed by Axiovision interactive measurement. For tracking the cells, DIC images were acquired simultaneously with the reflected light images using a TPMT module after setting the Köhler illumination with a fully opened condenser aperture (0.55 NA).

## 2.6 Experimental Synthetic Procedures

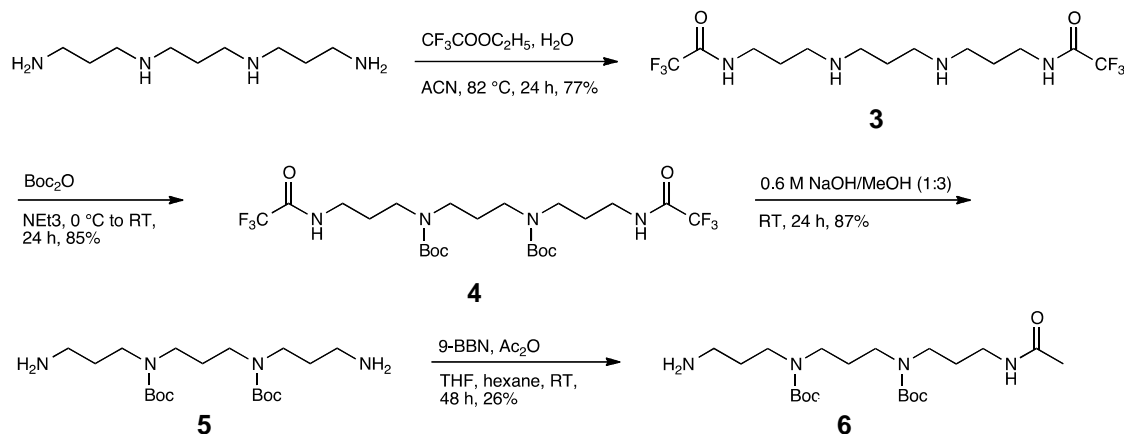


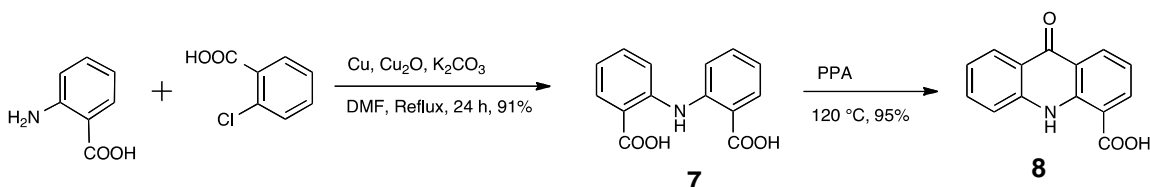
Figure 2.15. Synthesis of (6).

***N,N'*-((Propane-1,3-diylbis(azanediyl))bis(propane-3,1-diyl))bis(2,2,2-trifluoroacetamide) (3).** Title compound was synthesized from *N,N'*-bis(3-aminopropyl)propane-1,3-diamine as described previously,<sup>33</sup> in 77% yield, with minor changes.  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO-}d_6$ )  $\delta$  9.61 (s, 2H), 8.75 (bs, 2H), 3.29 – 3.23 (m, 4H), 2.98 (t,  $J = 7.5\text{ Hz}$ , 4H), 2.92 (t,  $J = 7.5\text{ Hz}$ , 4H), 1.91 (p,  $J = 7.8\text{ Hz}$ , 2H), 1.82 (p,  $J = 7.0\text{ Hz}$ , 4H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{DMSO-}d_6$ )  $\delta$  156.40, 117.08, 44.63, 44.02, 36.61, 25.22, 22.56;  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 381.2; found 381.2.

**Di-*tert*-Butyl propane-1,3-diylbis((3-(2,2,2-trifluoroacetamido)propyl)carbamate) (4).** Title compound was synthesized as described previously,<sup>33, 34</sup> with minor changes. The product was purified via flash chromatography ( $\text{SiO}_2$ ;  $\text{CH}_2\text{Cl}_2$ : $\text{MeOH}$ , 98:2 to 95:5) in 85% yield.  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO-}d_6$ )  $\delta$  9.42 (bs, 2H), 3.18 – 3.07 (m, 12H), 1.71 – 1.60 (m, 6H), 1.37 (s, 18H);  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 581.3; found 581.3.

**Di-*tert*-Butyl propane-1,3-diylbis((3-aminopropyl)carbamate) (5).** Title compound was synthesized as described previously,<sup>33</sup> with minor changes. The product was purified via flash chromatography ( $\text{SiO}_2$ ;  $\text{CH}_2\text{Cl}_2$ : $\text{MeOH}$ : $\text{NH}_4\text{OH}$ , 80:19:1 to 67:30:2) to yield the pure product in 87% yield.  $^1\text{H}$  NMR (500 MHz,  $\text{Chloroform-}d$ )  $\delta$  3.28 (bs, 4H), 3.16 (bs, 4H), 2.69 (t,  $J = 6.4\text{ Hz}$ , 4H), 1.79 – 1.71 (m, 2H), 1.64 (p,  $J = 6.8\text{ Hz}$ , 4H), 1.45 (s, 18H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{Chloroform-}d$ )  $\delta$  155.29, 79.06, 44.44, 38.96, 32.34, 31.73, 28.17;  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 389.3; found 389.3.

**tert-Butyl 3-((3-acetamidopropyl)(tert-butoxycarbonyl)amino)propyl(3-amino propyl) carbamate (6).** Compound **5** (1.1 g, 2.83 mmol, 1 equiv.) was dissolved in THF (15 mL). A 0.5 M solution of 9-Borabicyclo(3.3.1)nonane (5.66 mL, 2.83 mmol, 1 equiv.) was added and the reaction was stirred at room temperature for 3 hours.<sup>35</sup> Acetic anhydride (275 mg, 2.68 mmol, 0.95 equiv.) was added and the reaction was allowed to proceed at room temperature overnight. The reaction was quenched with water (10 mL) and extracted with EtOAc (3 x 50 mL). The combined organic layers were dried over magnesium sulfate and filtered through celite. The solution was concentrated by rotary evaporation, purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH:NH<sub>4</sub>OH, 94:5:0.5 to 89:10:1) to afford **6** (1.154 g, 26%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.32 – 3.10 (m, 10H), 2.71 (t, J = 6.7 Hz, 2H), 1.98 (s, 3H), 1.82 (s, 2H), 1.78 – 1.72 (m, 2H), 1.69 – 1.64 (m, 2H), 1.44 (s, 18H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 431.3; found 431.3.



**Figure 2.16.** Synthesis of 9-oxo-9,10-dihydroacridine-4-carboxylic acid (**8**).

**2,2'-Azanediyldibenzoic acid (7).** Title compound was synthesized as described previously,<sup>36</sup> in 91% yield, with minor changes in the work-up procedure. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 7.91 (dd, J = 7.9, 1.3 Hz, 2H), 7.50 – 7.40 (m, 4H), 6.95 (ddd, J = 8.0, 6.8, 1.5 Hz, 2H); <sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>) δ 168.41, 143.58, 133.38, 131.81, 119.99, 117.56, 113.56; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 258.1; found 258.1.

**9-Oxo-9,10-dihydroacridine-4-carboxylic acid (8).** Title compound was synthesized as described previously,<sup>36</sup> in 95% yield, with minor changes in the work-up procedure. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 12.01 (bs, 1H), 8.53 (dd, J = 8.0, 1.5 Hz, 1H), 8.45 (dd, J = 7.5, 1.6 Hz, 1H), 8.24 (d, J = 8.0 Hz, 1H), 7.83 – 7.73 (m, 2H), 7.41 – 7.29 (m, 2H); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>) δ 176.53, 169.14, 141.20, 139.92, 136.90, 134.11, 132.41, 125.89, 122.32, 121.63, 120.60, 120.24, 118.63, 115.01 *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 240.1; found 240.1.

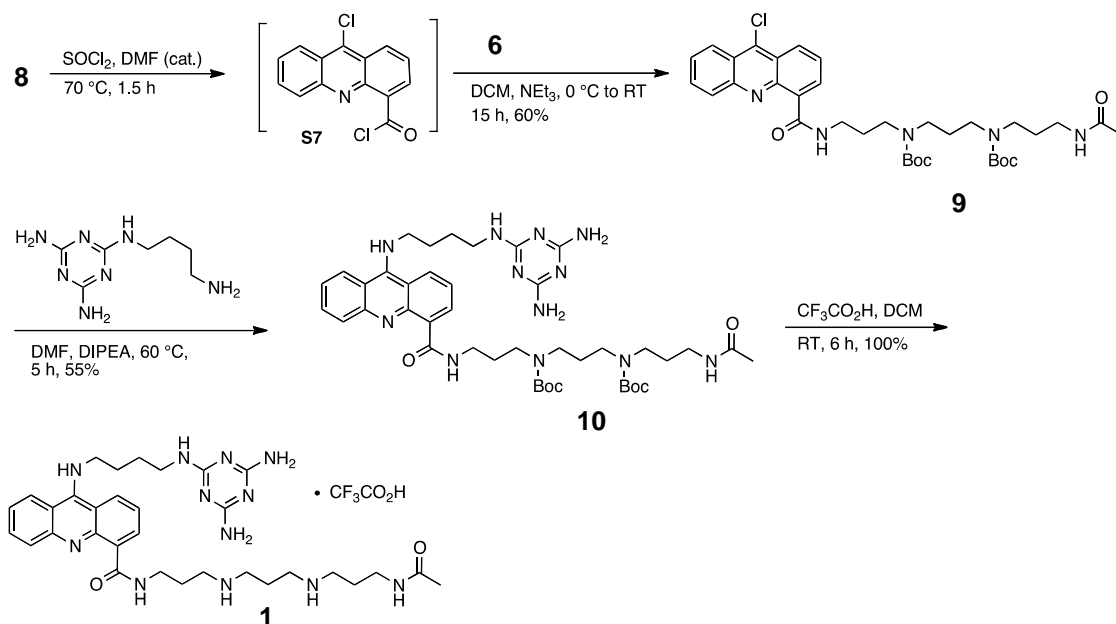


Figure 2.17. Synthesis of (1).

***tert*-Butyl (3-((3-acetamidopropyl)(*tert*-butoxycarbonyl)amino)propyl)(3-(9-Chloroacridine-4-carboxamido)propyl)carbamate (9).** A round-bottom flask, equipped with a stir bar, was charged with **8** (600 mg, 2.5 mmol, 1 equiv.) and freshly distilled thionyl chloride (3 mL, 41 mmol, 16.4 equiv.). A catalytic amount of DMF was added and heated gently under reflux at  $69\text{ }^\circ\text{C}$ , stirring until homogeneous and then for 1 h. The excess thionyl chloride was distilled off and the last traces of it were removed azeotropically via co-evaporation with DCM (3 x 50 mL). It was left under vacuum (minimally) for 1 h to afford the crude intermediate as a yellow powder. The crude intermediate was dissolved in anhydrous DCM. Anhydrous triethylamine was added to the solution until the pH was 11 and it was cooled to  $0\text{ }^\circ\text{C}$ . Compound **6** (1.065 g, 2.75 mmol, 1.1 equiv.) was added and the solution was stirred at  $0\text{ }^\circ\text{C}$  for 2 hours and slowly warmed to room temperature overnight. The solvent was removed by rotary evaporation and the crude mixture was purified via flash chromatography ( $\text{SiO}_2$ ;  $\text{CH}_2\text{Cl}_2$ :MeOH, 98:2 to 95:5) to yield **9** as a yellow solid (586 mg, 60%); *m/z* LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 670.3; found 670.3.

***tert*-Butyl(3-((3-acetamidopropyl)(*tert*-butoxycarbonyl)amino)propyl)(3-(9-((4-((4,6-di-amino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)propyl)carbamate (10).** A round-bottom flask, equipped with a stir bar, was charged with **9** (500 mg, 0.576 mmol, 1 equiv.) and of *N*-(4-Aminobutyl)-[1,3,5]triazine-2,4,6-triamine (125

mg, 0.635 mmol, 1.1 equiv.). DIPEA (163 mg, 1.26 mmol, 2.2 equiv.) and anhydrous DMF (25 mL) were added. The solution was heated at 70 °C for 5 hours. The solvent was removed by rotary evaporation and the product was purified via flash chromatography (Basic Alumina; DCM:Methanol:NH<sub>4</sub>OH, from 95:4.9:0.1 to 90:9.5:0.5) to yield **10** as a yellow solid (377 mg, 0.317 mmol, 55%); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 831.5; found 831.5.

***N*-(3-((3-((3-Acetamidopropyl)amino)propyl)amino)propyl)-9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide (1)**. A round-bottom flask, equipped with a stir bar, was charged with **10** (310mg, 0.261 mmol, 1equiv.). TFA (30 mL) and anhydrous DCM (70 mL) were added and stirred at room temperature for 6 h. The solvents were removed to yield **1** as a yellow solid (437 mg, 0.261 mmol, 100%). <sup>1</sup>H NMR (500 MHz, Deuterium Oxide) δ 8.33 (d, J = 8.6 Hz, 1H), 8.16 (dd, J = 13.0, 8.0 Hz, 2H), 7.90 – 7.84 (m, 2H), 7.58 (d, J = 8.5 Hz, 1H), 7.46 (t, J = 7.9 Hz, 2H), 4.02 (t, J = 6.7 Hz, 2H), 3.56 (t, J = 7.0 Hz, 2H), 3.29 – 3.10 (m, 10H), 3.06 – 3.03 (m, 2H), 2.14 – 2.06 (m, 4H), 1.96 (s, 3H), 1.93 – 1.85 (m, 4H), 1.70 – 1.63 (m, 2H); <sup>13</sup>C NMR (125 MHz, Deuterium Oxide) δ 174.89, 168.78, 165.12, 163.37, 163.09, 162.80, 162.52, 159.99, 159.45, 158.02, 156.64, 136.13, 134.88, 120.06, 119.18, 118.71, 117.74, 115.41, 113.09, 48.84, 45.74, 45.38, 44.74, 40.38, 37.07, 36.16, 31.59, 25.78, 24.87, 22.89, 21.91; *m/z* HRMS (ESI) calculated for [M+H]<sup>+</sup>: 631.3945; found 631.3944.

## 2.7 Acknowledgements

I would like to thank Dr. Yuan Fu and Professor Anne M. Baranger, for the MBNL1 protein expression and purification as well as the optical melting studies, Lien Nguyen, for cell culture and providing samples for confocal studies as well as splicing and cytotoxicity assay, Kali A. Miller, for her assistance with synthesis of **1**, and Dr. John Eichorst and Dr. Sivaguru Mayandi (Institute of Genomic Biology, University of Illinois) for their technical assistance with the confocal microscopy.

## 2.8 References

- (1) Wilson, W. D., and Li, K. (2000) Targeting RNA with small molecules, *Curr. Med. Chem.* 7, 73-98.
- (2) Pearson, N. D., and Prescott, C. D. (1997) RNA as a drug target, *Chem. Biol.* 4, 409-414.
- (3) Tor, Y. (2003) Targeting RNA with small molecules, *ChemBioChem* 4, 998-1007.
- (4) Guan, L., and Disney, M. D. (2012) Recent advances in developing small molecules targeting RNA, *ACS Chem. Biol.* 7, 73-86.
- (5) Thomas, J. R., and Hergenrother, P. J. (2008) Targeting RNA with small molecules, *Chem. Rev.* 108, 1171-1224.
- (6) Gallego, J., and Varani, G. (2001) Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges, *Acc. Chem. Res.* 34, 836-843.
- (7) Fulle, S., and Gohlke, H. (2010) Molecular recognition of RNA: challenges for modelling interactions and plasticity, *Journal of molecular recognition : JMR* 23, 220-231.
- (8) Chow, C. S., and Bogdan, F. M. (1997) A Structural Basis for RNAMinus signLigand Interactions, *Chem. Rev.* 97, 1489-1514.
- (9) Carlson, C. B., Stephens, O. M., and Beal, P. A. (2003) Recognition of double-stranded RNA by proteins and small molecules, *Biopolymers* 70, 86-102.
- (10) Cooper, T. A., Wan, L., and Dreyfuss, G. (2009) RNA and disease, *Cell* 136, 777-793.
- (11) Cooper, T. A. (2006) A reversal of misfortune for myotonic dystrophy?, *New Engl. J. Med.* 355, 1825-1827.
- (12) Foff, E. P., and Mahadevan, M. S. (2011) Therapeutics development in myotonic dystrophy type 1, *Muscle Nerve* 44, 160-169.
- (13) Gomes-Pereira, M., Cooper, T. A., and Gourdon, G. (2011) Myotonic dystrophy mouse models: towards rational therapy development, *Trends Mol. Med.* 17, 506-517.
- (14) Mirkin, S. M. (2007) Expandable DNA repeats and human disease, *Nature* 447, 932-940.
- (15) O'Rourke, J. R., and Swanson, M. S. (2009) Mechanisms of RNA-mediated disease, *J. Biol. Chem.* 284, 7419-7423.

- (16) Kanadia, R. N., Shin, J., Yuan, Y., Beattie, S. G., Wheeler, T. M., Thornton, C. A., and Swanson, M. S. (2006) Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A.* *103*, 11748-11753.
- (17) Grammatikakis, I., Goo, Y. H., Echeverria, G. V., and Cooper, T. A. (2011) Identification of MBNL1 and MBNL3 domains required for splicing activation and repression, *Nucleic Acids Res.* *39*, 2769-2780.
- (18) Wheeler, T. M., Krym, M. C., and Thornton, C. A. (2007) Ribonuclear foci at the neuromuscular junction in myotonic dystrophy type 1, *Neuromuscular disorders : NMD* *17*, 242-247.
- (19) Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., and Thornton, C. A. (2009) Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA, *Science* *325*, 336-339.
- (20) Wheeler, T. M., Leger, A. J., Pandey, S. K., MacLeod, A. R., Nakamori, M., Cheng, S. H., Wentworth, B. M., Bennett, C. F., and Thornton, C. A. (2012) Targeting nuclear RNA for in vivo correction of myotonic dystrophy, *Nature* *488*, 111-115.
- (21) Garcia-Lopez, A., Llamusi, B., Orzaez, M., Perez-Paya, E., and Artero, R. D. (2011) In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models, *Proc. Natl. Acad. Sci. U. S. A.* *108*, 11866-11871.
- (22) Warf, M. B., Nakamori, M., Matthys, C. M., Thornton, C. A., and Berglund, J. A. (2009) Pentamidine reverses the splicing defects associated with myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A.* *106*, 18551-18556.
- (23) Ofori, L. O., Hoskins, J., Nakamori, M., Thornton, C. A., and Miller, B. L. (2012) From dynamic combinatorial 'hit' to lead: in vitro and in vivo activity of compounds targeting the pathogenic RNAs that cause myotonic dystrophy, *Nucleic Acids Res.* *40*, 6380-6390.
- (24) Parkesh, R., Childs-Disney, J. L., Nakamori, M., Kumar, A., Wang, E., Wang, T., Hoskins, J., Tran, T., Housman, D., Thornton, C. A., and Disney, M. D. (2012) Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching, *J. Am. Chem. Soc.* *134*, 4731-4742.
- (25) Pushechnikov, A., Lee, M. M., Childs-Disney, J. L., Sobczak, K., French, J. M., Thornton, C. A., and Disney, M. D. (2009) Rational design of ligands targeting

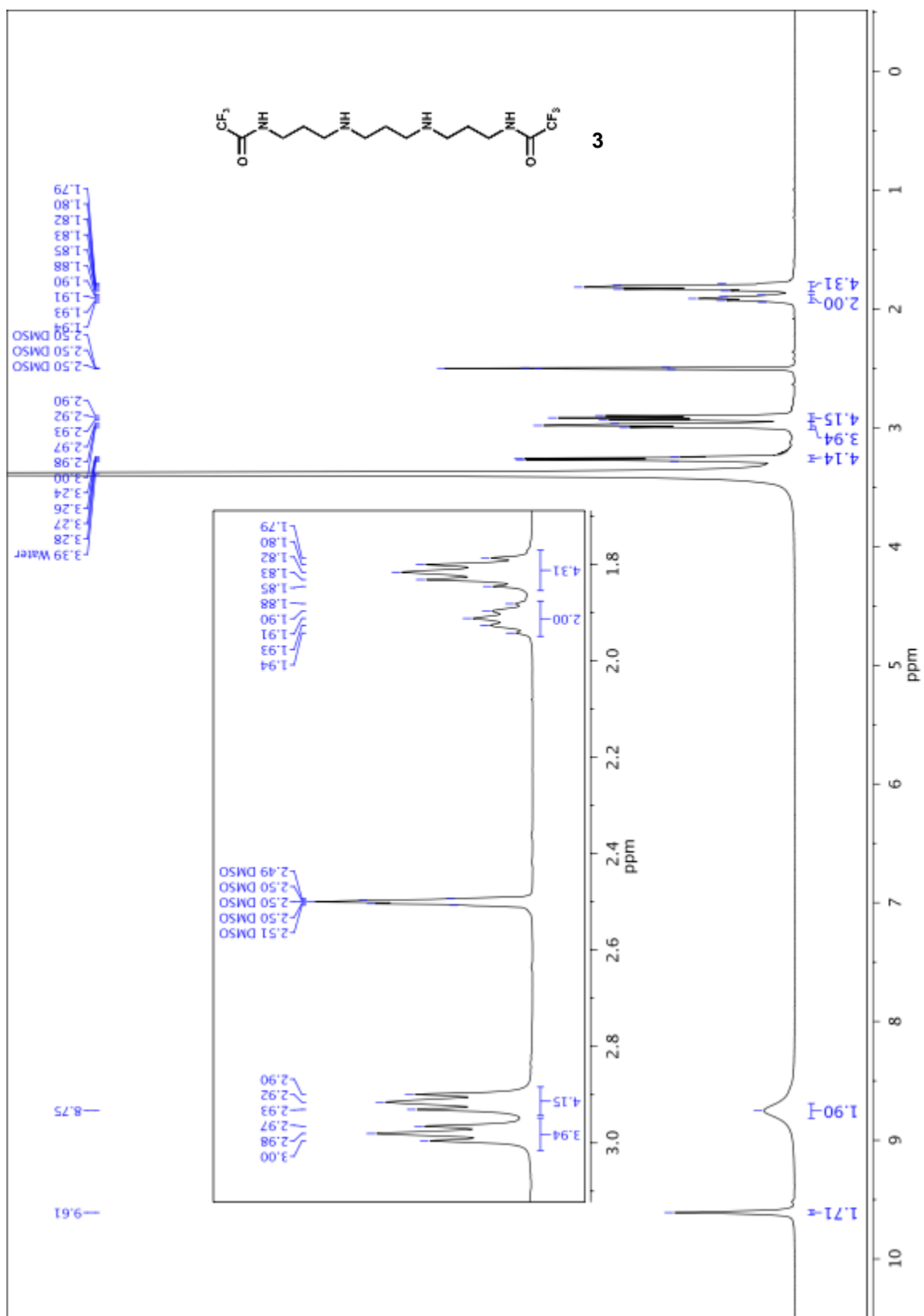
- triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3, *J. Am. Chem. Soc.* **131**, 9767-9779.
- (26) Childs-Disney, J. L., Hoskins, J., Rzuczek, S. G., Thornton, C. A., and Disney, M. D. (2012) Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potentially bioactive, *ACS Chem. Biol.* **7**, 856-862.
- (27) Arambula, J. F., Ramisetty, S. R., Baranger, A. M., and Zimmerman, S. C. (2009) A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16068-16073.
- (28) Fischer, W., Brissault, B., Prevost, S., Kopaczynska, M., Andreou, I., Janosch, A., Gradzielski, M., and Haag, R. (2010) Synthesis of linear polyamines with different amine spacings and their ability to form dsDNA/siRNA complexes suitable for transfection, *Macromol. Biosci.* **10**, 1073-1083.
- (29) Palmer, A. J., and Wallace, H. M. (2010) The polyamine transport system as a target for anticancer drug development, *Amino Acids* **38**, 415-422.
- (30) Delcros, J. G., Tomasi, S., Carrington, S., Martin, B., Renault, J., Blagbrough, I. S., and Uriac, P. (2002) Effect of spermine conjugation on the cytotoxicity and cellular transport of acridine, *J. Med. Chem.* **45**, 5098-5111.
- (31) Sanchez-Carrasco, S., Delcros, J. G., Moya-Garcia, A. A., Sanchez-Jimenez, F., and Ramirez, F. J. (2008) Study by optical spectroscopy and molecular dynamics of the interaction of acridine-spermine conjugate with DNA, *Biophys. Chem.* **133**, 54-65.
- (32) Perez-Flores, L., Ruiz-Chica, A. J., Delcros, J. G., Sanchez-Jimenez, F. M., and Ramirez, F. J. (2008) Effect of spermine conjugation on the interaction of acridine with alternating purine-pyrimidine oligodeoxyribonucleotides studied by CD, fluorescence and absorption spectroscopies, *Spectrochim Acta A Mol Biomol Spectrosc* **69**, 1089-1096.
- (33) Ilies, M. A., Seitz, W. a., Johnson, B. H., Ezell, E. L., Miller, A. L., Thompson, E. B., and Balaban, A. T. (2006) Lipophilic pyrylium salts in the synthesis of efficient pyridinium-based cationic lipids, gemini surfactants, and lipophilic oligomers for gene delivery., *J. Med. Chem.* **49**, 3872-3887.
- (34) Carta, F., Temperini, C., Innocenti, A., Scozzafava, A., Kaila, K., and Supuran, C. T. (2010) Polyamines inhibit carbonic anhydrases by anchoring to the zinc-coordinated water molecule, *J. Med. Chem.* **53**, 5511-5522.

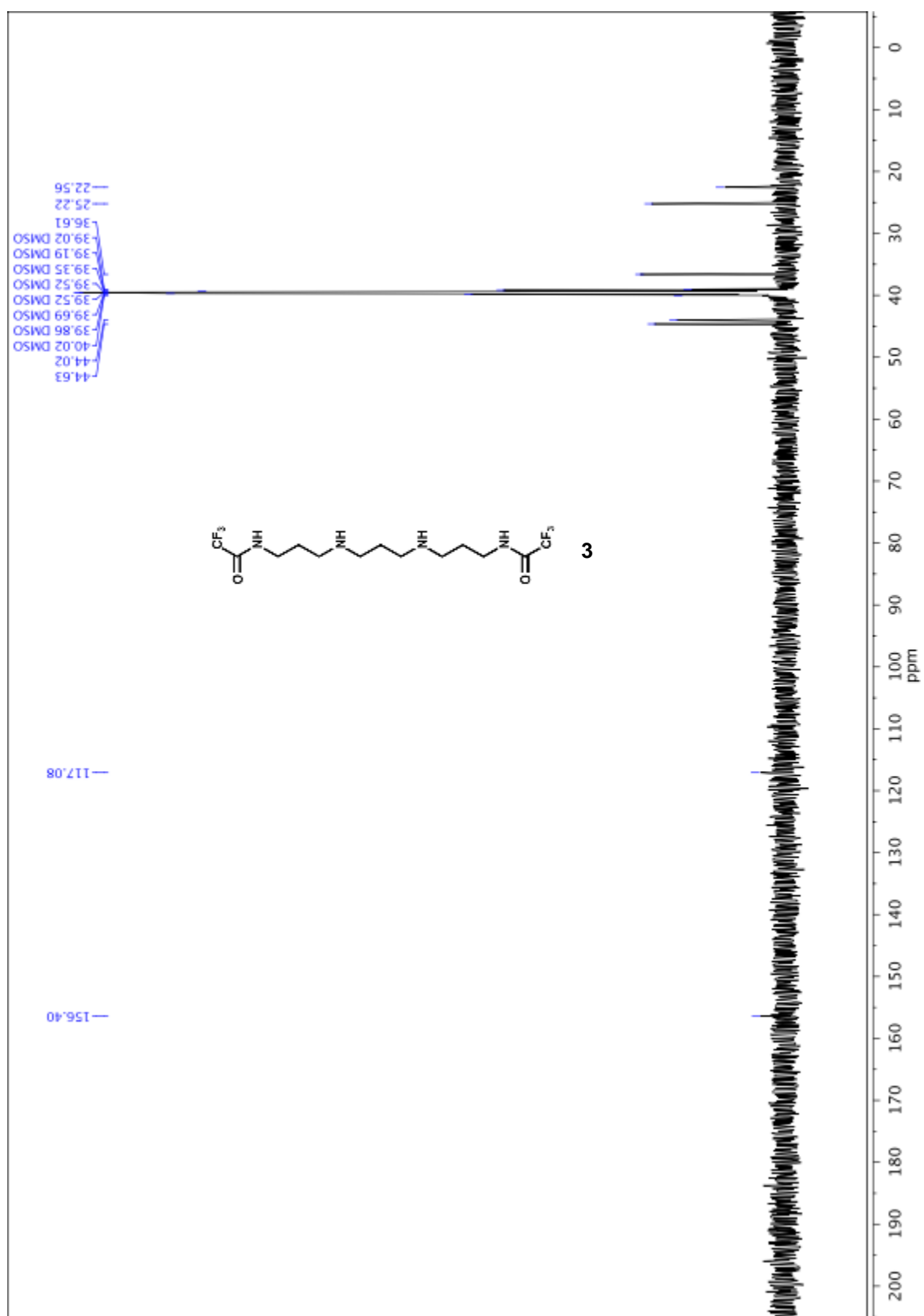


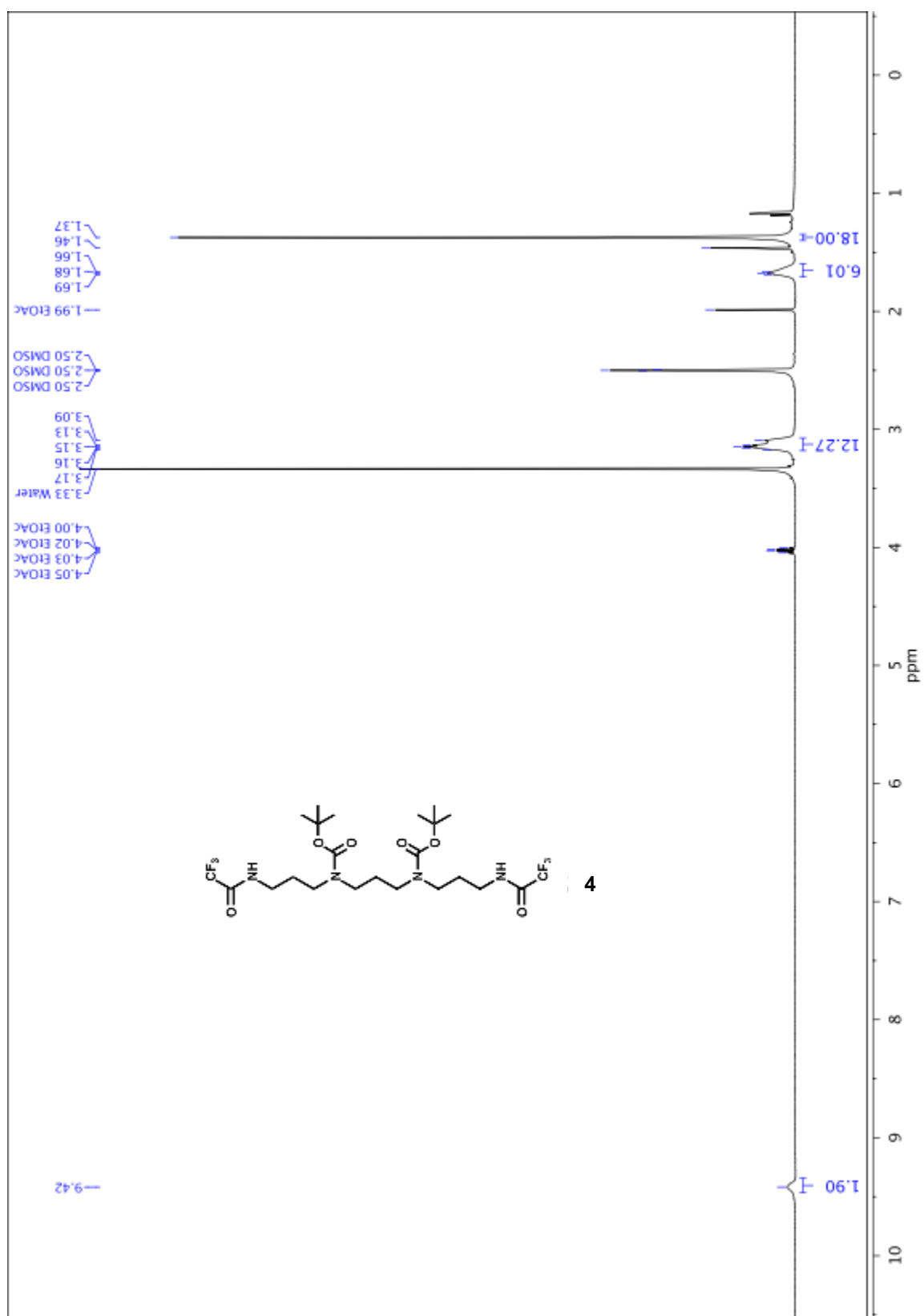
- (35) Zhang, Z., Yin, Z., Meanwell, N. a., Kadow, J. F., and Wang, T. (2003) Selective monoacylation of symmetrical diamines via prior complexation with boron., *Org. Lett.* **5**, 3399-3402.
- (36) Cuenca, F., Moore, M. J., Johnson, K., Guyen, B., De Cian, A., and Neidle, S. (2009) Design, synthesis and evaluation of 4,5-di-substituted acridone ligands with high G-quadruplex affinity and selectivity, together with low toxicity to normal cells, *Bioorg. Med. Chem. Lett.* **19**, 5109-5113.
- (37) Warf, M. B., and Berglund, J. A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T, *RNA* **13**, 2238-2251.
- (38) Fu, Y., Ramisetty, S. R., Hussain, N., and Baranger, A. M. (2012) MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation, *ChemBioChem* **13**, 112-119.
- (39) Laurent, F. X., Sureau, A., Klein, A. F., Trouslard, F., Gasnier, E., Furling, D., and Marie, J. (2012) New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats, *Nucleic Acids Res.* **40**, 3159-3171.
- (40) Mann, D. a., Kanai, M., Maly, D. J., and Kiessling, L. L. (1998) Probing Low Affinity and Multivalent Interactions with Surface Plasmon Resonance: Ligands for Concanavalin A, *J. Am. Chem. Soc.* **120**, 10575-10582.
- (41) Echeverria, G. V., and Cooper, T. A. (2012) RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity, *Brain Res.* **1462**, 100-111.
- (42) Ho, T. H., Savkur, R. S., Poulos, M. G., Mancini, M. A., Swanson, M. S., and Cooper, T. A. (2005) Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy, *J. Cell Sci.* **118**, 2923-2933.
- (43) Long, R. M., Elliott, D. J., Stutz, F., Rosbash, M., and Singer, R. H. (1995) Spatial consequences of defective processing of specific yeast mRNAs revealed by fluorescent in situ hybridization, *RNA* **1**, 1071-1078.
- (44) Sen, S., Talukdar, I., Liu, Y., Tam, J., Reddy, S., and Webster, N. J. (2010) Muscleblind-like 1 (Mbnl1) promotes insulin receptor exon 11 inclusion via binding to a downstream evolutionarily conserved intronic enhancer, *J. Biol. Chem.* **285**, 25426-25437.

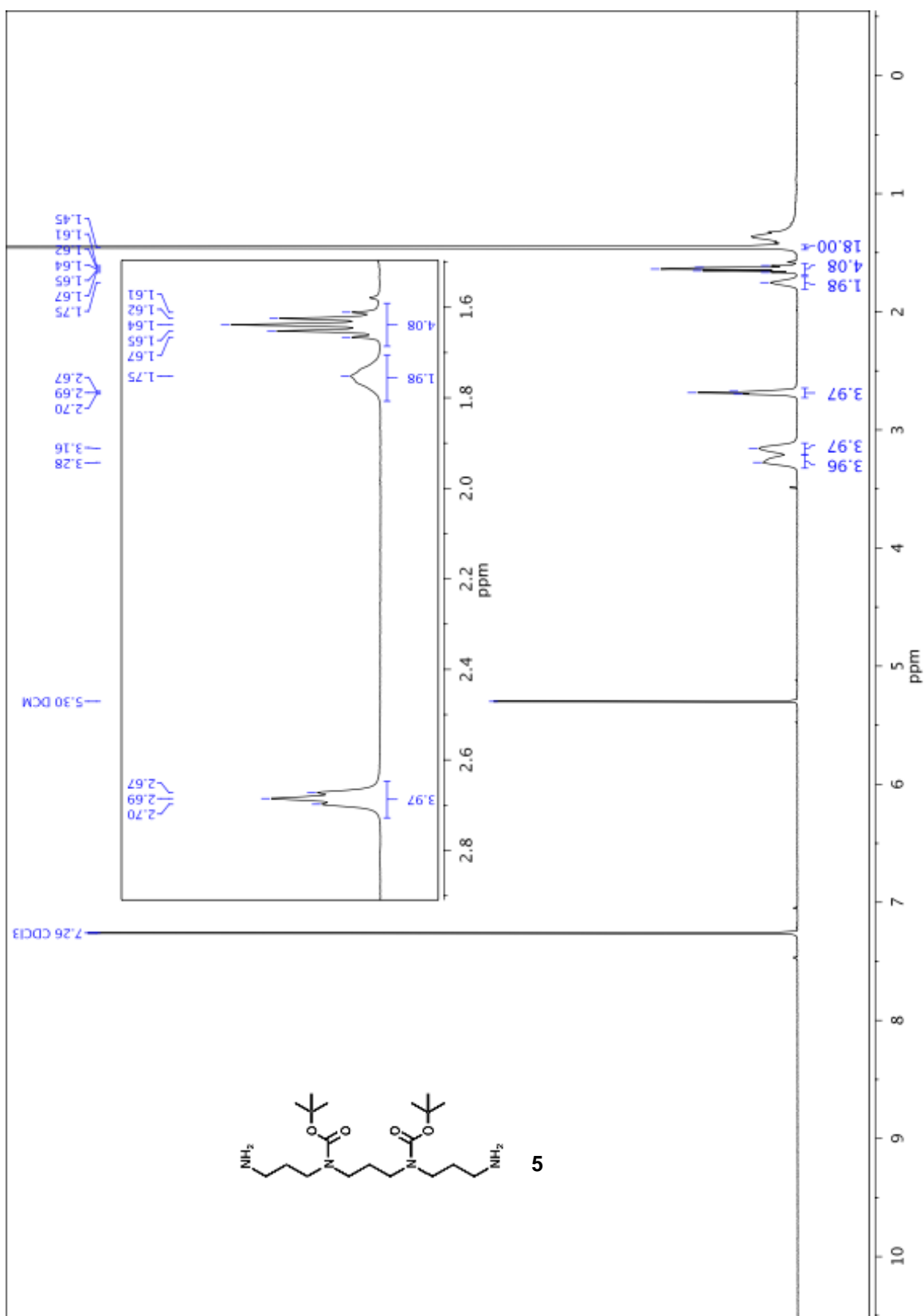
- (45) Savkur, R. S., Philips, A. V., and Cooper, T. A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy, *Nat. Genet.* 29, 40-47.
- (46) Kosaki, A., Nelson, J., and Webster, N. J. (1998) Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA, *J. Biol. Chem.* 273, 10331-10337.
- (47) Vichai, V., and Kirtikara, K. (2006) Sulforhodamine B colorimetric assay for cytotoxicity screening, *Nat. Protoc.* 1, 1112-1116.
- (48) Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., and Swanson, M. S. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs, *Nucleic Acids Res.* 35, 5474-5486.
- (49) Kosaki, a., Nelson, J., and Webster, N. J. (1998) Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA., *J. Biol. Chem.* 273, 10331-10337.

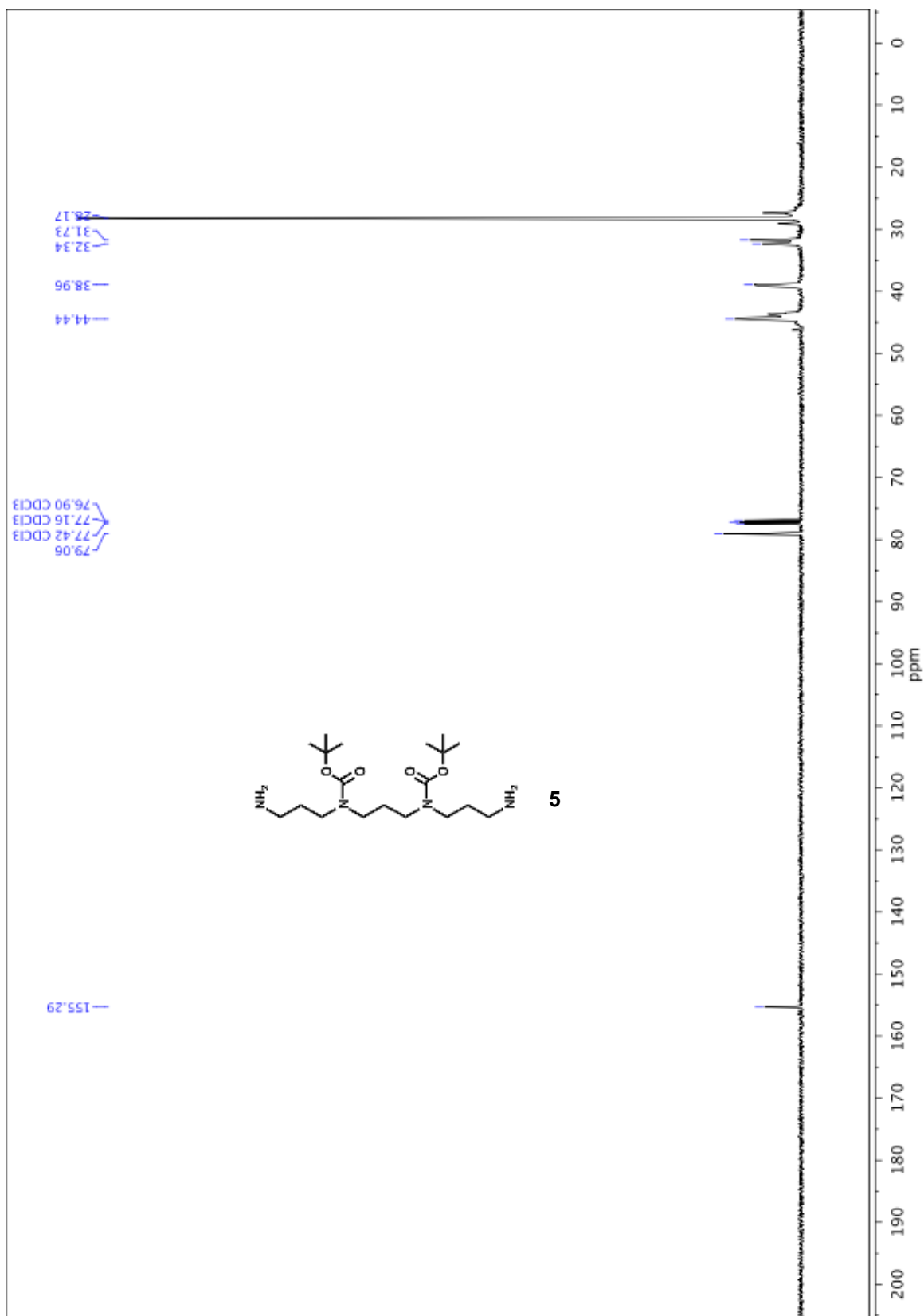
## 2.9 NMR Spectra

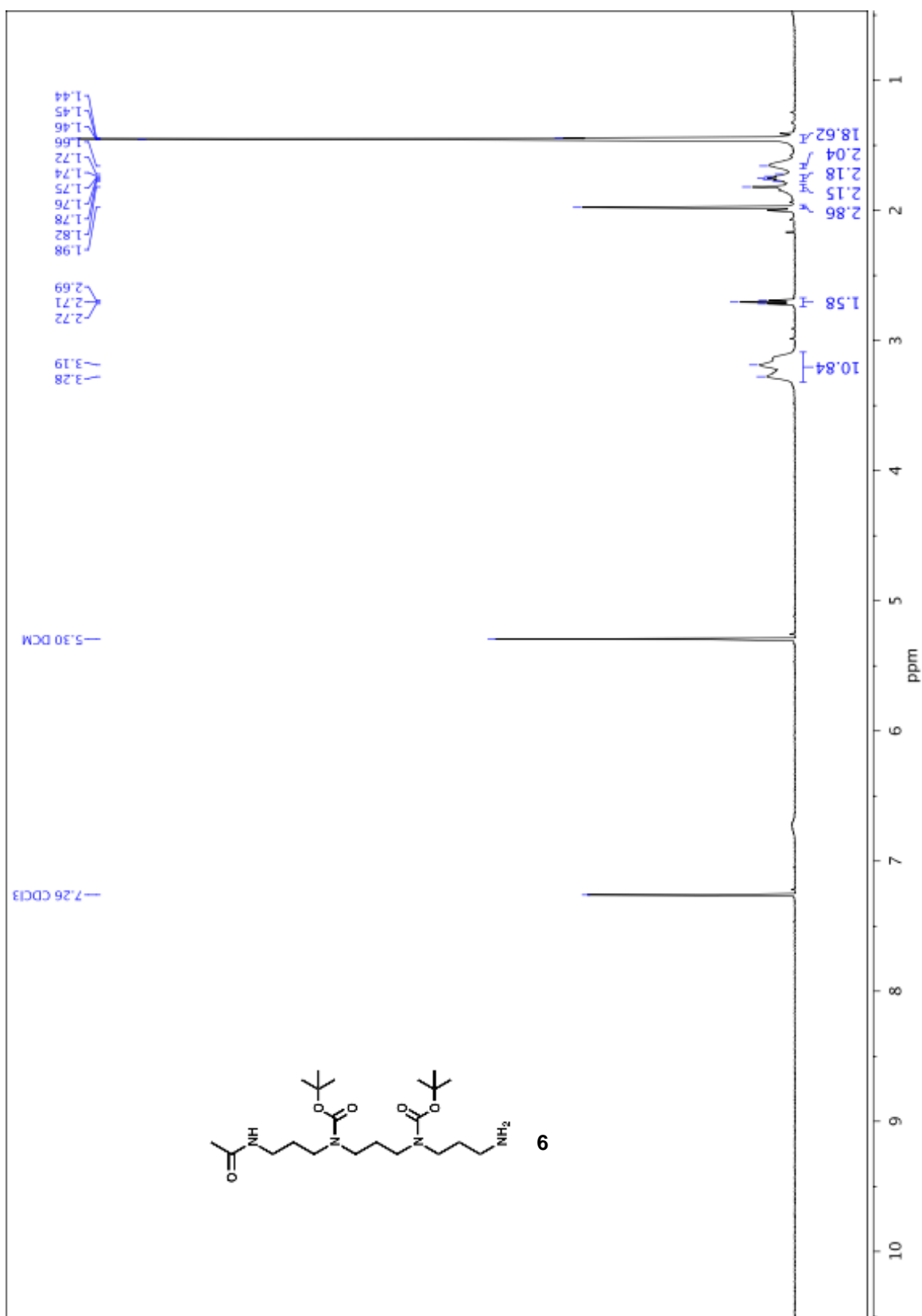




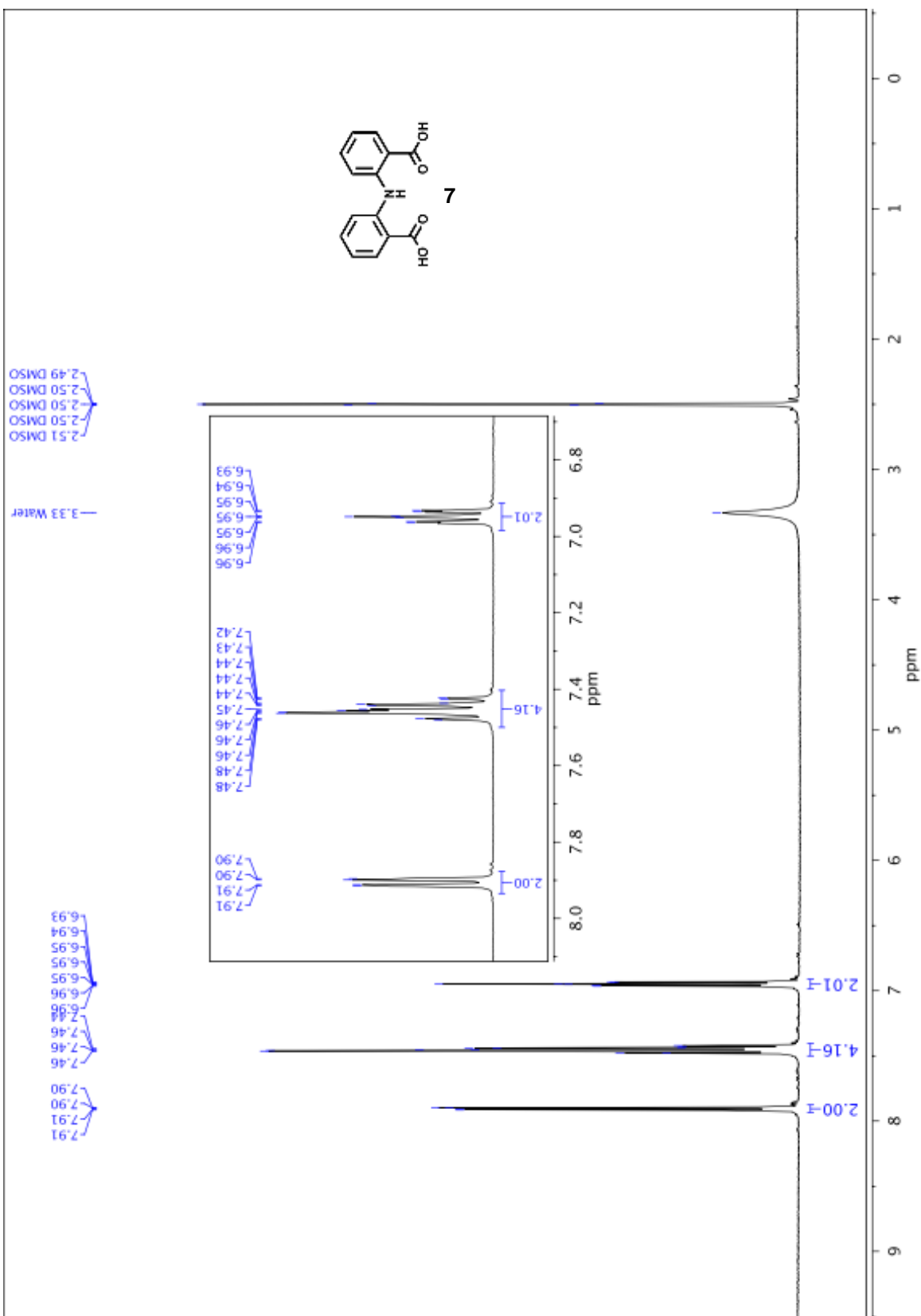


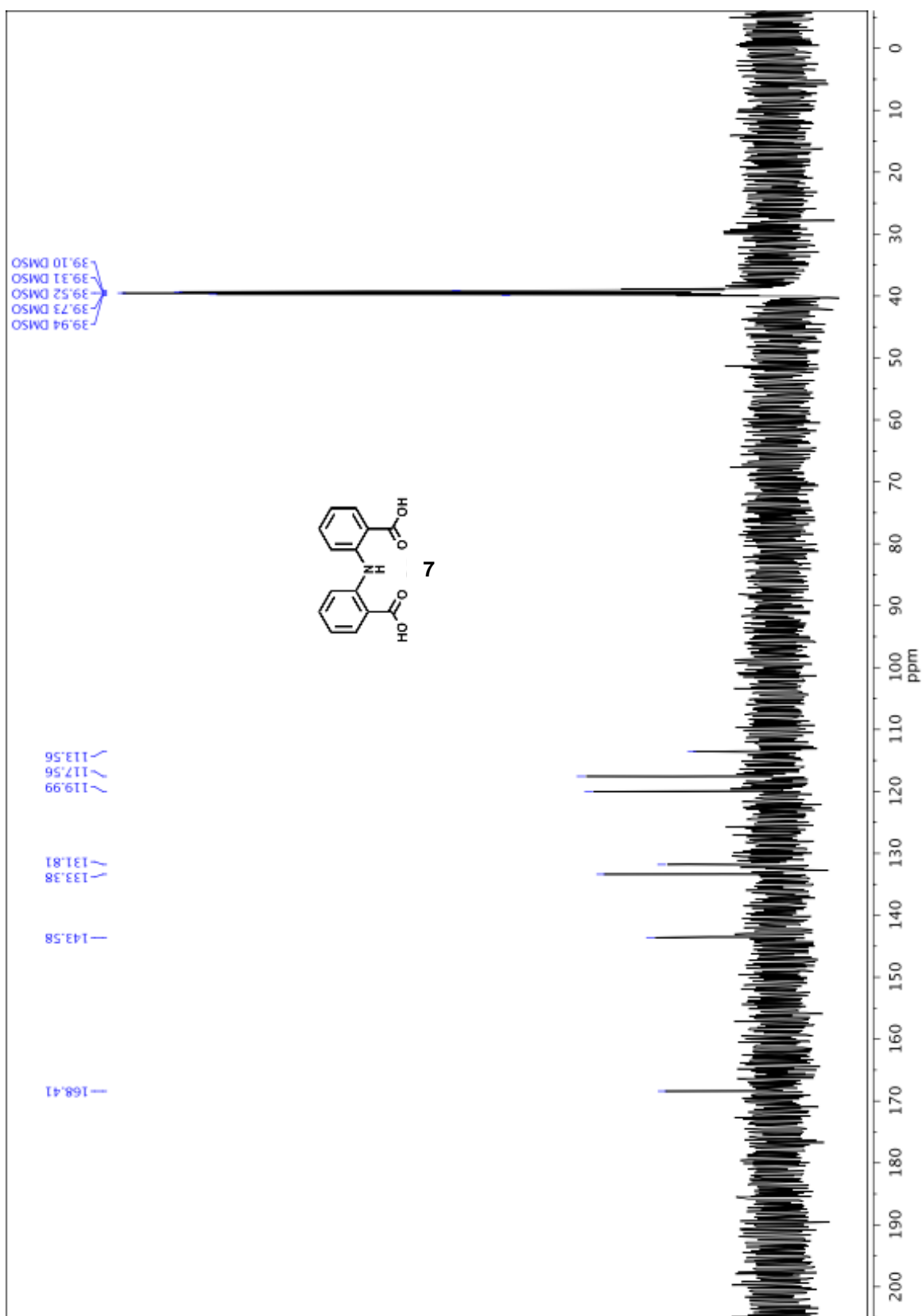


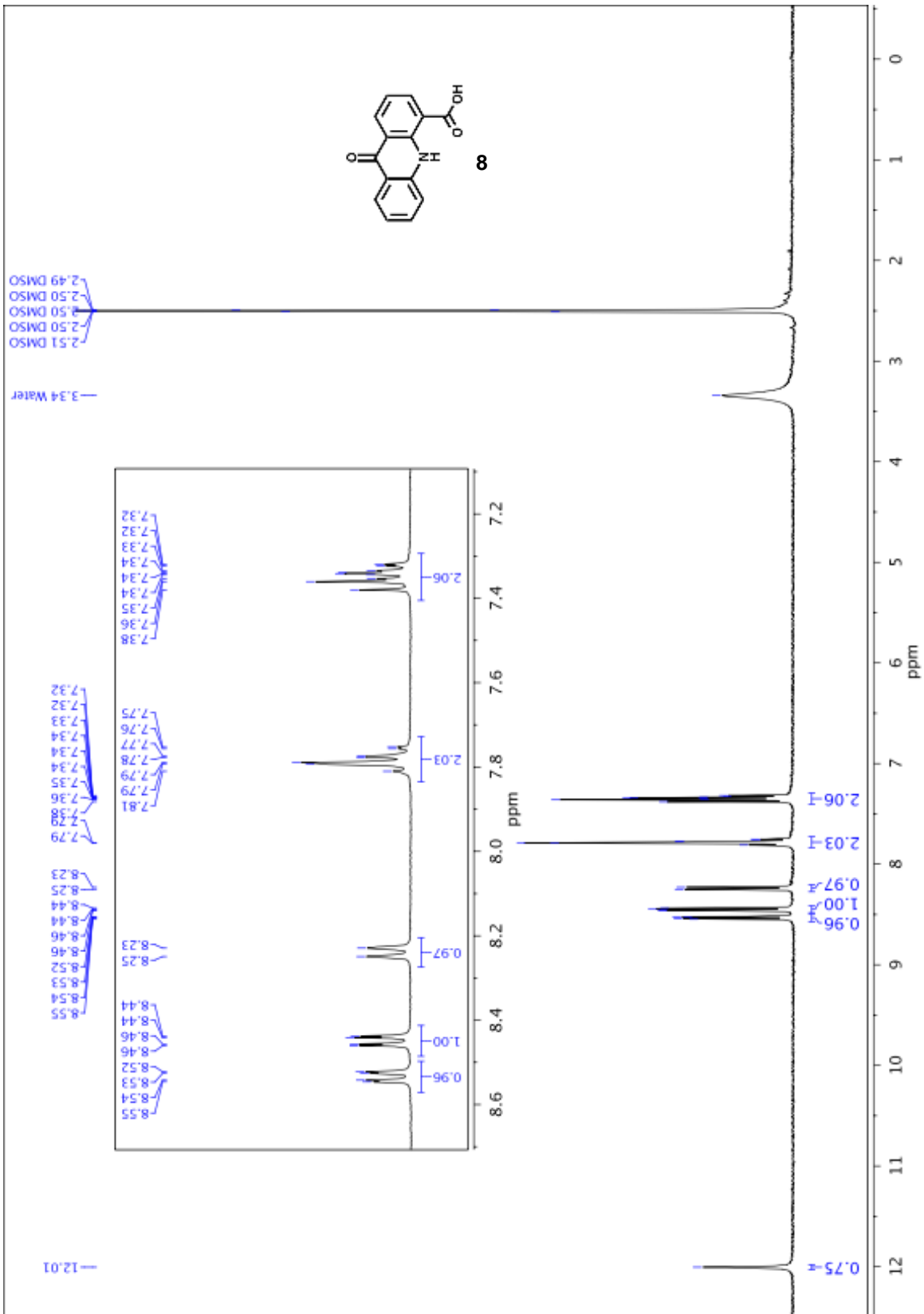


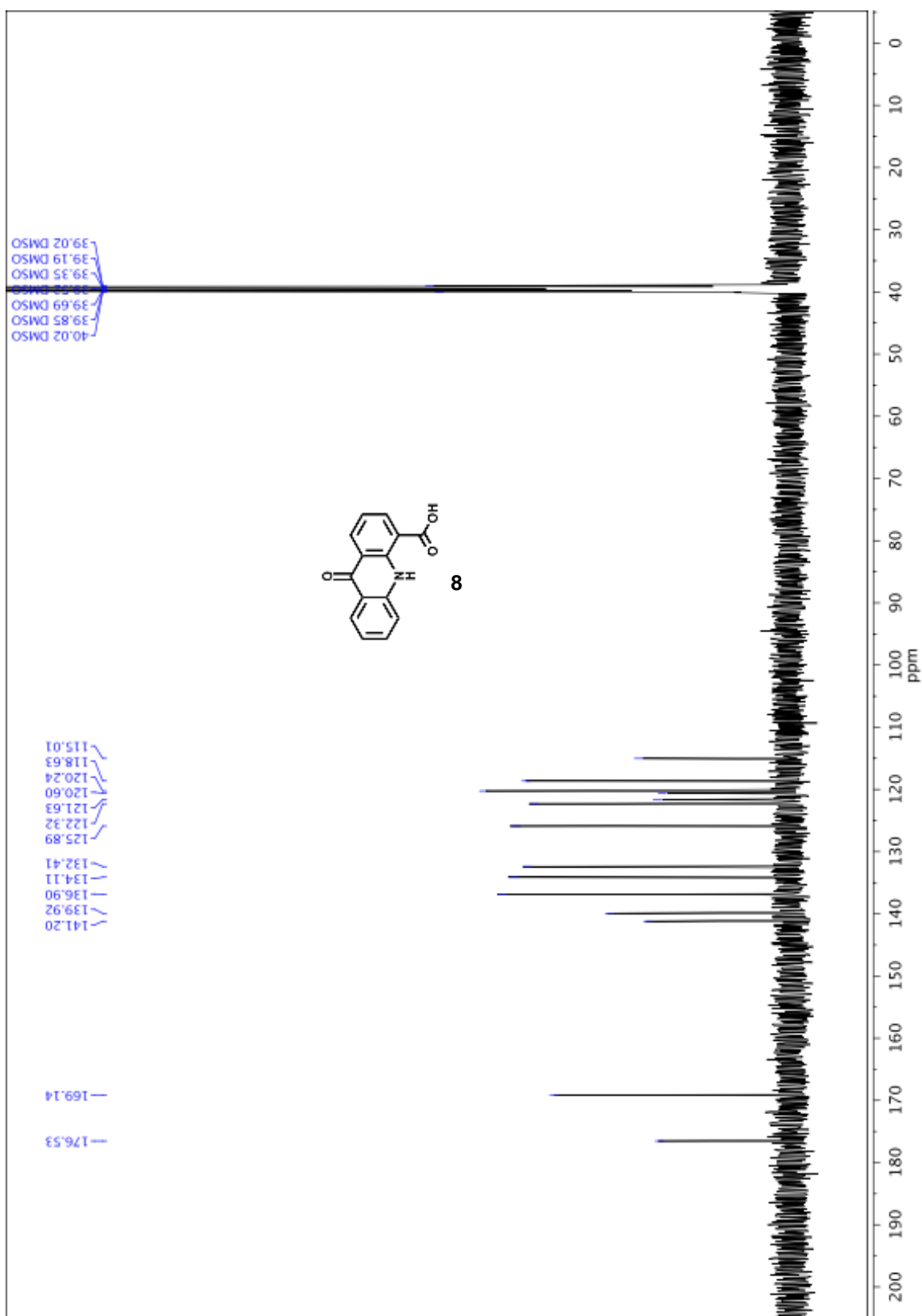


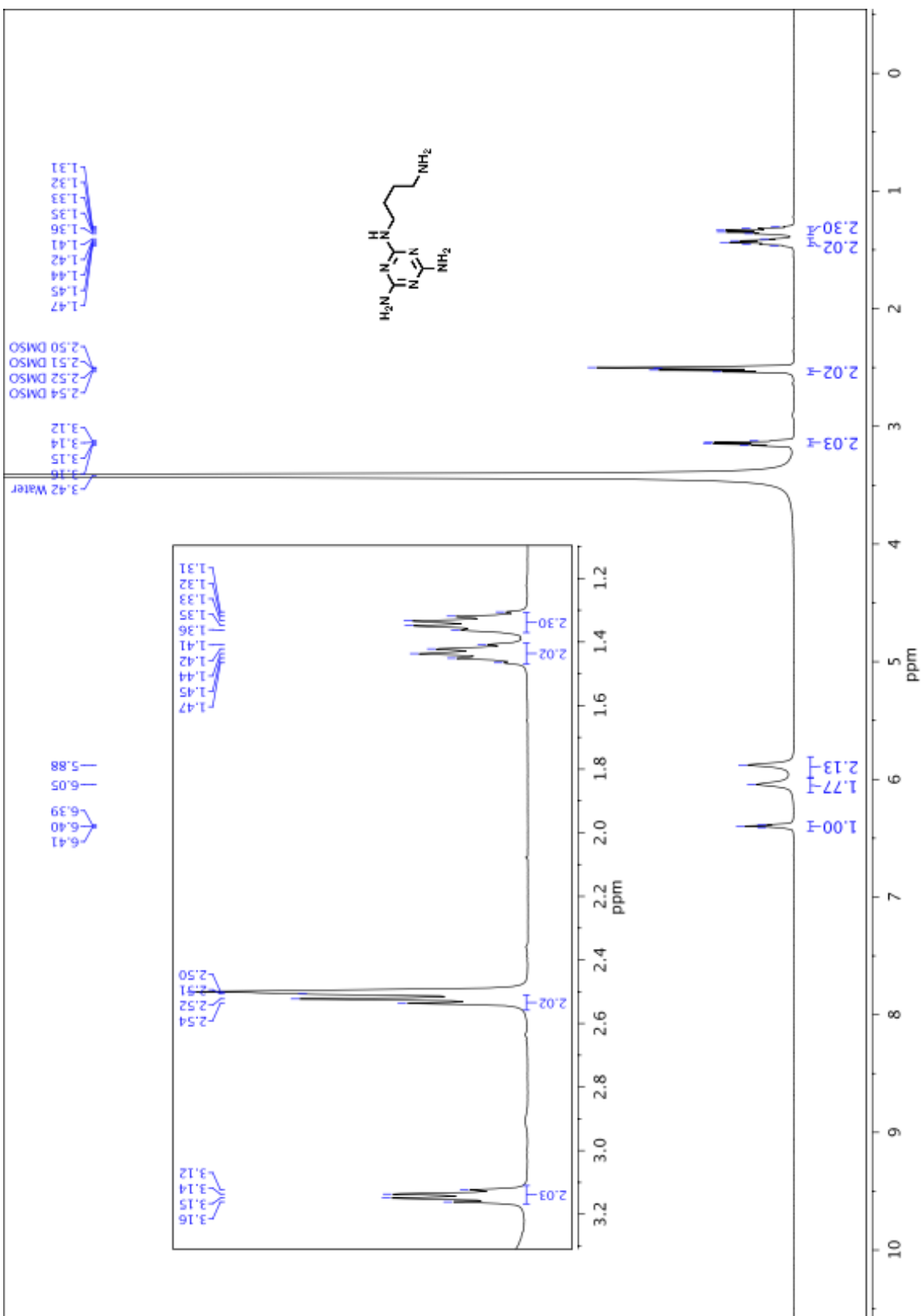


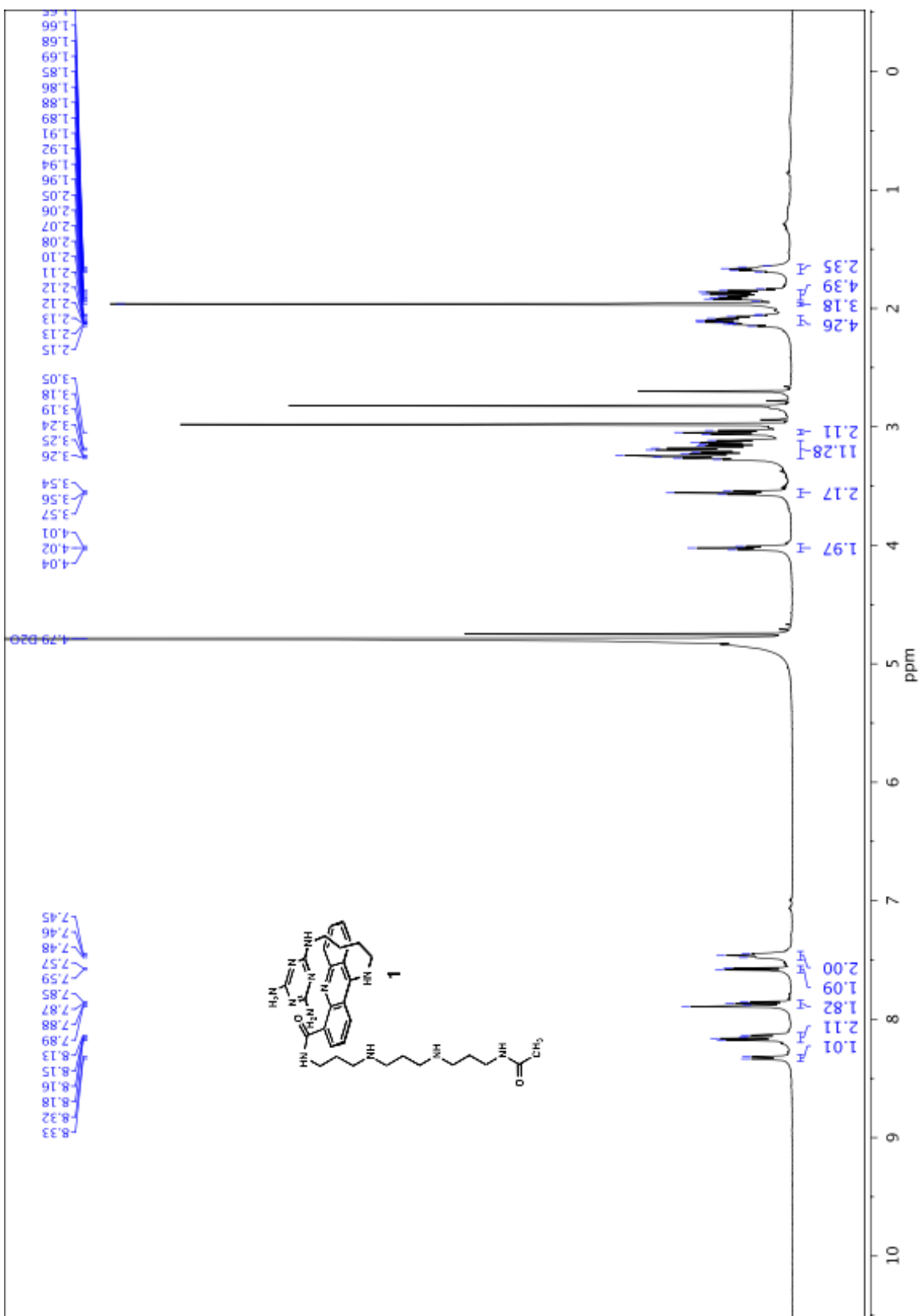


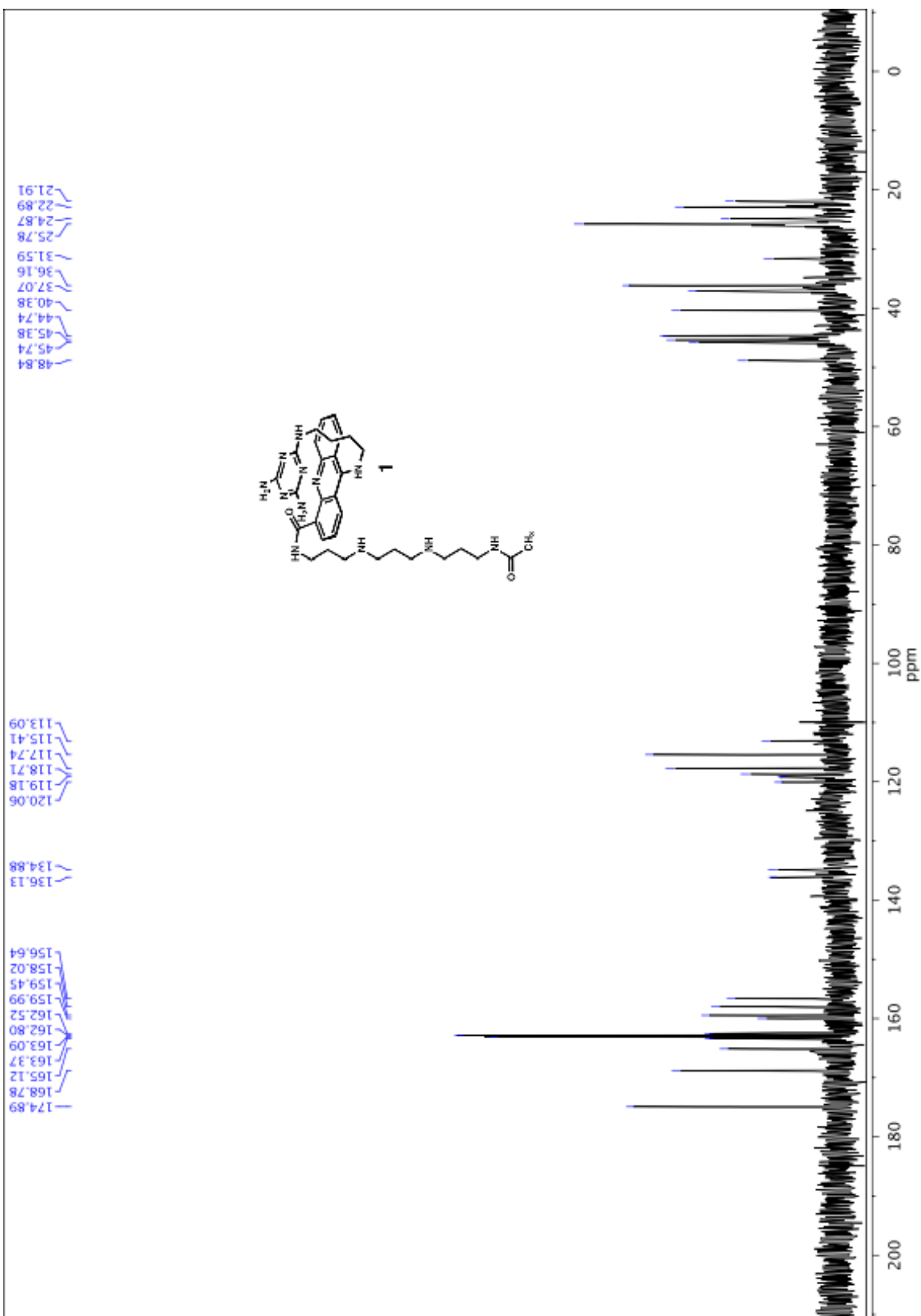










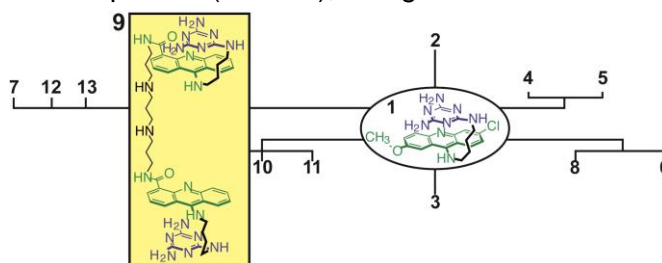


# Chapter 3.

## The Bivalent Ligand Approach Leads to a Bioactive Inhibitor of MBNL1·CUG<sup>exp</sup> Complex

### 3.1 Abstract

An expanded CUG repeat (CUG<sup>exp</sup>) is the causative agent of myotonic dystrophy type 1 (DM1) by sequestering muscleblind-like 1 protein (MBNL1), a regulator of alternative splicing. Inhibition of the pathologic MBNL1·CUG<sup>exp</sup> interaction by targeting aberrant CUG<sup>exp</sup> is the most reliable approach in DM1 drug discovery efforts thus far.



Herein, we present the design, synthesis and *in vitro* and cell-based evaluation of a small library of dimeric ligands based on a ligand that was previously reported to be active in an *in vitro* assay. This library includes 10 dimeric ligands that differ in length, composition and attachment point of the linking chain. From the results of *in vitro* assays, we propose that the potency of dimeric ligands depends on the composition more than the length of spacer. Oligoamino spacers give a greater gain in affinity to CUG<sup>exp</sup> as well as greater inhibition potency of MBNL1·CUG<sup>exp</sup> interaction. The most potent *in vitro* ligands were further evaluated in a DM1 cell model. The optimized bivalent ligand was an aqueous-soluble and cell permeable ligand that displayed almost complete dispersion of CUG<sup>exp</sup> foci. For direct proof of its bioactivity, we evaluated its activity by tracking a live DM1 cell model. This ligand was able to disperse MBNL1 from CUG<sup>exp</sup> foci in individual live DM1 model cells using time-lapse confocal fluorescence microscopy.

### 3.2 Introduction

Previously, ligand 1 was reported as a modest (CUG)<sub>12</sub>·MBNL1 inhibitor (IC<sub>50</sub> = 46 μM) in an *in vitro* assay.<sup>1</sup> However, this ligand had two shortcomings: Modest inhibition potency and more importantly, cell impermeability.



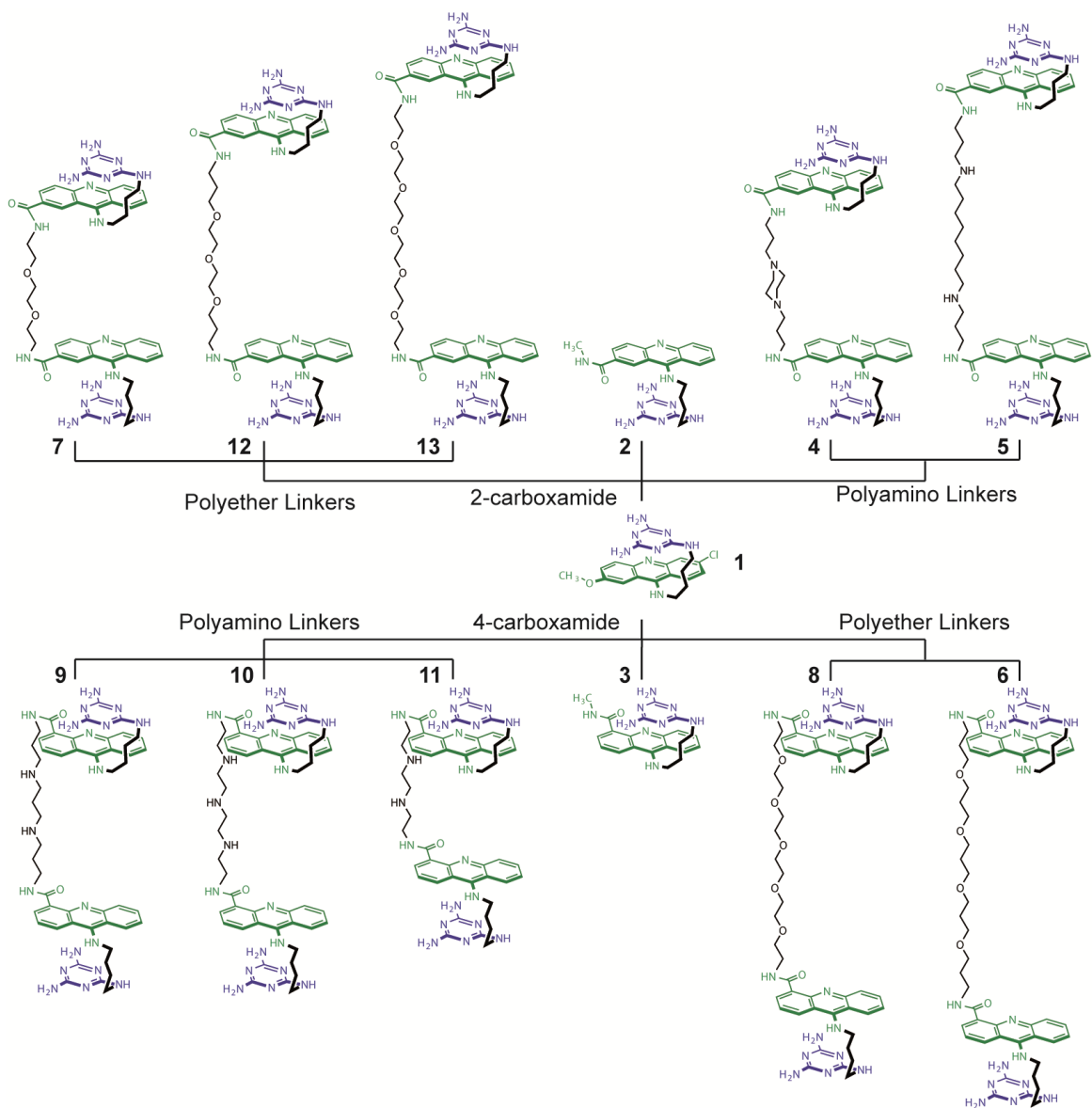
Considering the repeating nature of CUG<sup>exp</sup>, one approach to increase the affinity of this ligand to CUG<sup>exp</sup> is the generation of multivalent ligands.<sup>2, 3</sup> The multivalent effect has been previously proven useful to increase the binding affinity of ligands toward a wide variety of multivalent targets including CUG<sup>exp</sup>.<sup>4-12</sup> Another advantage of multivalent ligands is their potential to become cell-permeable, with appropriate spacers. Development of a bioactive multivalent ligand has some obstacles such as large size and molecular weight.<sup>12</sup> This is a serious drawback in terms of drugability. For example, previously Hoechst 33528 tetramer and pentamers have been developed as inhibitors for this interaction but were found insoluble and cell-impermeable.<sup>13</sup> Therefore this strategy is limited mostly to dimeric ligands with moderate molecular weights. The increase in affinity, attributed to dimeric ligands, arises from the thermodynamic advantage inherent in a cooperative binding system.<sup>3, 14</sup> In other words, the overall entropy of the ligand-CUG<sup>exp</sup> complex is significantly lowered by having the second binding module localized in the vicinity of its respective binding site, upon binding of the first module.<sup>2</sup>

The possibility of gaining increased selectivity and potency through the use of dimeric ligands is an attractive pathway that is becoming more prevalent in drug discovery efforts.<sup>15-20</sup> For almost all cases, due to entropic and enthalpic costs of bivalent binding, the bivalent effect is less than ideal, i.e., square binding constant.<sup>2</sup> On one hand, conformational rigidity can cause spatial mismatch and diminish the binding enthalpy of the second module. On the other hand, conformational flexibility can cause high entropic cost for the binding of the second module. Therefore, both rigidity and flexibility can potentially diminish the bivalent effect. So it is essential to have the right spacer to maximize the bivalent effect. To accomplish this goal, without a structural knowledge of the binding mode, a small library of dimeric ligands was considered necessary. Therefore, we pursued the rational design of a small library of dimeric ligands based on 1.

For the dimeric ligands to be cell-permeable, the composition of the spacers incorporated into the dimeric ligands play an important role. A successful combination of bivalent effect and drugability can yield an optimized dimeric ligand with increased potency and cell-permeability.

### 3.3 Results and Discussion

**Rational Design and Synthesis of dimeric ligands:** The design of a dimeric ligand requires the following selections: (1) a monomeric ligand with a modest affinity and selectivity for CUG<sup>exp</sup> (2) an appropriate handle, that would not interfere with the ligand·CUG<sup>exp</sup> interaction, for the attachment of the spacer to the monomeric ligand, and (3) an appropriate spacer to connect the two monomers so that each module can bind to CUG<sup>exp</sup>. From the previous studies in our lab, it was determined that **1** is a modest inhibitor for (CUG)<sub>12</sub>·MBNL1 (IC<sub>50</sub> = 46 μM) with over 50 fold selectivity for (CUG)<sub>4</sub> relative to a random duplex.<sup>1</sup> The covalent linkage of **1** to a spacer required attachment of a handle to it. The three possible sites for covalent modification of **1** were the acridine ring, the triaminotriazine recognition unit or the linking chain between these two components. Modification of the recognition unit is not favored due to possible interruption of hydrogen binding to UU mismatches. However, various bis-acridine intercalators spanning 2 or more base pairs, have been synthesized and studied before.<sup>21-25</sup> The most expeditious synthetic approach utilized an acridine ring containing a carboxylic acid group to connect two units through a spacer. Positioning the carboxylic acid handle in either position 1 or 3 in the acridine ring is not synthetically feasible because the synthetic route yields an inseparable mixture of both.<sup>26</sup> The remaining two positions of the acridine ring, i.e., positions 2 and 4, are synthetically attractive points for the preparation of dimeric ligands because by positioning the spacer and triaminotriazine recognition units on the same (2-carboxamido derivatives) or the opposite (4-carboxamido derivatives) sides of the acridine unit, the dimeric complex likely requires a non-threading or threading mechanism of binding, respectively, which increases the diversity of our library. It was found that the chloro- and methoxy-groups in **1** could be replaced with a 2- or 4-carboxamido group (see **2** and **3** in Figure 3.1, respectively) without reducing its affinity for CUG repeat or its inhibition potency of the MBNL1·CUG complex.



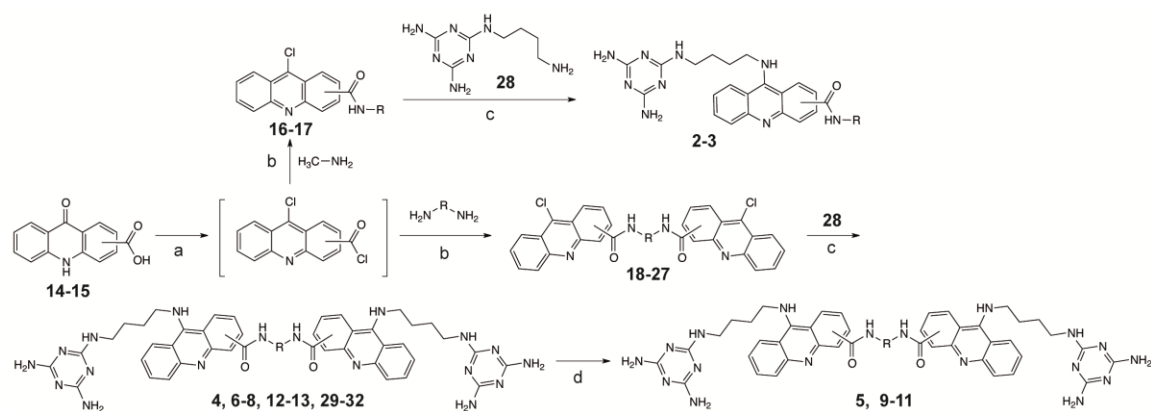
**Figure 3.1.** Library of synthesized dimeric ligands 4-13

Terminal diamine compounds were selected as the spacers for the synthesis of dimeric ligands, as they possess the desired functionality to form a stable amide bond to the carboxylic acid handle on the acridine ring.<sup>27, 28</sup> A rigid spacer can result in an 'all or nothing' effect;<sup>29</sup> whereby an appropriately designed rigid spacer increases the concentration of the second binding module at the UU mismatch site and an inappropriate spacer diminishes the binding of the second module. Without structural data for the binding mode, the spacer should be flexible enough to allow bivalent binding.

The polarity of the spacer is another important consideration when designing the dimeric ligands because solubility issues may arise from the inclusion of a polymethylene chain. The original monomeric ligand is not aqueous soluble, so polyamide and polyalkyl chains were ruled out as potential spacers because they can

become increasingly hydrophobic with increasing length. Thus, heteroatom-rich spacers, oligoamino or oligoether terminal diamines with various lengths from 10 to 21 atoms, were designed. Oligoethers consisted of oligoethylene or oligopropylene glycols. Oligoamino linkers can introduce positive charge to the dimeric ligand and increase its affinity to the polyphosphate backbone.<sup>30</sup> Spacers were designed so that they span two central GC base-pairs according to the nearest site exclusion principle and the two binding modules (triaminotriazine rings) could potentially bind to consecutive CUG sites. Although the binding mode was not firmly established for **1**,<sup>31</sup> it was designed to act as a “stacked intercalator” with the acridine and triaminotriazine rings  $\pi$ -stacked whereas the intercalator sits between the GC base pair and the U-triaminotriazine-U base triplet.<sup>1</sup>

Two series of dimeric ligands at distinct attachment points were synthesized as shown in Figure 3.2 which shows the general route for synthesis of the dimeric ligands. Spacers were either commercially available or synthesized as shown in Schemes 3.1-3.8 (See the Experimental Synthetic Procedures).



| Ring Sub. | Compound | R | Compound | R | Compound | R |
|-----------|----------|---|----------|---|----------|---|
| 2         | 16       |   |          |   | 2        |   |
| 4         | 17       |   |          |   | 3        |   |
| 2         | 18       |   |          |   | 4        |   |
| 2         | 19       |   | 29       |   | 5        |   |
| 4         | 20       |   |          |   | 6        |   |
| 2         | 21       |   |          |   | 7        |   |
| 4         | 22       |   |          |   | 8        |   |
| 4         | 23       |   | 30       |   | 9        |   |
| 4         | 24       |   | 31       |   | 10       |   |
| 4         | 25       |   | 32       |   | 11       |   |
| 2         | 26       |   |          |   | 12       |   |
| 2         | 27       |   |          |   | 13       |   |

<sup>a</sup>Reagents and conditions: (a) SOCl<sub>2</sub>, DMF (Cat.), 70 °C 2 h; (b) DCM, NEt<sub>3</sub>, 0 °C to RT, 15 h; 60-78%; (c) DMF, DIPEA, 80 °C, 6 h; 55-75%; (d) CF<sub>3</sub>COOH, DCM, RT, 6 h, 100%.

**Figure 3.2.** General synthetic scheme for dimeric ligands

**Optical Melting Experiments.** It has been proposed that MBNL1 displays a preference for single stranded (ss) CUG<sup>exp</sup>, and, thus, destabilizes the double stranded (ds) CUG<sup>exp</sup> upon binding.<sup>32, 33</sup> If this model is correct, any ligand that stabilizes the ds form of CUG<sup>exp</sup> may prove a more effective inhibitor of the MBNL1·CUG<sup>exp</sup> interaction. The increase in (CUG)<sub>12</sub> melting temperature upon ligand binding correlates with the (CUG)<sub>12</sub> stem loop stability, therefore we studied the effect of ligands on the T<sub>m</sub> of (CUG)<sub>12</sub>. Ligands **4**, **5** and **6** were not fully soluble in the reaction condition; therefore we could not get any ΔT<sub>m</sub> for them. We carried out thermal denaturation study of (CUG)<sub>12</sub>, in the presence of 1 equivalent of other ligands. Monophasic melting curves were obtained;

$\Delta T_m$  values are shown in table 3.1. For the monomeric ligands, **2**,  $\Delta T_m$  was 1.8 °C. For two of oligoether-spacer containing dimeric ligands, **7** and **8**,  $\Delta T_m$  was 2.7 and 1.8 °C respectively, close to that of the monomeric ligand, suggesting these ligands are binding to (CUG)<sub>12</sub> only with one of their binding modules. This can be explained by the possible coil-like conformation of oligoethylene glycol spacers,<sup>34</sup> which prevents the desired distance between binding modules. Oligopropylene glycol spacer in **6** was designed to prevent this coil-like conformation however it happened to be aqueous insoluble. Three oligoamino-spacer containing dimeric ligands, **9**, **10** and **11**, had a higher  $\Delta T_m$  of 9.3, 9.6 and 8.4 °C, respectively, suggesting a bivalent binding to (CUG)<sub>12</sub> and a greater stabilization of the ds form of (CUG)<sub>12</sub>. This finding suggests that **9**, **10** and **11**, unlike **7** and **8**, have the right type of spacer to bind to (CUG)<sub>12</sub> by two binding modules and stabilize the ds form of (CUG)<sub>12</sub> to a greater extent relative to the corresponding monomer. This finding shows the importance of the composition of the spacer is greater, rather than the number of atoms. Oligoethylene glycol spacers with similar numbers of atoms to oligoamino spacers failed to give the dimeric ligands the same  $\Delta T_m$  value. Although this is an encouraging result,  $\Delta T_m$  is considered a qualitative parameter.

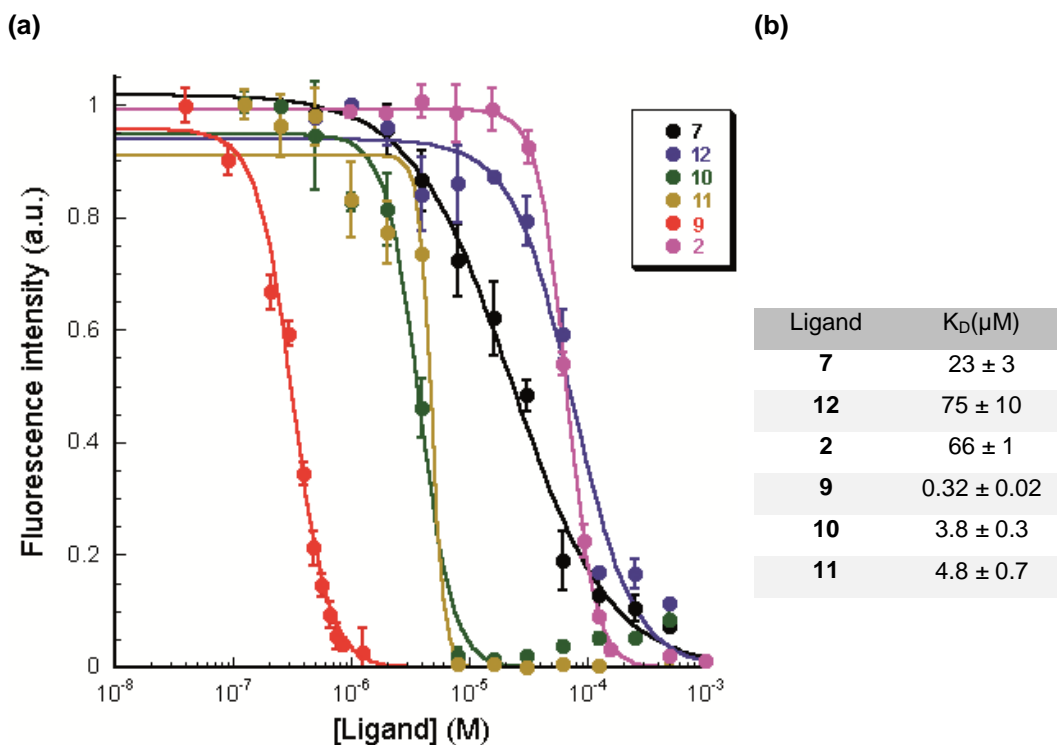
**Table 3.1.**  $\Delta T_m$  value of (CUG)<sub>12</sub> upon adding ligand in a 1:1 ratio.

| Ligand    | Solution  | $\Delta T_m$ (°C) |
|-----------|-----------|-------------------|
| <b>2</b>  | 5% DMSO   | 1.8 ± 0.6         |
| <b>7</b>  | 5% DMSO   | 2.7 ± 0.4         |
| <b>4</b>  | Insoluble | _____             |
| <b>12</b> | 5% DMSO   | N.D.              |
| <b>13</b> | 5% DMSO   | N.D.              |
| <b>5</b>  | Insoluble | _____             |
| <b>8</b>  | 5% DMSO   | 1.8 ± 0.2         |
| <b>6</b>  | Insoluble | _____             |
| <b>9</b>  | Aqueous   | 9.3 ± 0.6         |
| <b>10</b> | Aqueous   | 9.6 ± 1.3         |
| <b>11</b> | Aqueous   | 8.4 ± 2.1         |

**Steady State Fluorescence Studies.** To measure the binding affinity of the ligands to CUG repeat, a steady state fluorescence titration method with TAMRA-(CUG)<sub>6</sub> was utilized. It is known that guanosine quenches the TAMRA fluorescence through photoinduced electron transfer.<sup>35, 36</sup> Therefore, it is possible that binding of a ligand to the UU mismatch close to 5'-TAMRA makes a structural change that can lead to

quenching of the 5'-TAMRA by the 3'-G in 5'-TAMRA-(CUG)<sub>6</sub>-3'. Ligands were titrated into TAMRA-(CUG)<sub>6</sub> solution. Upon increasing the ligand concentration, TAMRA fluorescence intensity gradually decreased as a result of fluorophore quenching by the bound ligand. A plot of normalized fluorescence intensity versus increased concentrations of each ligand yielded a binding isotherm and K<sub>D</sub> value (Figure 3.3) which show a trend similar to the ΔT<sub>m</sub>.

Two of the dimeric ligands with oligoether-type spacers were evaluated. Ligand **12** gave a K<sub>D</sub> close to the monomeric ligand **2** and **7** showed only 3-fold decrease in K<sub>D</sub>. Whereas dimeric ligands with oligoamino-type spacer, **11**, **10** and **9**, show 14, 17 and 206 fold decrease in K<sub>D</sub>, respectively. Although **9** and **10** have the same atom numbers, the greater binding affinity of **9** might be because of the electrostatic interaction of the correctly positioned amino group in the spacer with the polyphosphate backbone of TAMRA-(CUG)<sub>6</sub>.<sup>30</sup>

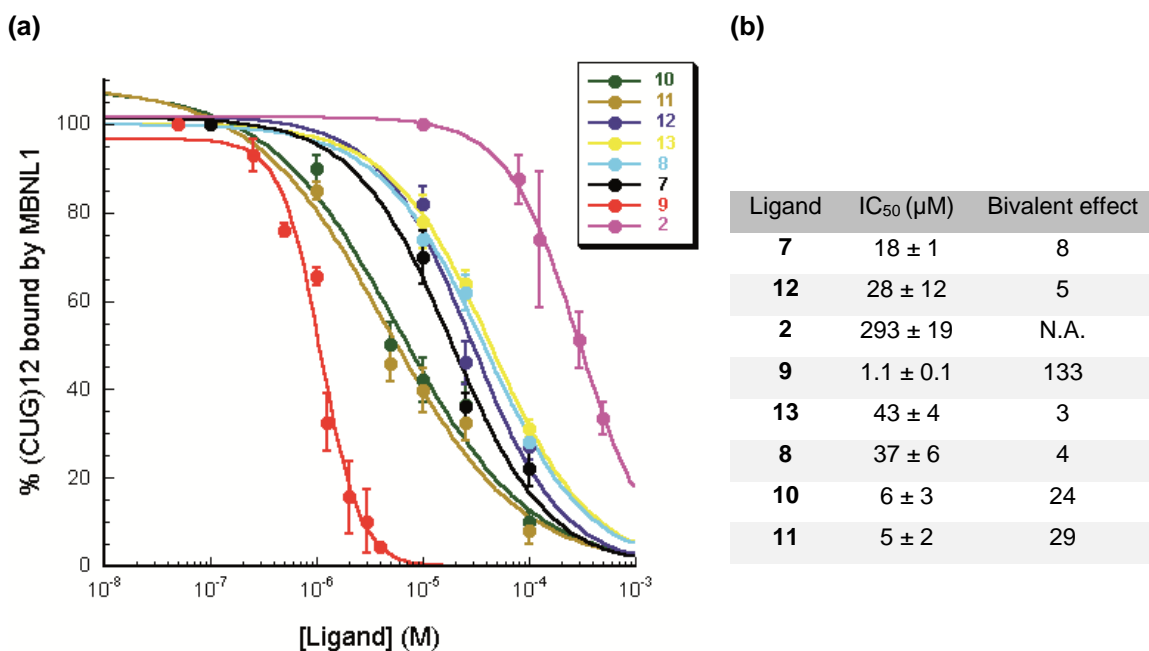


**Figure 3.3.** a) Fluorescence titrations of TAMRA-(CUG)<sub>6</sub> with ligands. Comparison of normalized fluorescence intensity change of TAMRA-(CUG)<sub>6</sub> in the presence of increasing concentrations of ligands. TAMRA was excited at 560 nm and its emission was recorded at 590 nm. Error bars represent mean ± s.d. of three replicates. b) K<sub>D</sub> was derived by fitting the Fluorescence intensities at different concentrations of ligands into the following equation:  $F = \frac{F_{max} - F_0}{1 + (\frac{K_D}{[L]})^n} + F_0$

**Inhibition of MBNL1-CUG Interaction Using a SPR-Based Biosensor.** Because SPR technique is particularly well suited for quantifying the binding of proteins to a target on the SPR chip, we developed a simple SPR-based method to directly measure MBNL1 complexation of (CUG)<sub>12</sub> in real time under equilibrium condition and in a label-free format. Further, we were able to quantify the inhibition potency of ligands. To rule out non-specific inhibition due to aggregation and non-selective RNA binding the assays were performed in the presence of Tween-20 and excessive competitor t-RNA.<sup>37</sup> Thus, biotinylated (CUG)<sub>12</sub> was immobilized on a streptavidin coated SPR sensor chip and incubated with different concentrations of each ligand to reach a steady state response units (RU) over 150 s. The response to the binding of each ligand was negligible in comparison to protein binding so the direct contribution of each ligand could be ignored. Successive injections of a 0.65  $\mu$ M solution of MBNL1 containing the same concentration of each ligand as in the pre-incubation, led to varying responses depending on the concentration of each ligand. Because the SPR signal directly reflects the binding of MBNL1 to the biotinylated (CUG)<sub>12</sub>, the differences in the response curves are a direct result of inhibition by each ligand. The maximum RU at 150 s was recorded for each concentration of ligands and converted to the fraction of (CUG)<sub>12</sub> bound by MBNL1, all values normalized to that measured in the absence of each ligand. Fitting the data points in the plot of %(CUG)<sub>12</sub> bound by MBNL1 versus increasing concentrations of each ligand (Figure 3.4a) gave an apparent IC<sub>50</sub> value for each ligand (Figure 3.4b).

Normalized IC<sub>50</sub> values can be calculated by dividing the actual IC<sub>50</sub> by the number of binding modules it has. The ratio between normalized IC<sub>50</sub> values of monomeric and dimeric ligand is called bivalent effect. The IC<sub>50</sub> values (Figure 3.4) show that more than the spacer length, its composition is critical for an optimal bivalent effect. The dimeric ligands with oligoether-type spacer, **13**, **8**, **12** and **7**, show a small bivalent effect of 3, 4, 5 and 8, respectively, whereas dimeric ligands with oligoamino-type spacer, **9**, **10**, **11**, show a large bivalent effect of 24, 29 and 133, respectively.

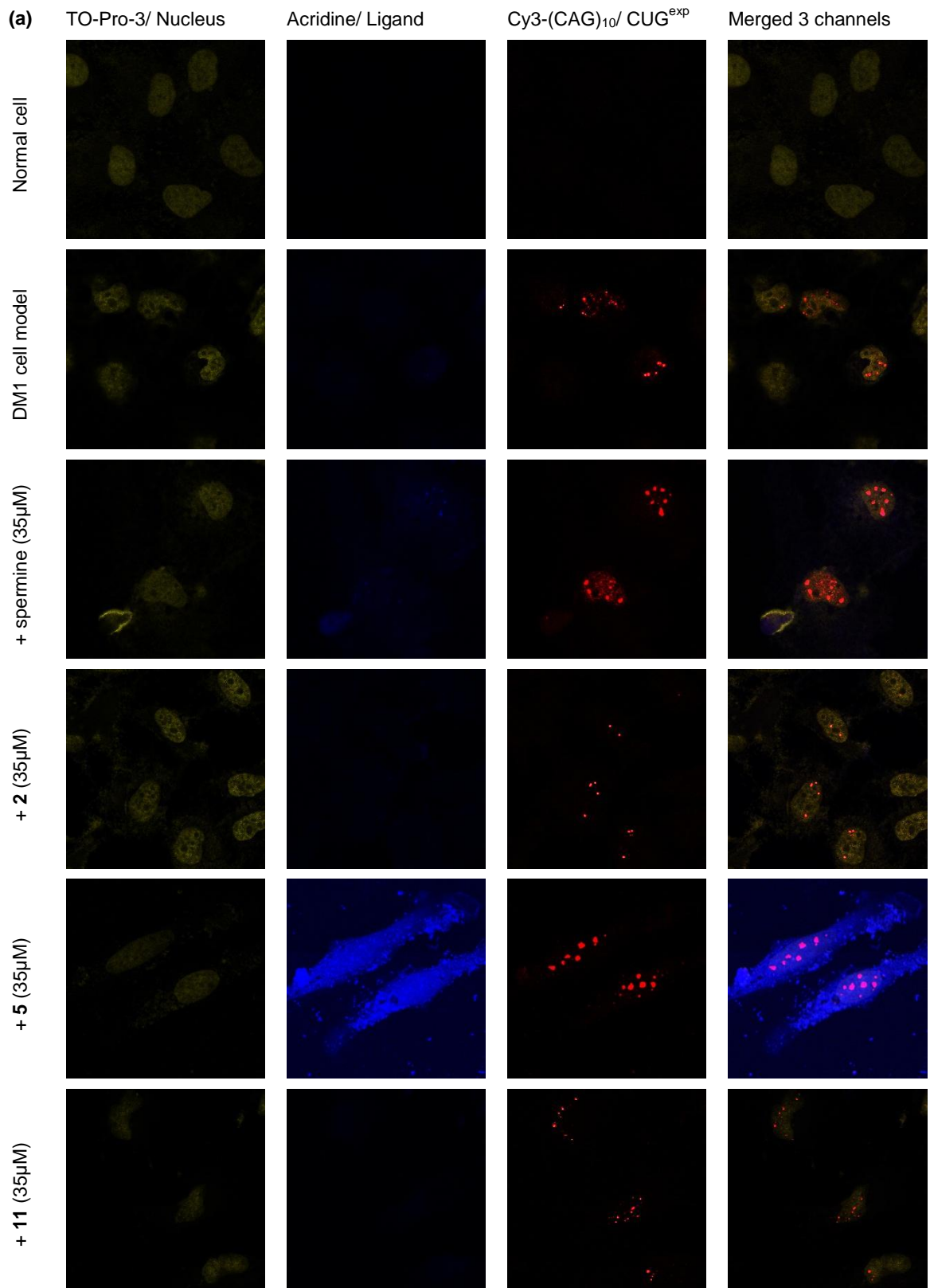


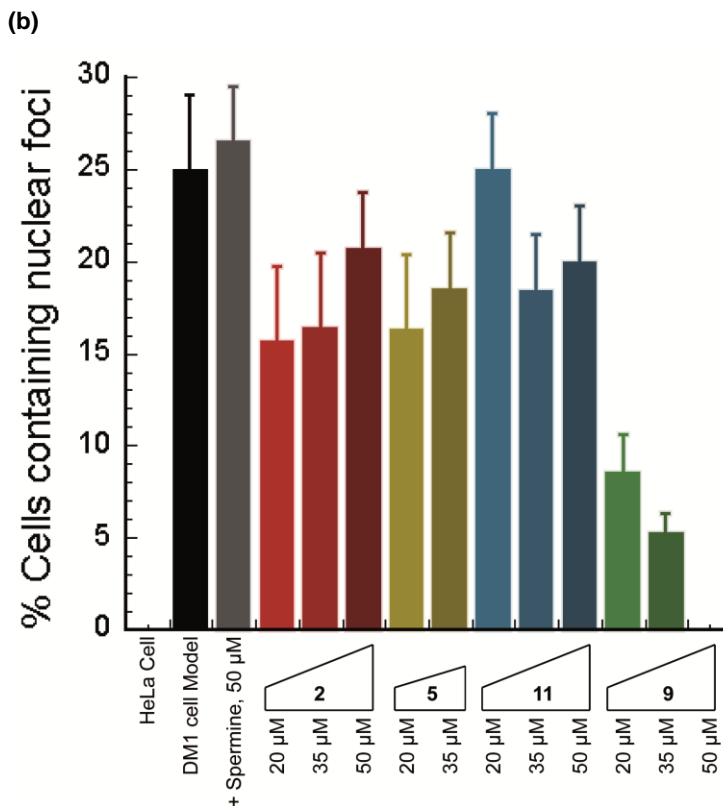
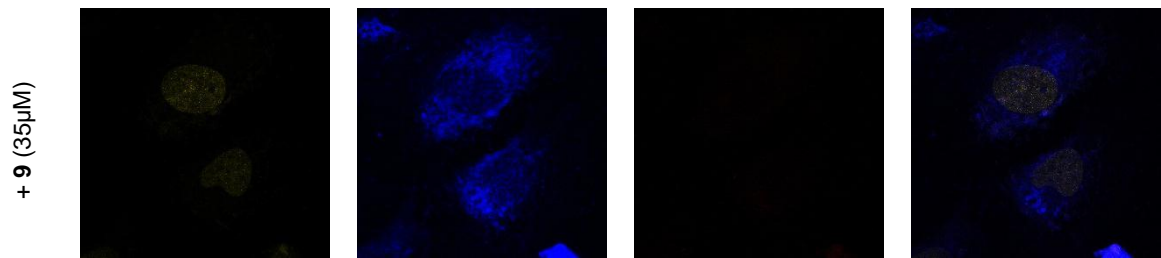


**Figure 3.4.** a) Fitting the response unit data in the plot of % (CUG)<sub>12</sub> bound by MBNL1 versus increasing concentrations of each ligand to a dose-response curve by SPR assay. b) IC<sub>50</sub> values and bivalent effects were derived. Error bars represent mean ± s.d. of three replicates.

These *in vitro* experiments demonstrate that **9** can bind to (CUG)<sub>12</sub>, stabilize the hairpin structure and inhibit the (CUG)<sub>12</sub>·MBNL1 interaction, selectively. It is noteworthy that all of the *in vitro* experiments above were carried out with (CUG)<sub>12</sub> in 1X PBS buffer. This particular buffer was chosen because it is the closest of common buffers to physiological conditions. It is also a more challenging buffer for small molecule inhibitors because it increases the (CUG)<sub>12</sub>·MBNL1 stability.

**Ligand 9 is active in nuclear foci dispersion in CUG repeat-transfected cells.** After successful *in-vitro* studies, we sought to study the most potent dimeric ligands, **9**, in cell-based assays. Two other dimeric ligands, **5** and **11**, and the corresponding monomeric ligand, **2**, were studied for a comparison. Spermine, which has a similar structure to the spacer in **9**, was studied as a negative control. We studied cytotoxicity of **9** as well as **2** in HeLa cells and both showed IC<sub>50</sub> values higher than 100 μM by Sulforhodamine B in 24 hours. The cellular permeability of **9** as well as **2**, **5** and **11** was tested by fluorescence microscopy tracking the fluorescence of acridine ring in the ligands. We found that **9** and **5** are cell-permeable whereas **2** and **11** are not (Figure 3.5, column 2).

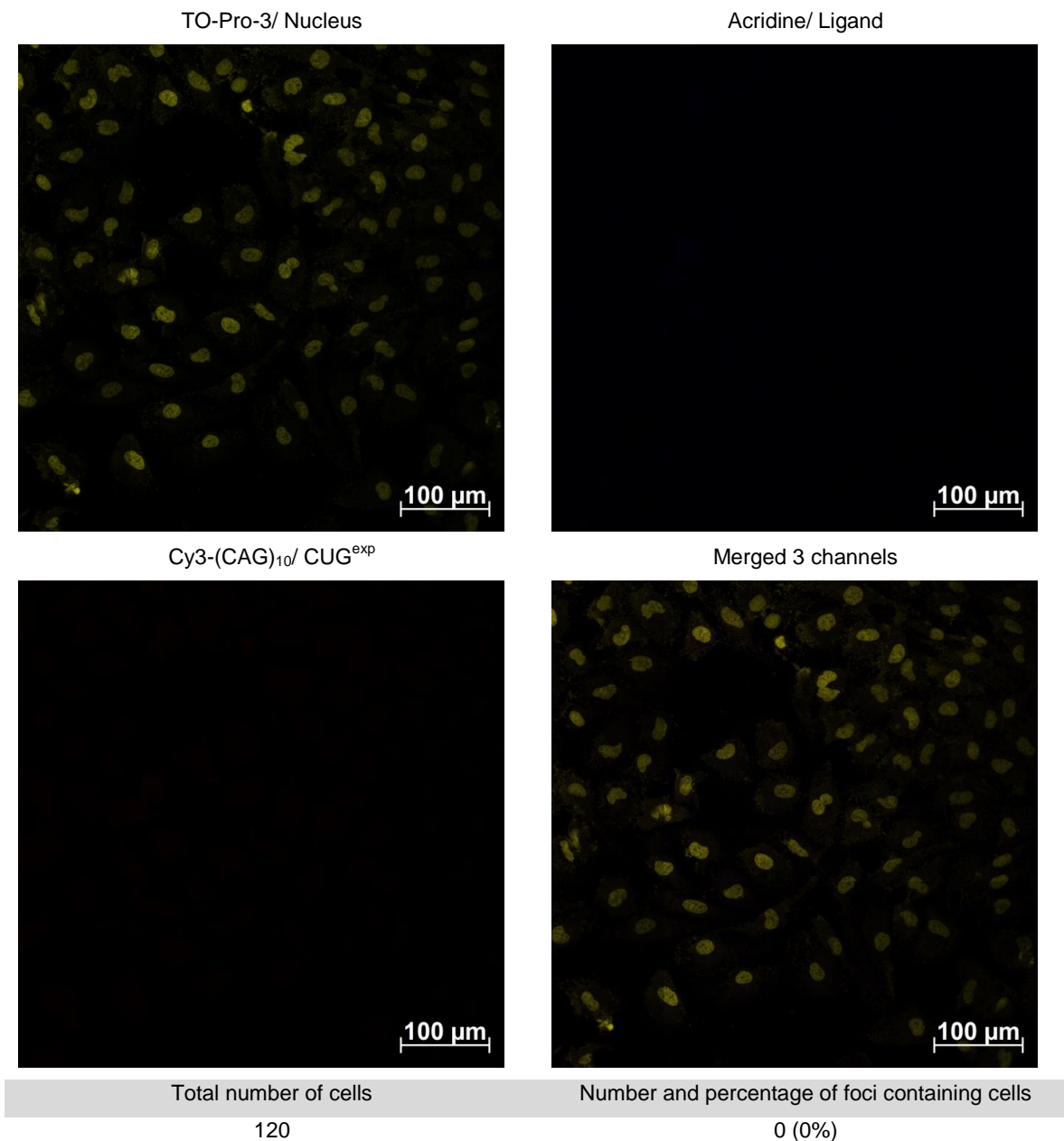




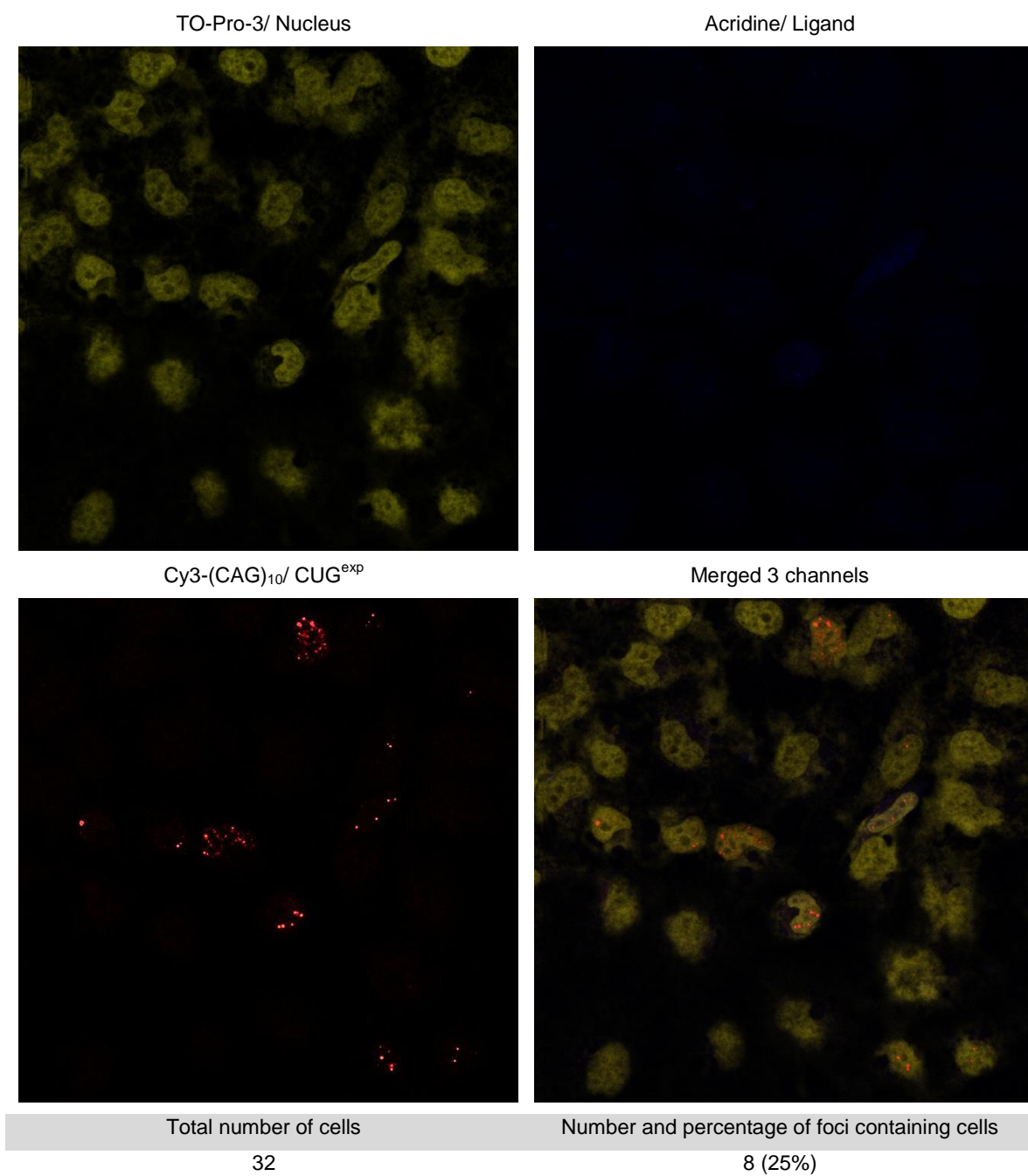
**Figure 3.5.** a) Confocal fluorescent images show CUG<sup>exp</sup> foci in DM1 model cells are present in rows 2-6, no ligand treatment, treated with spermine (negative control), **2**, **9** and **5**, respectively. CUG<sup>exp</sup> foci are not present in negative control cells, row 1, as well as row 7 where DM1 cell model is treated with **9**, for 48 h. b) Plot of CUG<sup>exp</sup> foci-containing cell fraction at various concentrations of ligands. These data are gathered from scoring over 100 cells. The error bars represent mean  $\pm$  standard error of at least three independent experiments.

CUG<sub>960</sub>-transfected HeLa cells were used as our DM1 model cells. They were treated with three different concentrations, 20, 35 and 50  $\mu$ M, of each ligand for 36 h followed by Fluorescence in situ hybridization (FISH) study with Cy3-CAG<sub>10</sub> probe. This probe targets CUG<sub>960</sub> nuclear foci and its fluorescence was tracked by confocal fluorescence microscopy. We found **2**, **5** and **11**, and our negative control, Spermine, show no significant foci dispersion in any concentration; however we observed partial foci

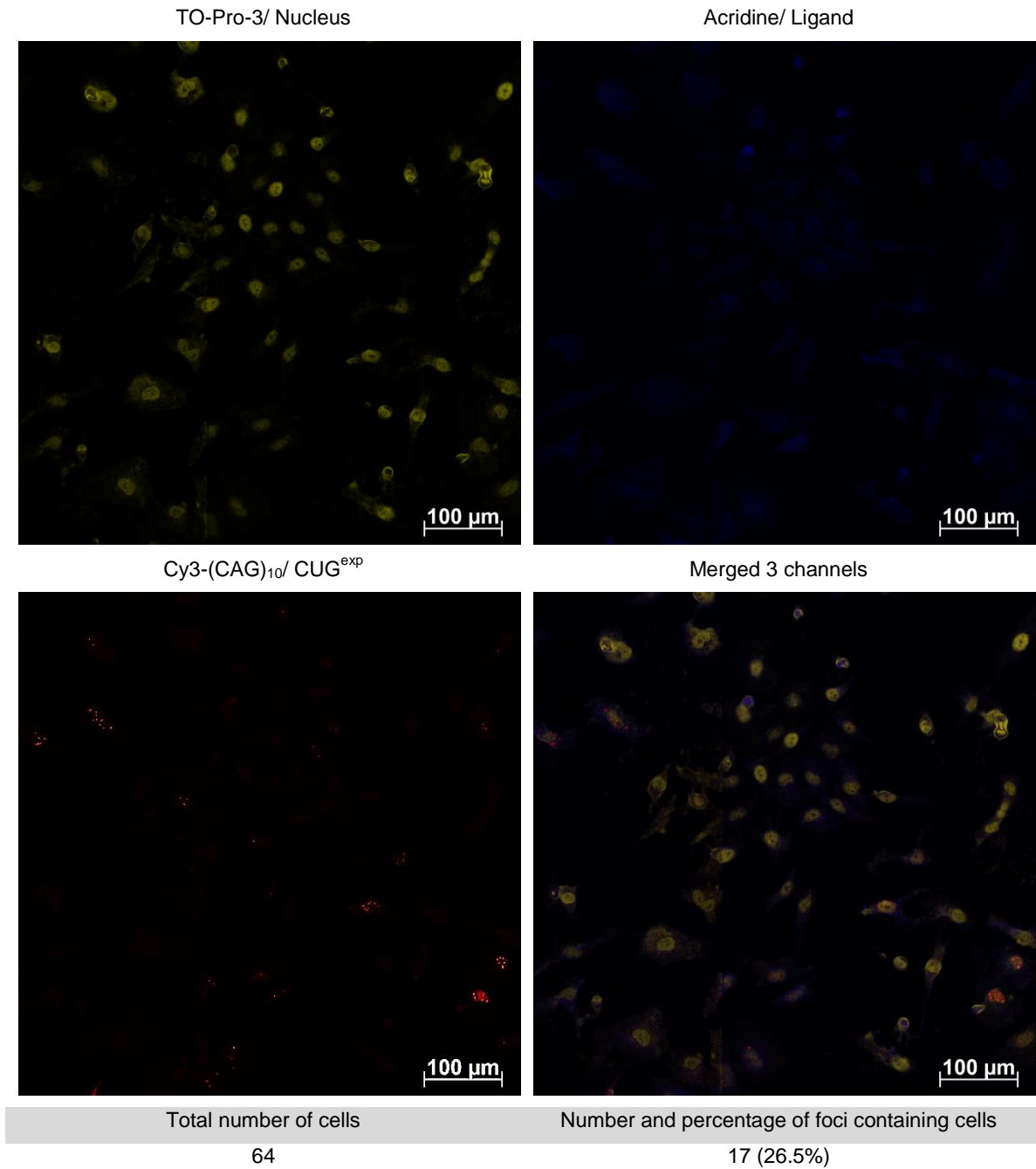
dispersion for **9** at 20  $\mu\text{M}$  and 35  $\mu\text{M}$  and full foci dispersion for **9** at 50  $\mu\text{M}$  concentration (Figure 3.5a, column 3). This study of the foci-containing fixed cells was a statistical comparison and represents an indirect measurement of the foci dispersion by ligands.



**Figure 3.6.** HeLa cells transfected with CUG<sub>0</sub> served as our negative control cell.



**Figure 3.7.** HeLa cells transfected with CUG<sub>960</sub> served as our DM1 cell model.



**Figure 3.8.** DM1 cell model treated with Spermine (negative control compound), 50  $\mu$ M.

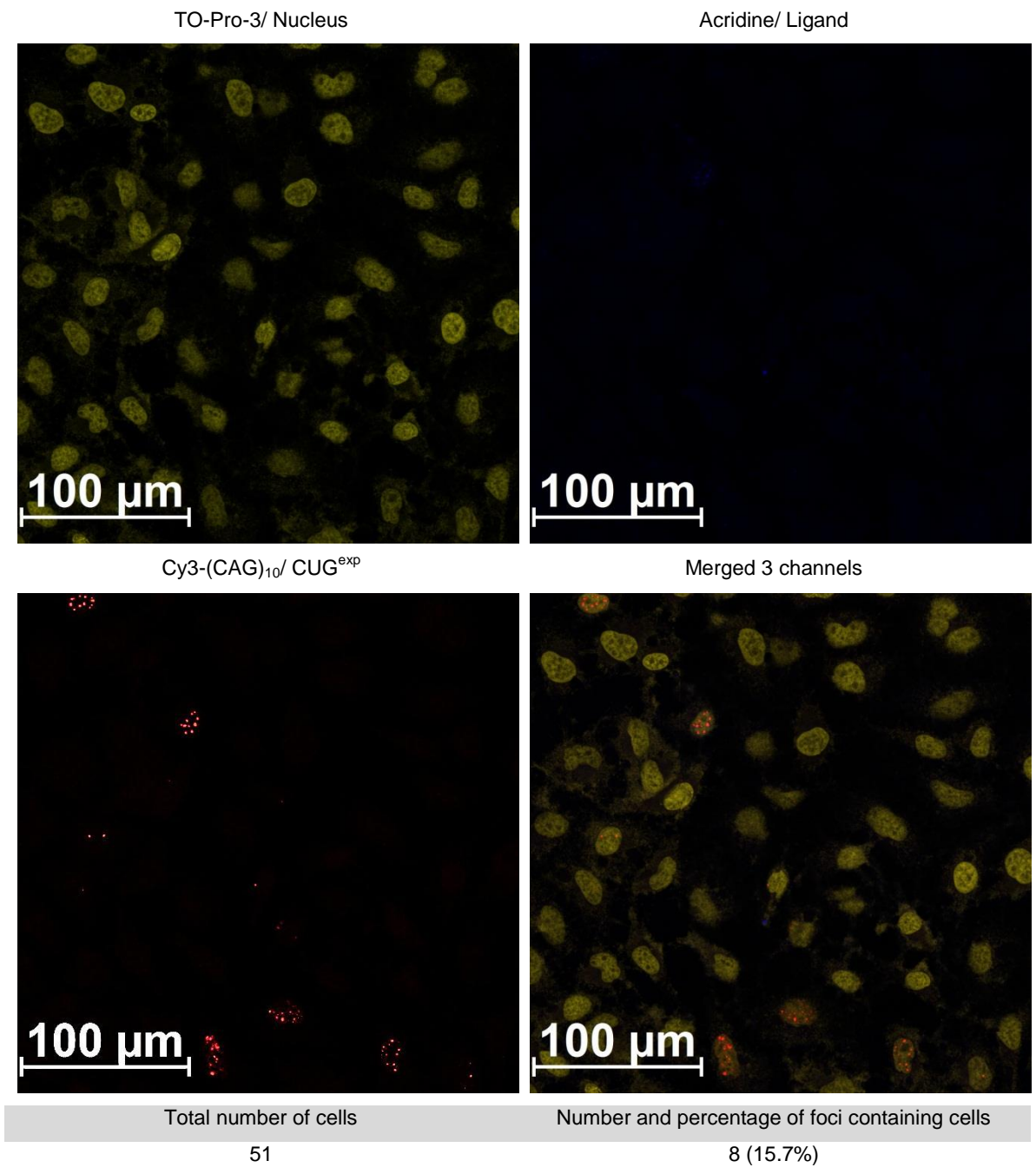
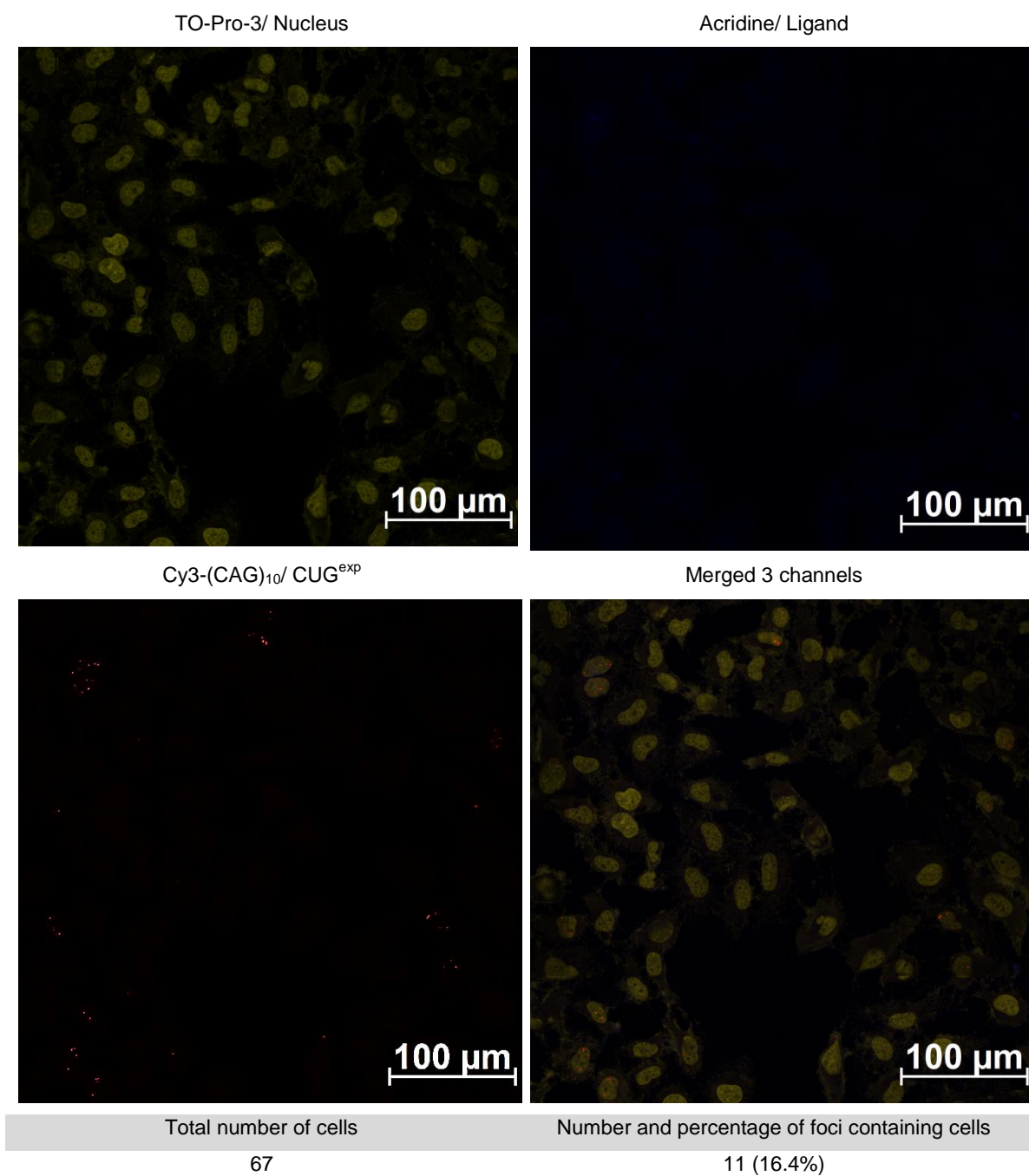


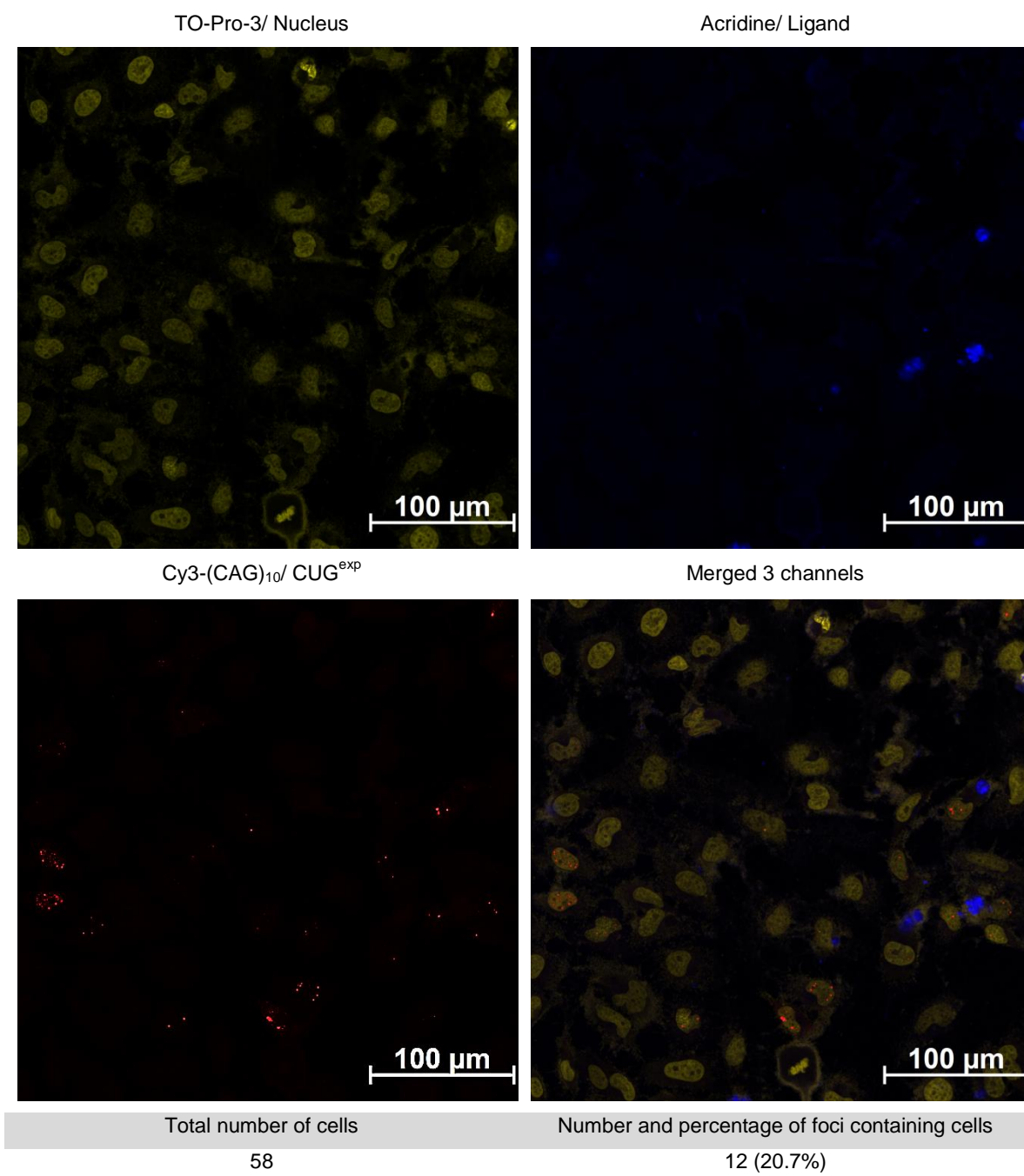
Figure 3.9. DM1 cell model treated with 2, 20 μM.



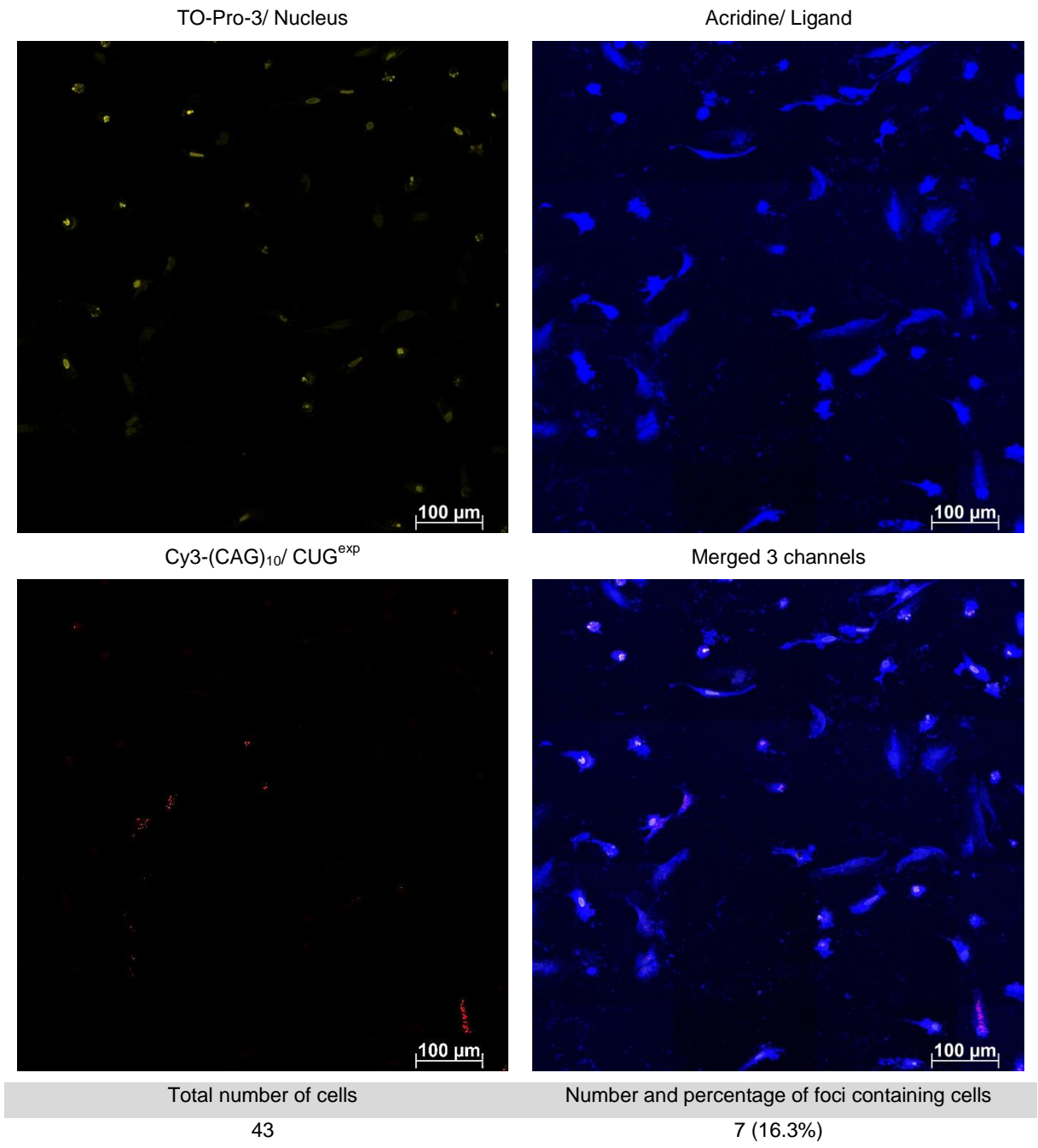


**Figure 3.10.** DM1 cell model treated with **2**, 35 μM.

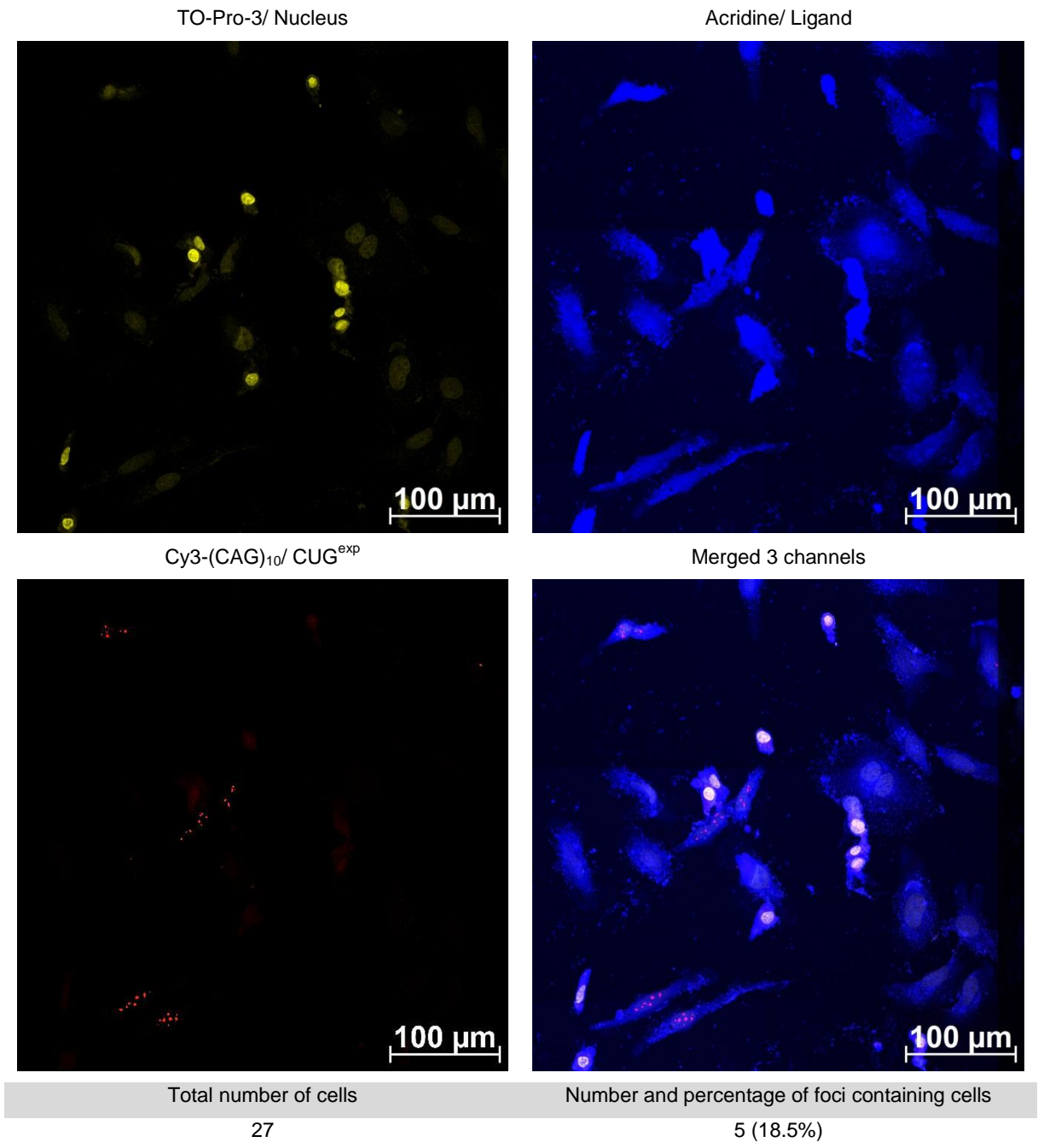




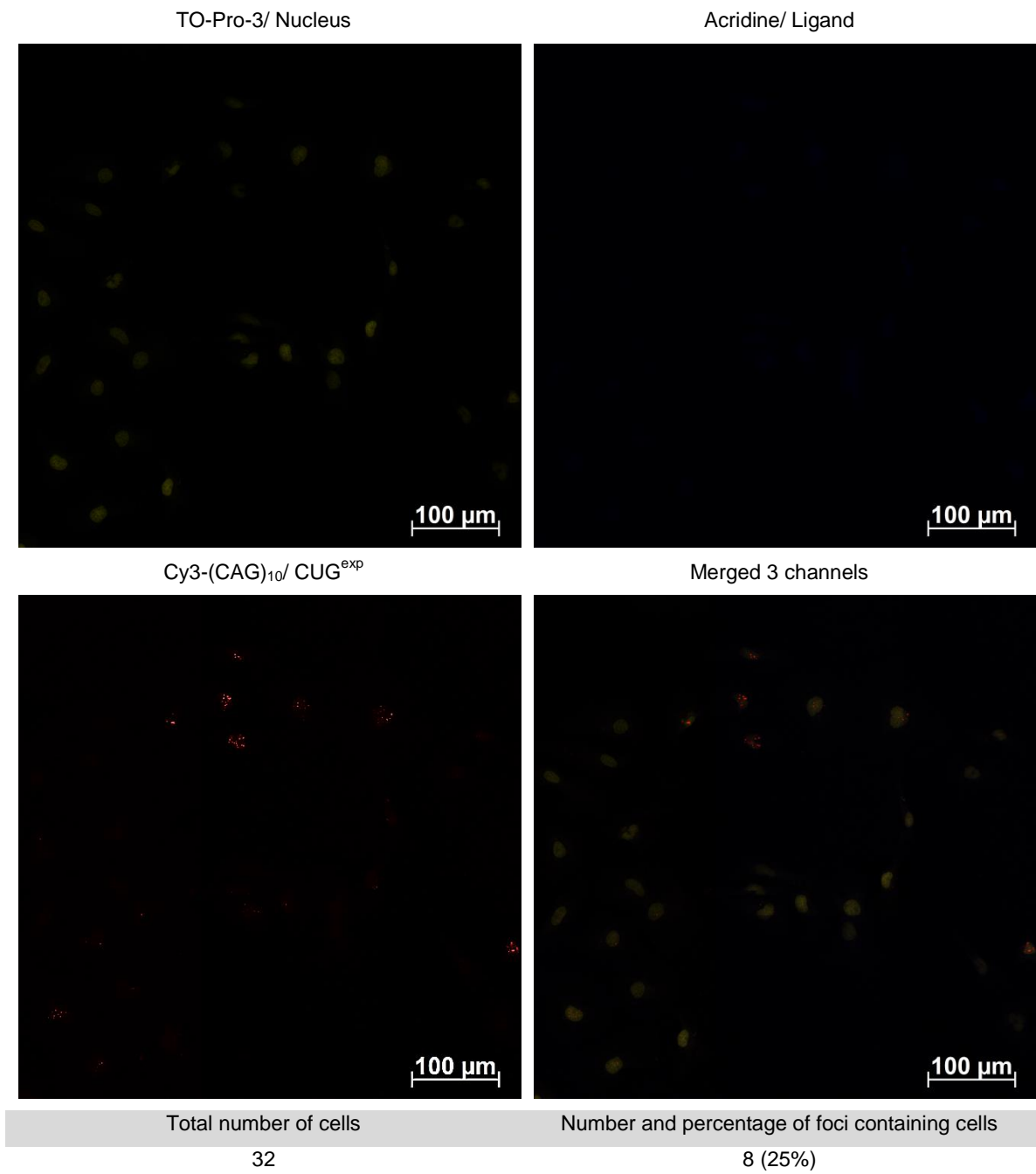
**Figure 3.11.** DM1 cell model treated with **2**, 50 μM.



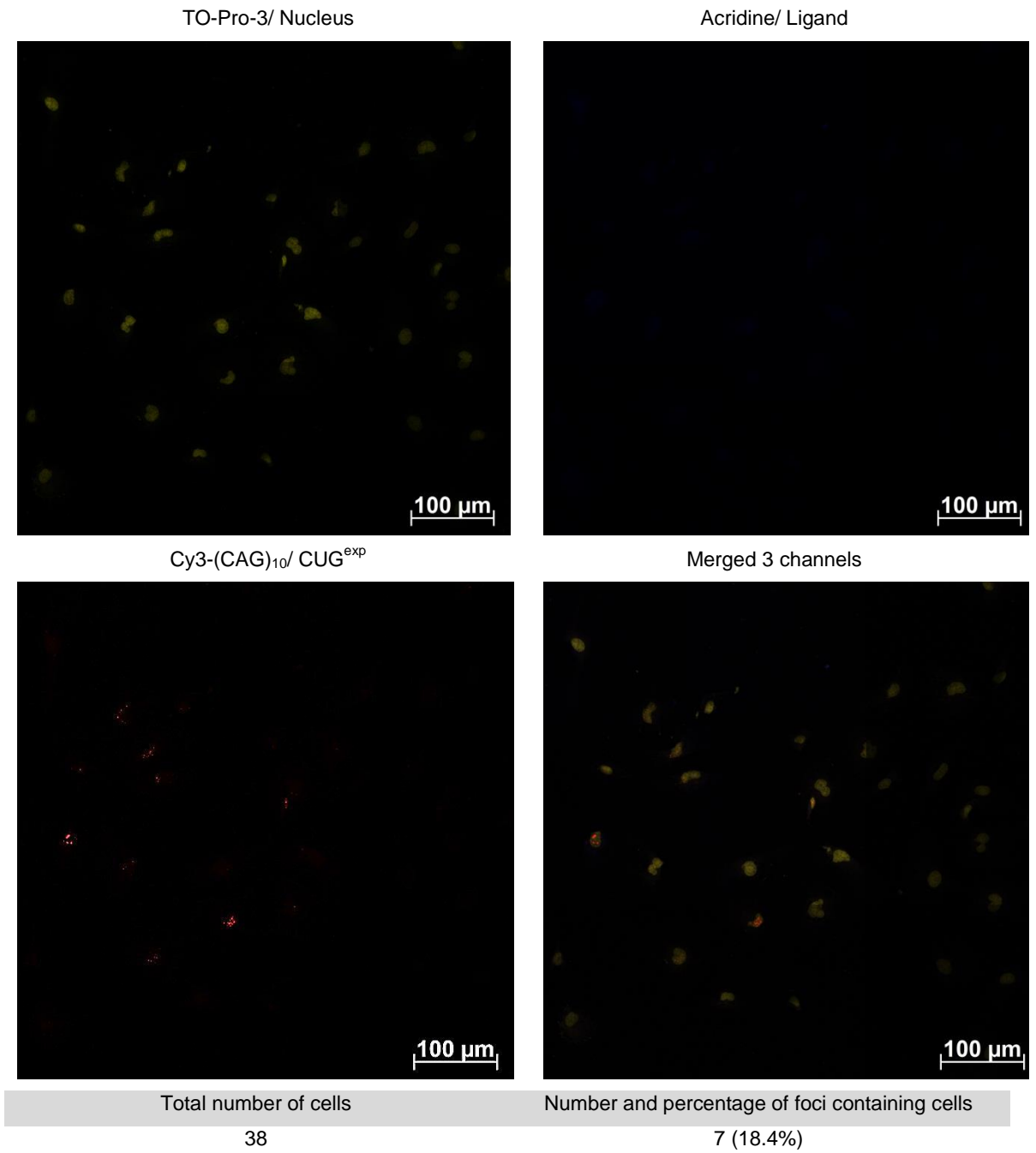
**Figure 3.12.** DM1 cell model treated with 5, 20 μM.



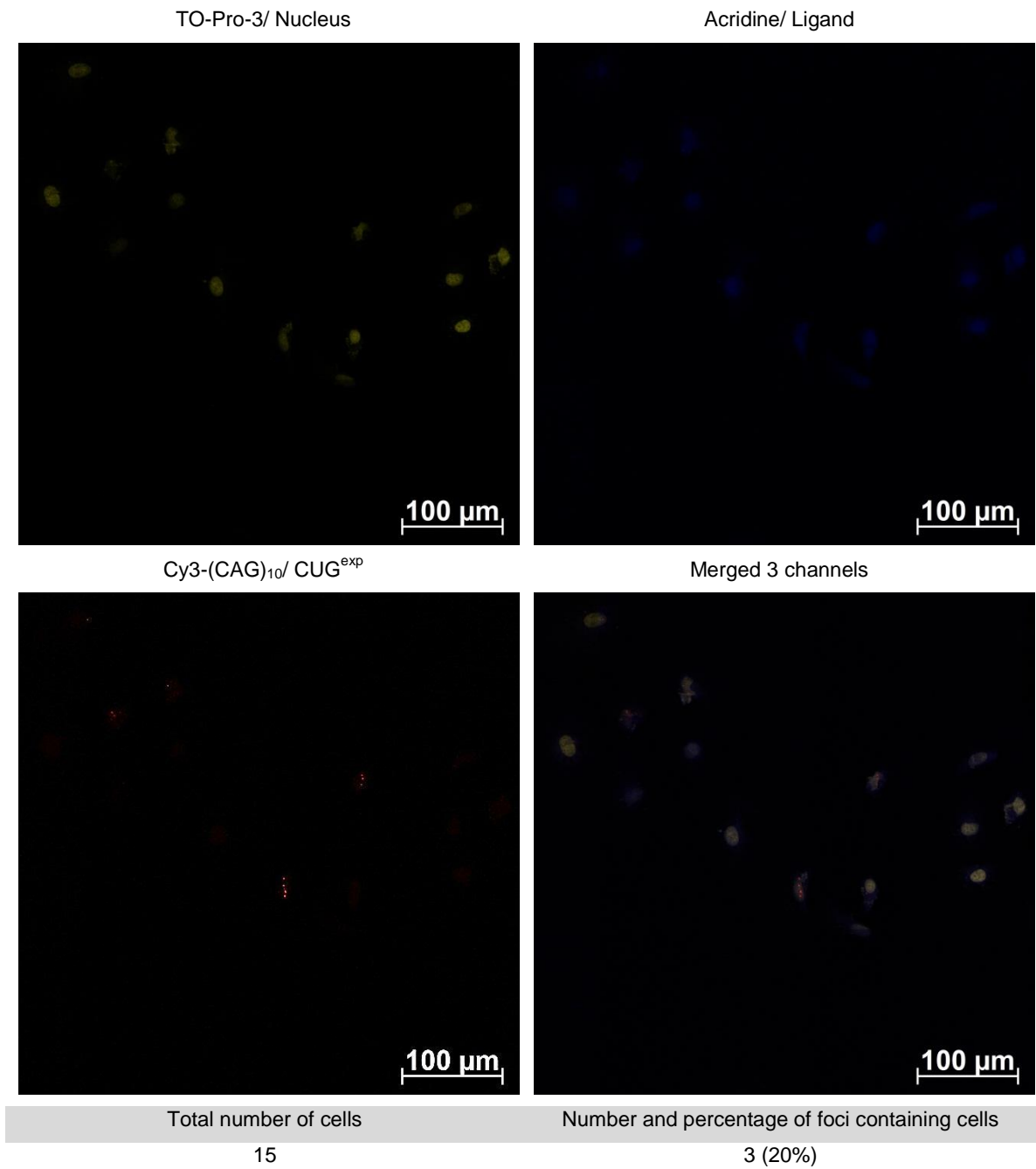
**Figure 3.13.** DM1 cell model treated with **5**, 35 μM.



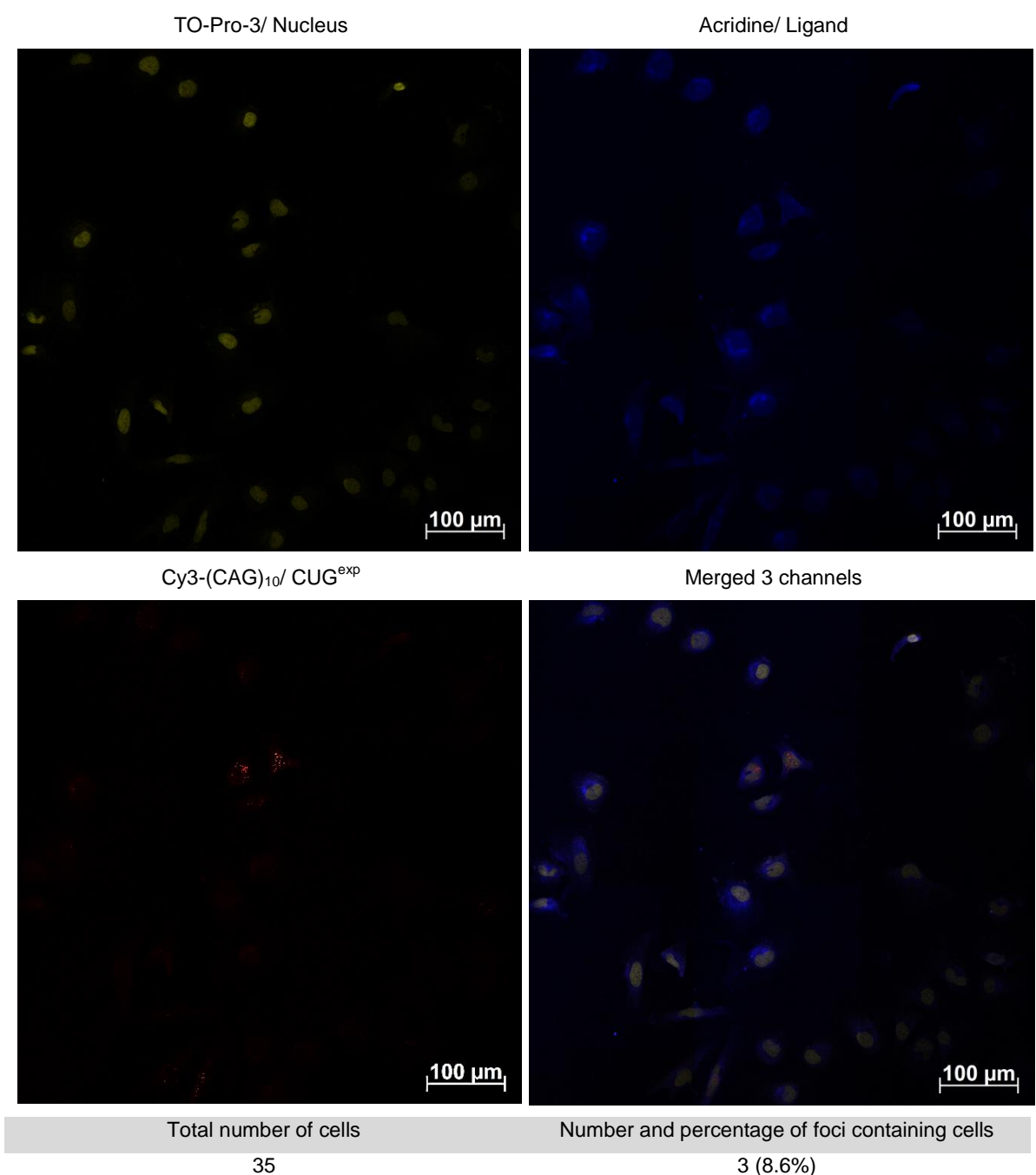
**Figure 3.14.** DM1 cell model treated with 11, 20  $\mu$ M.



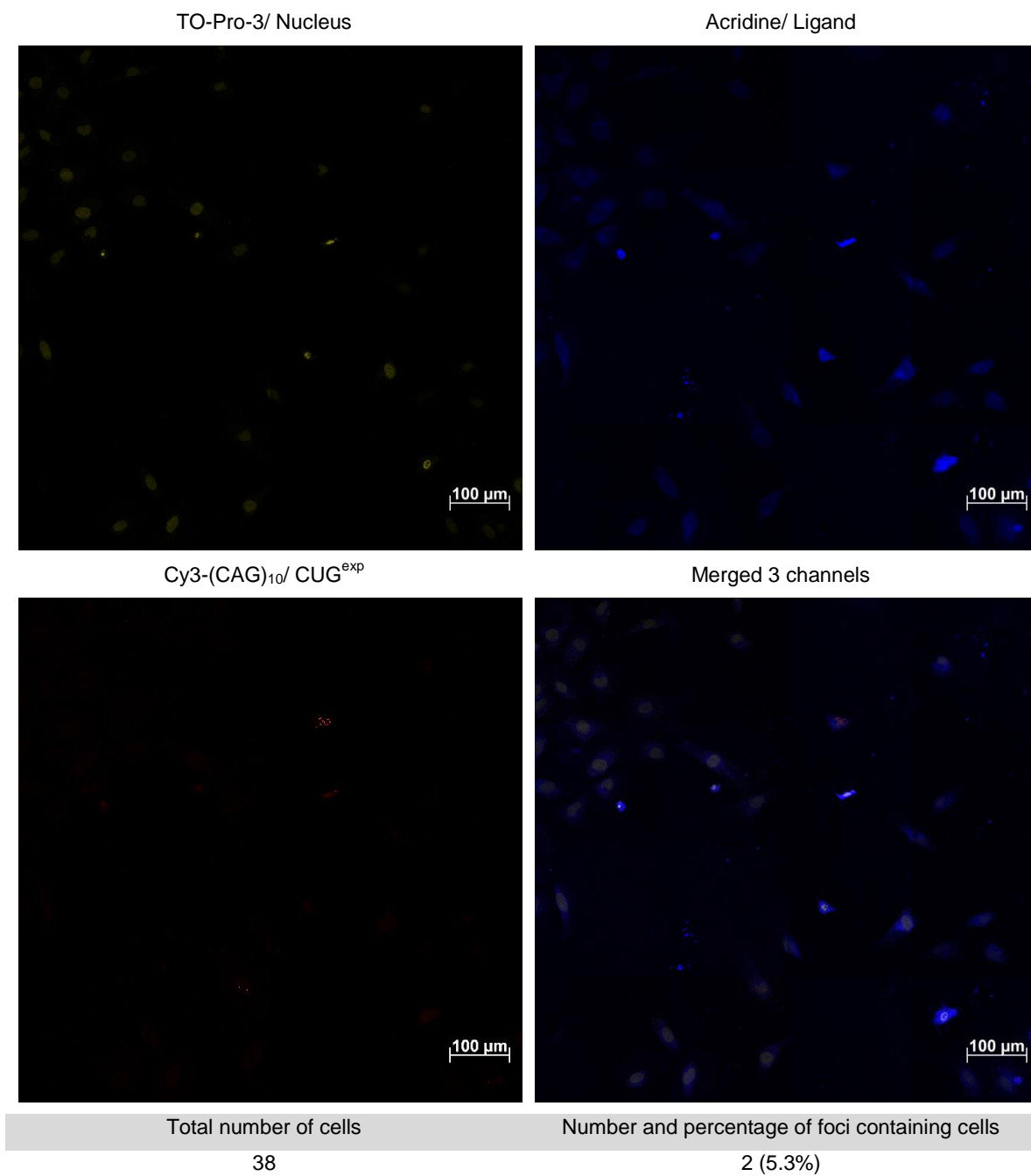
**Figure 3.15.** DM1 cell model treated with 11, 35 μM.



**Figure 3.16.** DM1 cell model treated with **11**, 50 μM.

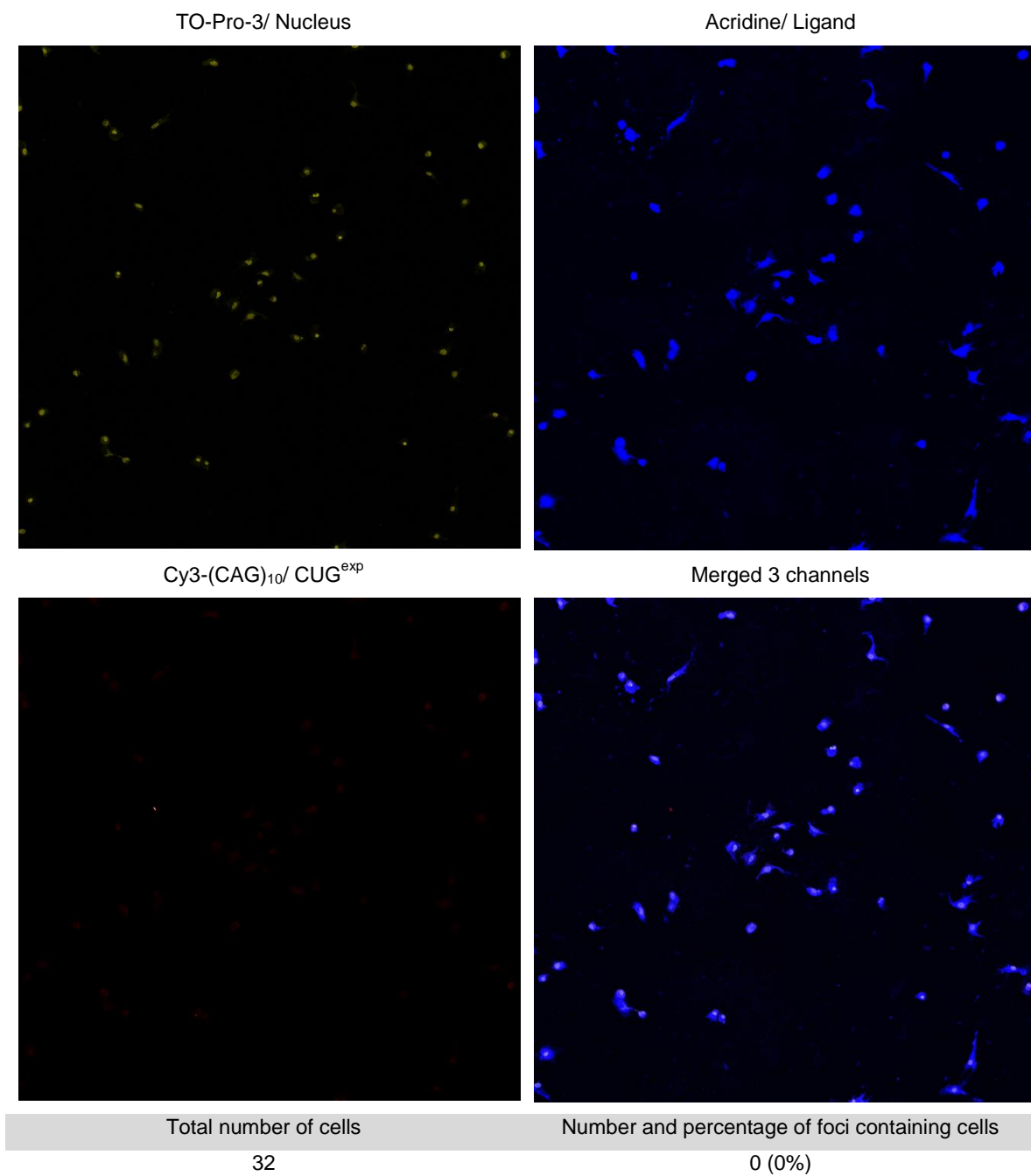


**Figure 3.17.** DM1 cell model treated with **9**, 20 μM.



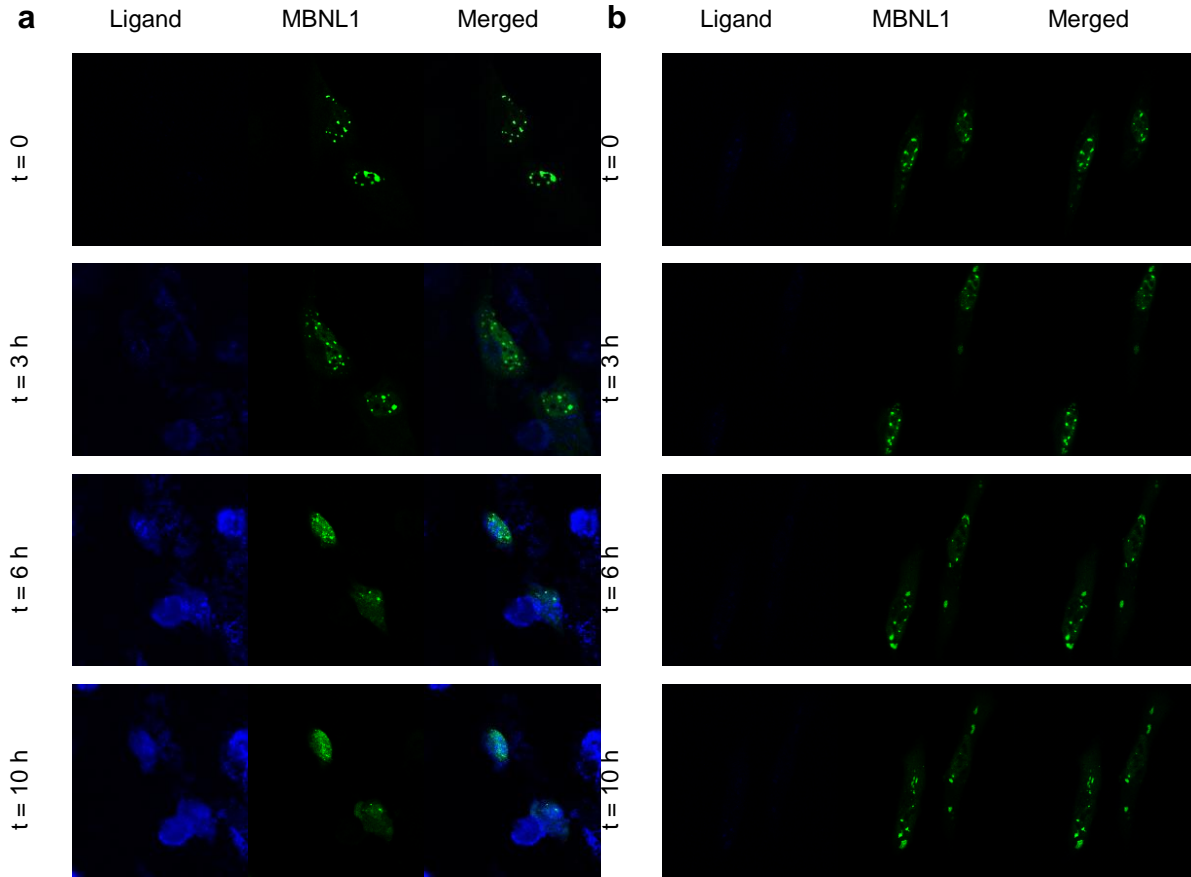
**Figure 3.18.** DM1 cell model treated with **9**, 35  $\mu$ M..



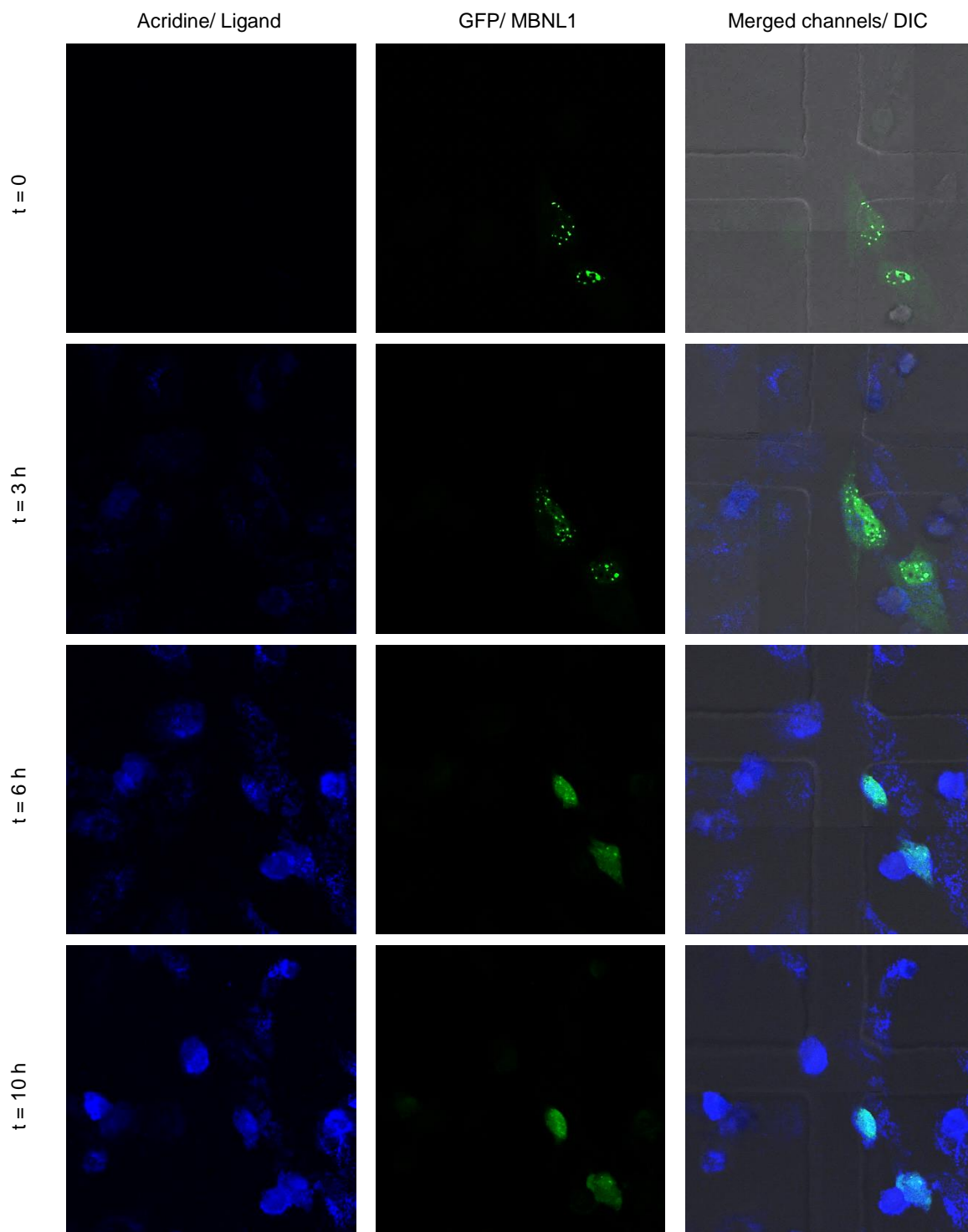


**Figure 3.19.** DM1 cell model treated with **9**, 50  $\mu$ M.

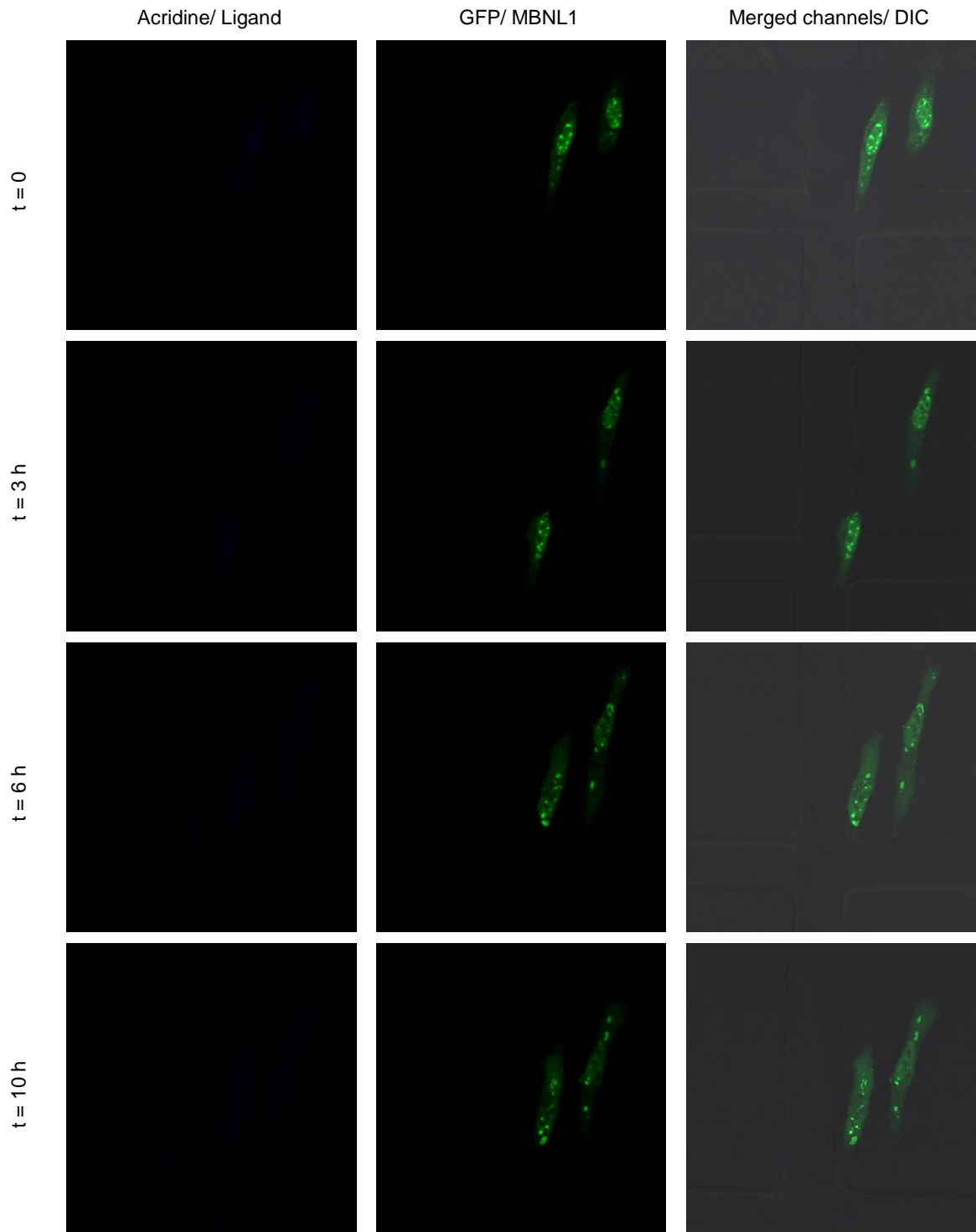
**Ligand 9 is active in MBNL1 foci dispersion in Live DM1 cell microscopy.** FISH experiment was performed on fixed cells which compared ligand-treated cells with untreated ones, indirectly. Because the effect of **9** is an indirect measurement of its activity, we sought to study its activity, directly in a single live DM1 model cell. Live cells were tracked in a Petri dish with an imprinted 500  $\mu\text{m}$  grid, to relocate the cell following the incubation interval, and incubated with **9** at 50  $\mu\text{M}$  at  $t = 0$  and monitored by confocal microscopy at several time points (Figures 3.21 and 3.22). The first observation at time point,  $t = 0$ , was made immediately before addition of **9** and MBNL1 nuclear foci were clearly present (Figure 3.20a,  $t = 0$ ). The ability to monitor the location of **9** was made possible by the inherent fluorescence of the acridine unit. Over time, **9** penetrated the cellular and nuclear membrane and the MBNL1 foci gradually dispersed over the nucleus (Figure 3.20a,  $t = 3, 6$  and 10 h). To validate that foci are dispersed over the entire nucleus a Z-stacked image, containing 1  $\mu\text{m}$ -apart slices from the whole DM1 model cell, was obtained which confirmed almost all the foci are dispersed over the entire nucleus at  $t = 10$  h, upon treatment with **9** at 50  $\mu\text{M}$  (Figure 3.23). As a negative control, to rule out the possibility of spontaneous MBNL1 foci dispersion over time, DM1 model cell without ligand treatment was studied the same way. The stability of foci was confirmed by the presence of foci in all the time points (Figure 3.20b). It is noteworthy that the increase in cell size and number of MBNL1 foci is caused by natural growth of live cells and continuous expression of GFP-MBNL1 in real time. To compare MBNL1 nuclear foci in untreated cells with **9**-treated cells a Z-stacked image, containing 1  $\mu\text{m}$  apart slices, from the whole cell was obtained which confirmed that several number of foci are present at different slices of the untreated cell (Figure 3.24).



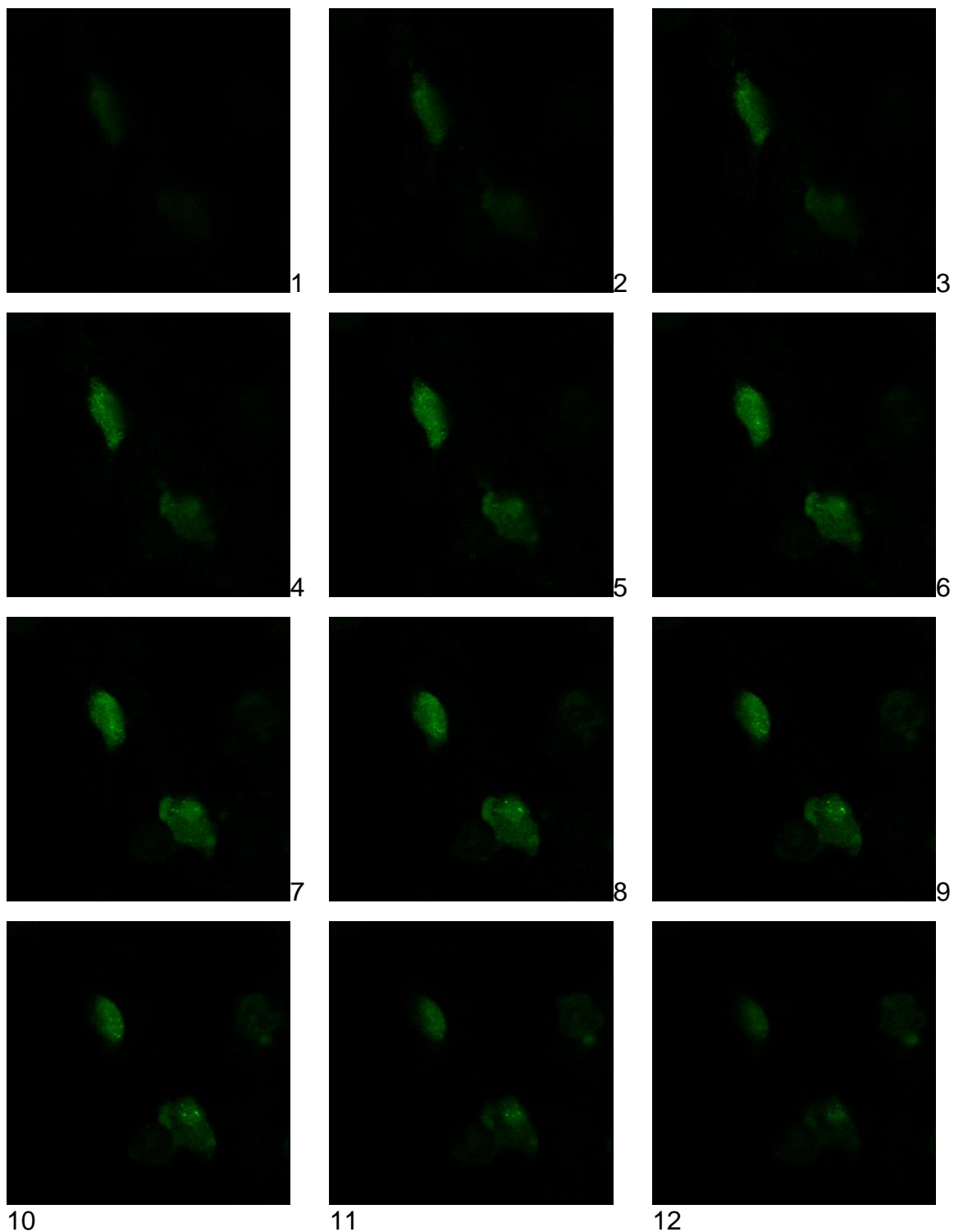
**Figure 3.20.** Live cell microscopy demonstrates a direct evidence for MBNL1 foci dispersion with **9**. a) Live DM1 model cells are treated with **9** (50 μM) at t = 0, immediately after the first image is taken. Fluorescence of **9**, confirms its penetration to the nucleus. MBNL1 nuclear foci are gradually dispersing over time in two cells. b) Two live cell shows stability of foci in a DM1 cell, in the absence of **9**, over the period of 10 h. Each box shows 120 μM X 120 μM.



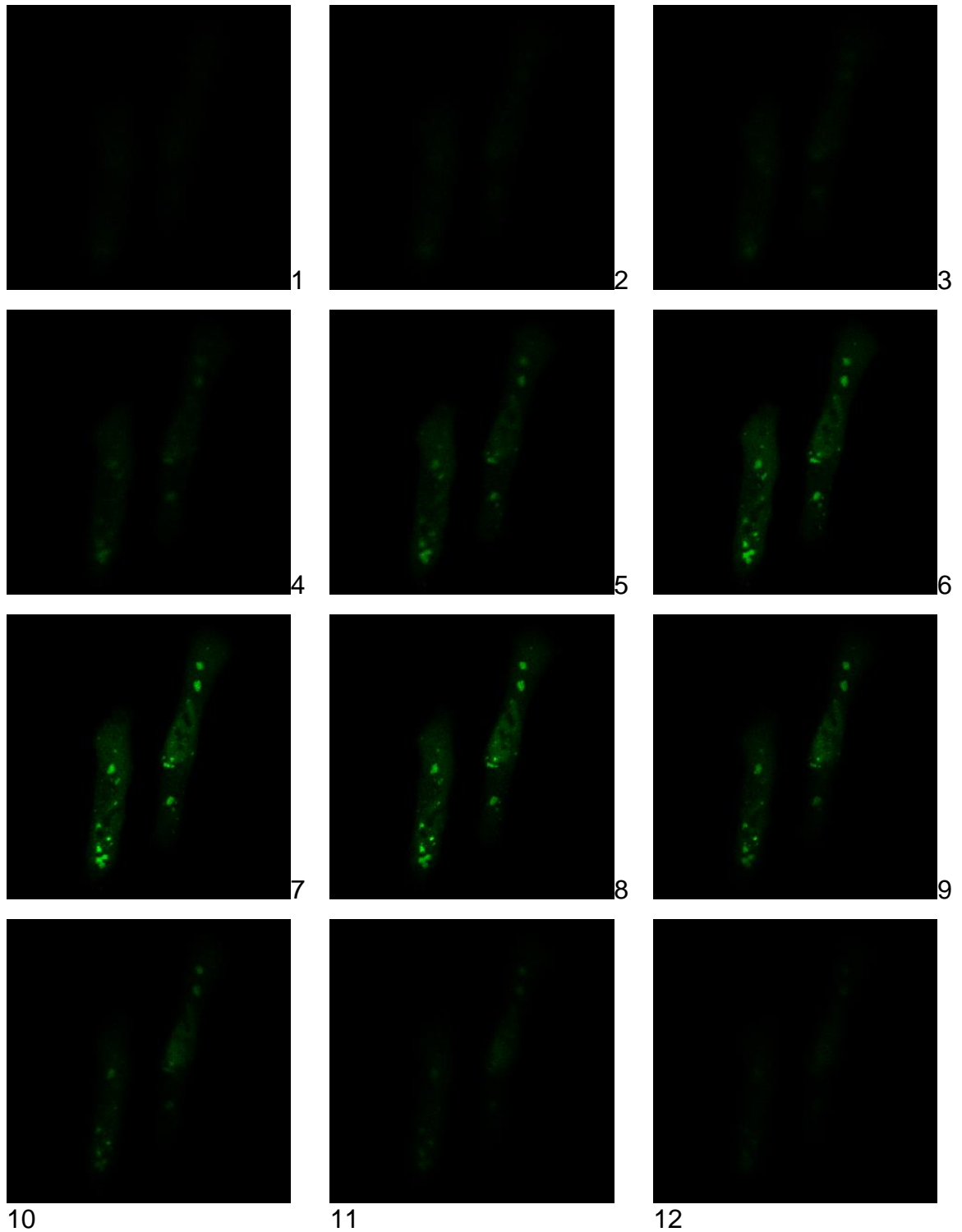
**Figure 3.21.** Merged channels of two DM1 cells treated with **9** (50  $\mu$ M), corresponding to two cells in Figure 3.20a are combined with the DIC channel to show the absolute position of each individual cell. Each box shows 200  $\mu$ M X 200  $\mu$ M.



**Figure 3.22.** Merged channels of two untreated DM1 cells, corresponding to Figure 3.20b are combined with the DIC channel to show the absolute position of each individual cell. Each box shows 200  $\mu$ M X 200  $\mu$ M.



**Figure 3.23.** Z-stacked image of two DM1 cells treated with **9** (50  $\mu$ M) at  $t = 10$  h, corresponding to two cells in the last row of figure 3.20a. It contains 12 Slices, 1  $\mu$ m apart from each other. Almost all MBNL1 foci are dispersed throughout the nuclei at  $t = 10$  h. Each box shows 120  $\mu$ m X 120  $\mu$ m.



**Figure 3.24.** Z-stacked image of two untreated DM1 cells at  $t = 10$  h, corresponding to two cells in the last row of figure 3.20b. It contains 12 Slices,  $1 \mu\text{m}$  apart from each other. Many MBNL1 foci are present in the nuclei at  $t = 10$  h. Each box shows  $120 \mu\text{m} \times 120 \mu\text{m}$ .

### 3.4 Conclusion

To improve the potency and drugability of a previously reported hit molecule we applied a dimeric ligand approach to target CUG repeat. Limited by the absence of structural data on the ligand·CUG<sup>exp</sup> complex, we designed and synthesized a small library of dimeric ligands that could potentially bind to consecutive CUG sites. This small library consisting of 10 dimeric ligands varied in composition, length and attachment point of spacer. These ligands were assessed for their binding potency to (CUG)<sub>12</sub> as well as inhibition potency of (CUG)<sub>12</sub>·MBNL1 interaction. We found the bivalent effect depends more on the composition of the spacer rather than the length of it. Dimeric ligands with oligoamino spacers were the most potent ones, whereas the dimeric ligands containing oligoether spacers, with the same or similar number of atoms, improved the binding and inhibition potency to a much lesser extent. Moreover, short oligoamino spacers made the dimeric ligands aqueous soluble whereas those containing oligoether spacers were aqueous insoluble. Among four dimeric ligands containing oligoamino spacers **9** was the most potent ligand by *in vitro* assays. In cell-based assays **9** was low-cytotoxic (IC<sub>50</sub> > 75 μM) and bioactive at 35 μM. To obtain a direct evidence of its bioactivity, we studied **9** by time-lapse confocal fluorescence microscopy and observed a gradual MBNL1 foci dispersion in individual live DM1 model cells at 50 μM over 10 h. To validate that foci are dispersed over the entire nucleus a Z-stacked image, containing 1 μm-apart slices from the whole DM1 model cell, was obtained which confirmed almost all the foci are dispersed over the entire nucleus at t = 10 h, upon treatment with **9** at 50 μM. The positive results with **9** suggest that it is the optimized dimeric ligand and a good candidate for further lead development.

### 3.5 Materials and Methods

**MBNL1N, CUG<sub>0/960</sub> Plasmid and RNAs.** The expression vector pGEX-6p-1/MBNL1N was obtained from Maurice S. Swanson (University of Florida, College of Medicine). Wild type DMPK-CUG<sub>960</sub>, DMPK-CUG<sub>0</sub> and GFP-MBNL1 mini-genes were obtained from the lab of Thomas Cooper (Baylor College of Medicine). The insulin receptor (IR) mini-gene was obtained from the lab of Nicholas Webster (University of California, San Diego).

The MBNL1 used here is MBNL1N containing the four zinc finger motifs of MBNL1 and a hexa His tag (C-terminus). MBNL1N is known to bind RNA with a similar affinity as



the full-length MBNL1 and is commonly used in such studies.<sup>38</sup> All the oligonucleotides were purchased from Integrated DNA Technology and were HPLC purified. The sequences and modifications for RNA constructs used in this study are as follows:

(CUG)<sub>12</sub> construct for optical melting experiments:

5'-GCCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGGCC-3'

(CUG)<sub>12</sub> construct for SPR experiments:

5'-GCCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGGCC-TEG-biotin-3'

**MBNL1N Protein Expression and Purification.** Using BL21-CodonPlus(DE3)-RP competent cells (Stratagene), the expression of MBNL1N protein was induced with 1 mM IPTG at OD<sub>600</sub> 0.6 in LB media with ampicillin for 2 h at 37 °C. Bacterial cells were collected by centrifugation and were then resuspended in a lysis buffer containing 25 mM Tris-Cl (pH = 8), 0.5 M NaCl, 10 mM imidazole, 2 mM BME, 5% glycerol, 0.1% Triton X-100, 2 mg mL<sup>-1</sup> lysozyme, 0.1 mM PMSF, 1 μM pepstatin, and 1 μM leupeptin, and sonicated six times for 15 s each. The cell pellet was centrifuged, and the supernatant was collected and filtered through a 45 μm Millex filter. To purify MBNL1N, Ni-NTA agarose was incubated with the lysate for 1 h at 4 °C and washed with a washing buffer containing 25 mM Tris-Cl (pH=8), 0.5 M NaCl, 20 mM imidazole and 0.1% Triton X-100, followed by elution with elution buffer of 25 mM Tris-Cl (pH = 8), 0.5 M NaCl, 250 mM imidazole and 0.1% Triton X-100. The eluate containing the GST fusion protein was dialyzed against 1X PBS buffer for using in SPR analysis. The molecular weight was confirmed by MALDI mass spectrometry and the concentration was determined by Bradford assay.

**Optical Melting Experiments.** The melting temperature of the (CUG)<sub>12</sub> was measured on a Shimadzu UV2450 spectrophotometer equipped with a temperature controller. The path length of the cuvettes used was 1 cm. The absorbance of 3.3 μM (CUG)<sub>12</sub> in 1X PBS buffer in the absence and presence of 3.3 μM of each was recorded at 260 nm with a slit width of 1 nm from 10 °C to 95 °C at a ramp rate of 0.5 °C min<sup>-1</sup>. Each profile for melting temperature analysis was generated by subtracting the absorbance of the solution of each ligand in 1X PBS buffer from the (CUG)<sub>12</sub>/ligand solution. Melting temperatures were determined by fitting the melting curve using Meltwin 3.5 software.

**Surface Plasmon Resonance (SPR) Analysis.** All SPR experiments were conducted on a streptavidin coated sensor chip using a Biacore 3000 instrument. Streptavidin coated research grade sensor chips were preconditioned with three consecutive 1-min injections of 1 M NaCl/ 50 mM NaOH before the immobilization was started. 3'-biotin labeled (CUG)<sub>12</sub> was captured on flow cell 2 (Response Unit, RU, between 100-1100). Flow cell 1 was used as a reference. Inhibition analysis was carried out in PBS 1X buffer, pH = 7.4, containing 0.05% Tween-20 and 0.2 mg/mL (7.4 μM or 580 μM nucleotides) bulk yeast t-RNA to confirm the specificity of inhibition. Various concentrations of each ligand were passed over the immobilized RNA at a rate of 20 μL min<sup>-1</sup> for 300 s. After the initial 150 s, a solution of GST-MBNL1 protein, 650 nM, in the same buffer was flowed over the surface for 150 s. The reference-subtracted sensograms were recorded. After the dissociation phase, the surface was regenerated, with a pulse of 0.5% SDS and/or 100 mM NaOH, for a few times followed by a buffer wash to reestablish baseline. The measured RU upon injection of PBS buffer was subtracted from the sensograms. For inhibition studies, the resulting sensograms were set to the baseline at t = 150 s to offset the binding of each ligand to the immobilized (CUG)<sub>12</sub> surface. The peak RU at t = 150 s was recorded and converted to the percentage of (CUG)<sub>12</sub> bound by MBNL1. All values normalized to that measured in the absence of each ligand. The data points were fit to a four parameter logistic curve to determine the apparent IC<sub>50</sub> using the following equation by Kaleidagraph software:

$$Y = \frac{Y_{max} - Y_{min}}{1 + \left(\frac{[ligand]}{IC_{50}}\right)^n} + Y_{min}$$

where Y is the percentage of (CUG)<sub>12</sub> bound by MBNL1,  $Y_{max}$  and  $Y_{min}$  are the maximum and minimum of this percentage and n is the Hill coefficient. Two or three separate SPR experiments on different sensor chips with different levels of (CUG)<sub>12</sub> immobilization were carried out to verify that the values are not affected by surface RNA density.

**Steady State Fluorescence-based Binding Assays.** To determine the equilibrium parameters for binding of ligands to CUG<sup>exp</sup>, we followed quenching of TAMRA in TAMRA-(CUG)<sub>6</sub> at various ligand concentrations. TAMRA-(CUG)<sub>6</sub> was excited at 560 nm and its emission was recorded at 590 nm. Stoichiometric titrations were carried out at 20 °C in PBS, 1X buffer. The baseline fluorescence was recorded before addition of 20 nM TAMRA-(CUG)<sub>6</sub>. Increase in the fluorescence was recorded and attributed to the

fluorescence of TAMRA-(CUG)<sub>6</sub>. Upon addition of each aliquot of each ligand, the fluorescence signal was allowed to equilibrate. Titration was continued at a series of increasing final concentrations of the ligand until the fluorescence was completely quenched. Fluorescence intensities at different concentrations of ligands were fit to the following equation using Kaleidagraph software:

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

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

**FISH (Fluorescence *In Situ* Hybridization).** A total of ca. 120,000 HeLa cells were seeded in each well of a 6-well plate on coverslips. After a day, the cells were transfected with 500 ng DMPK-CUG<sub>0</sub> or DMPK-CUG<sub>960</sub> plasmid using Lipofectamine following the manufacturer's protocol at cell confluence of 70–80%.<sup>39</sup> After 4 hours, the media were changed and each ligand was added to each well at different concentrations (20, 35 and 50  $\mu$ M). After two days, the cells were fixed with 4% PFA then washed five times with 1x PBS. Fixed cells were permeabilized with 0.5% triton X-100 in 1x PBS at room temperature for 5 min. Cells were prewashed with 30% formamide in 2x SSC for 10 min at room temperature. Cells were probed with FISH probe (1 ng  $\mu$ L<sup>-1</sup> of Cy3 CAG<sub>10</sub> in 30% formamide, 2x SSC, 2  $\mu$ g mL<sup>-1</sup> BSA, 66  $\mu$ g mL<sup>-1</sup> yeast tRNA) for 2 h at 37 °C. Cells were then washed with 30% formamide in 2x SSC for 30 min at 37 °C, followed by washing with 1x SSC for 30 min at room temperature. The cells were washed twice with 1x PBS and then nuclei were stained with 1  $\mu$ M To-Pro-3 and washed twice. Cells were mounted onto glass sides with ProLong® Gold. Slides were imaged at RT by LSM 710, AxioObserver confocal microscopy equipment using a confocal single photon technique with a plan-Apochromat 20x/0.8 M27 objective. Image analysis was performed by Axiovision interactive measurement. The following table indicates the excitation filters used in these experiments.

| Fluorophore | Component          | Excitation wavelength (nm) |
|-------------|--------------------|----------------------------|
| Acridine    | Ligands            | 405                        |
| Cy3         | CUG <sub>960</sub> | 555                        |
| TO-PRO-3    | Nucleus            | 639                        |

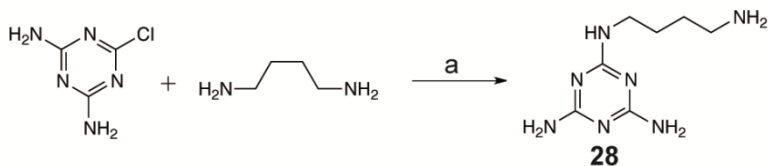
**Live Cell Imaging.** A total of ca. 120,000 HeLa cells were grown in an Ibidi 35 mm Petri dish with a standard bottom, high walls and an imprinted 500  $\mu\text{m}$  relocation grid. After a day, cell confluence reached to about 70–80%, cells were transfected with 500  $\mu\text{g}$  DMPK–CUG<sub>960</sub> plasmid and 500 ng GFP-MBNL1 plasmid using Lipofectamine following standard protocol. After 4 h, media were changed and cells were incubated at 37 °C, 5% CO<sub>2</sub>. 24 h post-transfection, ligand **9** was added to final concentration of 50  $\mu\text{M}$ . Live-cell, time-lapse images were taken before addition of **9** as well as at 3, 6 and 10 h time points at RT by a LSM 710, AxioObserver confocal microscopy equipment using a confocal single photon technique with a plan-Apochromat 20x/0.8 M27 objective. Image analysis was performed by Axiovision interactive measurement. For tracking the cells, DIC images were acquired simultaneously with the reflected light images using a TPMT module after setting the Köhler illumination with a fully opened condenser aperture (0.55 NA).

**Instrumentation and Chemicals.** All reagents were purchased from commercial suppliers and used without further purification. Anhydrous solvents were obtained from an anhydrous solvent dispensing system. For all reactions employing anhydrous solvents, glassware was oven-dried, cooled under vacuum, and then purged and conducted under dry nitrogen. Purified compounds were further dried under high vacuum (0.01–0.05 Torr) or lyophilized using a Labconco lyophilizer. Yields refer to purified and spectroscopically pure compounds. NMR spectra were recorded at 23 °C on either Varian Unity 500 or Varian Unity Inova 500NB, operating at 500 MHz and 125 MHz for <sup>1</sup>H and <sup>13</sup>C acquisitions, respectively. NMR spectra were processed using MestReNova software. Chemical shifts are reported in ppm and referenced to the corresponding residual nuclei in the following deuterated solvents: CDCl<sub>3</sub> (7.26 ppm <sup>1</sup>H, 77.16 ppm <sup>13</sup>C); DMSO (2.50 ppm <sup>1</sup>H, 39.52 ppm <sup>13</sup>C); D<sub>2</sub>O (4.79 ppm <sup>1</sup>H); CD<sub>3</sub>OD (3.31 ppm <sup>1</sup>H, 50.41 ppm <sup>13</sup>C). Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), sext (sextet), dd (doublet of doublets), td (triplet of doublets), dt (doublet of triplets), m (multiplet), b (broad). Integration is provided and coupling constants, *J*, are reported in Hertz (Hz). ESI mass spectra were recorded using the

Quattro or ZMD mass spectrometer. High resolution mass spectra (HRMS) were obtained at the University of Illinois mass spectrometry facility. All compounds described herein gave NMR and mass spectral data in accord with their structures. Ligand **1** gave a HRMS within 1 ppm of calculated values.

### 3.6 Experimental Synthetic Procedures

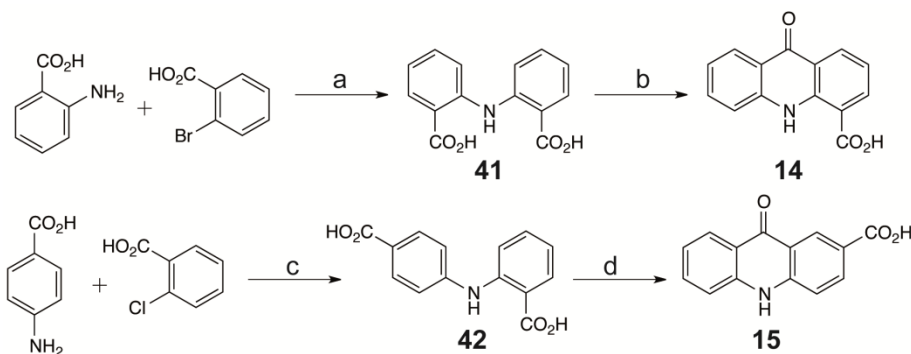
**Scheme 3.1.** Synthesis of *N*<sup>2</sup>-(4-Aminobutyl)-1,3,5-triazine-2,4,6-triamine<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) 130 °C, 80%.

***N*<sup>2</sup>-(4-Aminobutyl)-1,3,5-triazine-2,4,6-triamine (28).** Title compound was synthesized as described previously in 80% yield,<sup>1</sup> with minor changes in the work-up procedure. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 6.40 (t, *J* = 5.6 Hz, 1H), 6.05 (s, 2H), 5.88 (s, 2H), 3.14 (q, *J* = 6.3 Hz, 2H), 2.56 – 2.51 (m, 2H), 1.44 (p, *J* = 7.1 Hz, 2H), 1.37 – 1.31 (m, 2H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 198.1; found 198.1.

**Scheme 3.2.** Synthesis of 9-Oxo-9,10-dihydroacridine-4-carboxylic acid<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) Cu, Cu<sub>2</sub>O, K<sub>2</sub>CO<sub>3</sub>, DMF, Reflux, 24 h, 91%; (b) PPA, 120 °C, 95%; (c) Cu, Cu<sub>2</sub>O, K<sub>2</sub>CO<sub>3</sub>, 2-ethoxyethanol, 130 °C, 70%; (d) PPA, 120 °C, 95%.

**,2'-Azanediylidibenzoic acid (41).** Title compound was synthesized as described previously in 91% yield,<sup>40</sup> with minor changes in the work-up procedure. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 7.91 (dd, *J* = 7.9, 1.3 Hz, 2H), 7.50 – 7.40 (m, 4H), 6.92 – 6.97 (m,

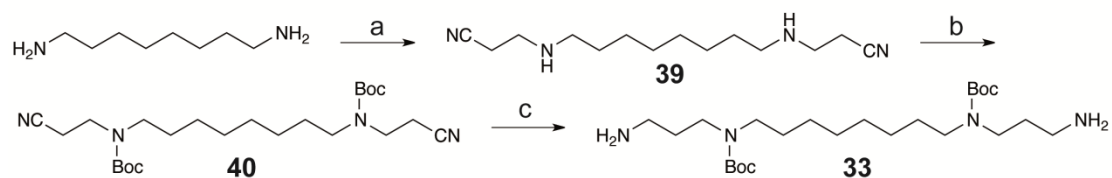
2H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{DMSO-}d_6$ )  $\delta$  168.41, 143.58, 133.38, 131.81, 119.99, 117.56, 113.56;  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 258.1; found 258.1.

**9-Oxo-9,10-dihydroacridine-4-carboxylic acid (14).** Title compound was synthesized as described previously in 95% yield,<sup>40</sup> with minor changes in the work-up procedure.  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-}d_6$ )  $\delta$  12.01 (s, 1H), 8.53 (dd,  $J = 8.0, 1.5$  Hz, 1H), 8.45 (dd,  $J = 7.5, 1.6$  Hz, 1H), 8.24 (d,  $J = 8.0$  Hz, 1H), 7.83 – 7.73 (m, 3H), 7.41 – 7.29 (m, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{DMSO-}d_6$ )  $\delta$  176.53, 169.14, 141.20, 139.92, 136.90, 134.11, 132.41, 125.89, 122.32, 121.63, 120.60, 120.24, 118.63, 115.01;  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 240.1; found 240.1.

**2-((4-Carboxyphenyl)amino)benzoic acid (42).** Title compound was synthesized as described previously in 70% yield,<sup>41</sup> with minor changes in the work-up procedure.  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO-}d_6$ )  $\delta$  9.87 (s, 1H), 7.94 (d,  $J = 8.0$  Hz, 1H), 7.87 (d,  $J = 8.7$  Hz, 2H), 7.51 – 7.45 (m, 2H), 7.27 (d,  $J = 8.7$  Hz, 2H), 6.96 – 6.91 (m, 1H);  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 258.1; found 258.1.

**9-Oxo-9,10-dihydroacridine-2-carboxylic acid (15).** Title compound was synthesized as described previously in 95% yield,<sup>42</sup> with minor changes in the work-up procedure.  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-}d_6$ )  $\delta$  12.90 (s, 1H), 12.06 (s, 1H), 8.83 (d,  $J = 2.0$  Hz, 1H), 8.26 – 8.18 (m, 2H), 7.75 – 7.82 (m, 1H), 7.56 – 7.61 (m, 2H), 7.36 – 7.27 (m, 1H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{DMSO-}d_6$ )  $\delta$  177.00, 167.41, 143.13, 140.90, 134.12, 133.88, 133.62, 128.39, 126.10, 121.76, 120.85, 119.69, 117.69, 117.38;  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 240.1; found 240.1.

**Scheme 3.3.** Synthesis of  $N,N'$ -Bis(3-aminopropyl)-1,8-octanediamine<sup>a</sup>



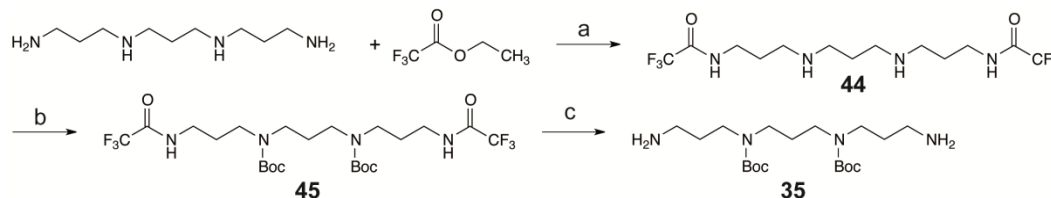
<sup>a</sup>Reagents and conditions: (a) 2 equiv. acrylonitrile,  $\text{EtOH}$ ,  $0\text{ }^\circ\text{C}$  to  $\text{RT}$ , 48 h, 83%; (b)  $\text{Boc}_2\text{O}$ ,  $\text{NEt}_3$ ,  $\text{MeOH}$ ,  $0\text{ }^\circ\text{C}$  to  $\text{RT}$ , 24 h; (c)  $\text{H}_2$ ,  $\text{Ra-Ni}$ ,  $\text{Pd/C}$ ,  $\text{LiOH}$ ,  $\text{dioxane}$ ,  $\text{H}_2\text{O}$ , 24 h, 65%.

**N,N'-Bis(2-cyanoethyl)-1,8-octanediamine (39).** 1,8-octanediamine (1.8 g, 12.5 mmol, 1 equiv.) was dissolved in ethanol. Acrylonitrile (1.33 g, 25 mmol, 2 equiv.) was added dropwise to the solution and stirred for 48 h at room temperature to yield the product in 83% yield.  $^1\text{H}$  NMR (500 MHz, Chloroform-*d*)  $\delta$  2.93 (td,  $J = 6.6, 2.2$  Hz, 4H), 2.66 – 2.59 (m, 4H), 2.52 (td,  $J = 6.6, 2.2$  Hz, 4H), 1.48 (s, 4H), 1.31 (s, 8H);  $^{13}\text{C}$  NMR (125 MHz, Chloroform-*d*)  $\delta$  118.87, 49.36, 45.27, 30.16, 29.54, 27.28, 18.90;  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 251.2; found 251.2.

**1,17-Diazido-3,6,9,12,15-pentaoxaheptadecane (40).** Compound **39** (3.13 g, 12.5 mmol, 1 equiv.) was treated with Di-tert-butyl dicarbonate (10.91 g, 50 mmol, 4 equiv.) and triethylamine (5.06 g, 50 mmol, 4 equiv.) in methanol at room temperature for 24 h to give **40** in 85% yield.  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 333.2; found 333.2.

**N,N'-Bis(3-aminopropyl)-Bis(t-butylcarbamate)-1,8-octanediamine (33).** Compound **40** was hydrogenated in the presence of Pd and Raney nickel and LiOH in a water/ dioxane mixture for 24 h at 100 atm to give the product in 65% yield.  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 281.2; found 281.2.

**Scheme 3.4.** Synthesis of N,N'-Bis(3-aminopropyl)-1,3-propanediamine<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) H<sub>2</sub>O, ACN, 80 °C, 24 h, 77%; (b) Boc<sub>2</sub>O, NEt<sub>3</sub>, 0 °C to RT, 24 h, 85%; (c) 0.6 M NaOH/MeOH, RT, 24 h, 87%.

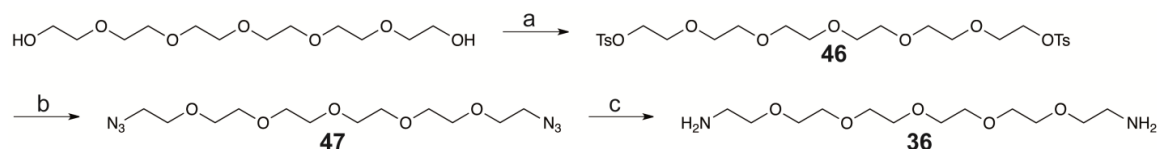
**3,6,9,12,15-Pentaoxaheptadecane-1,17-diyl bis(4-methylbenzenesulfonate) (44).** N,N'-Bis(3-aminopropyl)-1,3-propanediamine was dissolved in acetonitrile. Ethyltrifluoroacetate and water were added dropwise at room temperature and then stirred at 82°C for 24 h to give **44** in 77% yield.<sup>43, 44</sup>  $^1\text{H}$  NMR (500 MHz, Chloroform-*d*)  $\delta$  7.79 (d,  $J = 8.3$  Hz, 4H), 7.34 (d,  $J = 8.1$  Hz, 4H), 4.15 (t,  $J = 5$  Hz, 4H), 3.69 – 3.67 (m, 4H), 3.64 – 3.56 (m, 16H), 2.44 (s, 6H);  $m/z$  LRMS (ESI) calculated for  $[\text{M}+\text{H}]^+$ : 591.2; found 591.2.

**1,17-Diazido-3,6,9,12,15-pentaoxaheptadecane (45).** Compound **44** was treated with NEt<sub>3</sub> and slow addition of Di-tert-butyl dicarbonate at 0 °C. The solution was stirred

at RT for 24 h and worked up as described,<sup>43, 44</sup> to yield the product in 85% yield. <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.69 – 3.65 (m, 20H), 3.42 – 3.36 (m, 4H); <sup>13</sup>C NMR (126 MHz, Chloroform-*d*) δ 70.85, 70.82, 70.78, 70.73, 70.18, 50.83; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 333.2; found 333.2.

**3,6,9,12,15-Pentaoxaheptadecane-1,17-diamine (35).** Compound **45** was stirred in a methanolic sodium hydroxide solution to deprotect the terminal trifluoroacetyl groups and give **35** in 87% yield.<sup>43, 44</sup> <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.67 – 3.60 (m, 20H), 3.50 (t, *J* = 5.2 Hz, 4H), 2.85 (t, *J* = 5.2 Hz, 4H); <sup>13</sup>C NMR (125 MHz, Chloroform-*d*) δ 73.69, 70.77, 70.76, 70.73, 70.45, 42.01; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 281.2; found 281.2.

**Scheme 3.5.** Synthesis of hexaethylene glycol diamine<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) TsCl, CH<sub>2</sub>Cl<sub>2</sub>, DMAP, NEt<sub>3</sub>, 0 °C to RT, 24 h, 77%; (b) NaN<sub>3</sub>, DMF, 80 °C, 24 h, 85%; (c) H<sub>2</sub>, Pd/C, CH<sub>3</sub>OH, 87%.

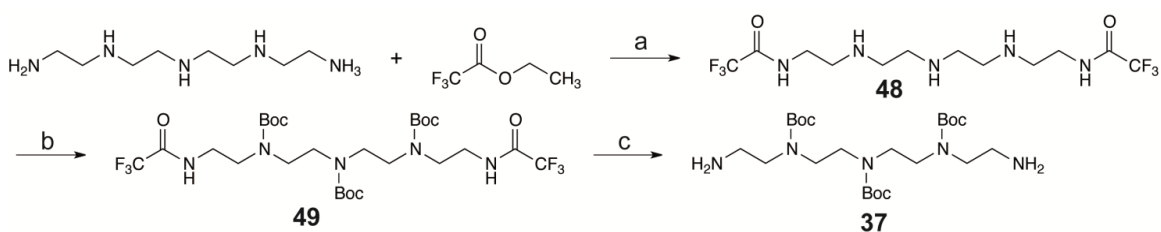
**Hexaethylene glycol 1,20-ditosylate (46).** Hexaethylene glycol (0.705 g, 2.5 mmol, 1 equiv.), DMAP (60 mg, 0.5 mmol, 0.2 equiv.) and NEt<sub>3</sub> (3.5 mL, 24 mmol, 10 equiv.) were dissolved in anhydrous DCM (50 mL) and cooled to 0°C. Tosyl Chloride (1.605 g, 10 mmol, 4 equiv.) was added and stirred for 24 h at room temperature to give **46** in 77% yield.<sup>45</sup> <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 9.61 (s, 2H), 8.75 (s, 2H), 3.26 (q, *J* = 6.5 Hz, 4H), 2.98 (t, *J* = 7.5, 4H), 2.94 – 2.88 (m, 4H), 1.91 (p, *J* = 7.8 Hz, 2H), 1.82 (p, *J* = 7.0 Hz, 4H); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 156.40, 117.08, 44.63, 44.02, 36.61, 25.22, 22.56; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 381.2; found 381.2.

**Hexaethylene glycol 1,20-diazide (47).** Excess sodium azide (663 mg, 10.2 mmol, 10.2 equiv.) was combined with **46** (1 g, 1.7 mmol, 1 equiv.), dissolved in anhydrous DMF (15 mL) and stirred at 80°C for 24 h to yield hexaethylene glycol diazide. The product was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH, 98:2 to 95:5) in 85% yield.<sup>45</sup> <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 9.41 (bs, 2H), 3.18 – 3.08 (m, 12H), 1.72 – 1.62 (m, 6H), 1.40 – 1.35 (m, 18H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 581.3; found 581.3.



**Hexaethylene glycol 1,20-diamine (36).** Compound **47** was dissolved in methanol and hydrogenated in the presence of 5% Pd/C for 24h to yield hexaethylene glycol diamine. The product was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH:NH<sub>4</sub>OH, 80:19:1 to 67:30:2) to give the pure product in 87% yield. <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.28 (bs, 4H), 3.16 (bs, 4H), 2.69 (t, *J* = 6.4 Hz, 4H), 1.81 – 1.70 (m, 2H), 1.64 (p, *J* = 6.8 Hz, 4H), 1.45 (s, 18H); <sup>13</sup>C NMR (126 MHz, Chloroform-*d*) δ 155.29, 79.06, 44.44, 43.68, 38.96, 32.34, 31.73, 28.17; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 389.3; found 389.3.

**Scheme 3.6.** Synthesis of **37**<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) DCM, 0 °C to RT, 24 h (b) Boc<sub>2</sub>O, NEt<sub>3</sub>, Boc<sub>2</sub>O, 0 °C to RT, 48 h; (c) methanol, NaOH, RT, 48 h, 34% overall.

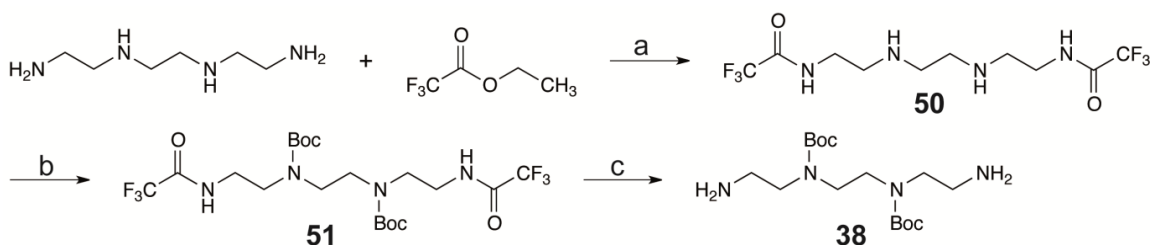
**N,N'-(((azanediylbis(ethane-2,1-diyl))bis(azanediyl))bis(ethane-2,1-diyl))bis(2,2,2-trifluoroacetamide) (48).** Tetraethylenepentamine (5 g, 26.4 mmol, 1 equiv.) was dissolved in DCM (100 mL) and cooled to 0 °C. Ethyl trifluoroacetate (7.5 g, 52.8 mol, 2 equiv.) was added dropwise and the reaction was stirred for 30 min at 0 °C and then for 24 h at RT and worked up as described previously.<sup>46</sup> *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 382.2; found 382.2.

**Di-tert-Butyl (((tert-butoxycarbonyl)azanediyl)bis(ethane-2,1-diyl))bis((2-(2,2,2-trifluoroacetamido)ethyl)carbamate) (49).** Without isolation, TEA (14.6 mL, 105.6 mmol, 4 equiv.) was added to the previous reaction mixture. A solution of Boc<sub>2</sub>O (23 g, 105.6 mmol, 4 equiv.) in DCM (30 mL) was added dropwise and then then the reaction mixture was stirred for 48 h. The mixture was washed with NaHCO<sub>3</sub> and water, dried over sodium sulfate. The solution was concentrated by rotary evaporation and worked up.<sup>46</sup> *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 682.3; found 682.3.

**Di-tert-Butyl(((tert-butoxycarbonyl)azanediyl)bis(ethane-2,1-diyl))bis((2-amino ethyl)carbamate) (37).** Methanol (200 mL) was and NaOH (8 g) was added to **49** and the reaction mixture was stirred for 48 h. The reaction mixture was concentrated by

rotary evaporation, dissolved in chloroform and NaOH, and filtered through celite. The filtrate was washed with water and the organic layers dried over sodium sulfate and concentrated. The solvent was removed by rotary evaporation and the crude mixture was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH:NH<sub>4</sub>OH, 90:10:0 to 90:10:1) to yield the product in 33.8% overall yield. *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 490.4; found 490.4.

**Scheme 3.7.** Synthesis of **38**<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) DCM, methanol, RT, 24 h (b) Boc<sub>2</sub>O, DCM, NEt<sub>3</sub>, 0 °C to RT, 3 h; (c) NaOH, methanol, RT, 24 h, 14% overall.

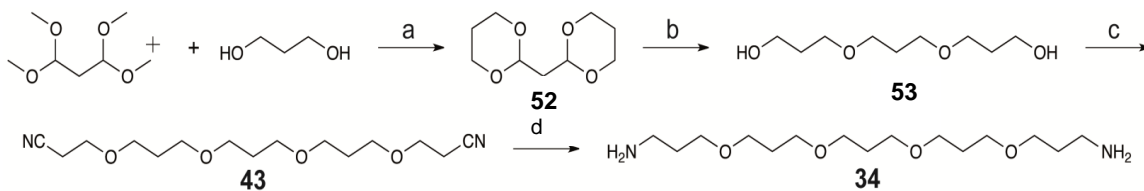
***N,N'*-((Ethane-1,2-diylbis(azanediyl))bis(ethane-2,1-diyl))bis(2,2,2-trifluoroacetamide) (50).** Triethylenepropylamine (5 g, 34.1 mmol, 1 equiv.) was dissolved in DCM (20 mL) and methanol (10 mL). A solution of Ethyl trifluoroacetate (10.2 g, 71.8 mmol, 2.1 equiv.) in DCM (50 mL) as added dropwise over 1 h at 0 °C. This was stirred at RT for 3 h and worked up as described previously;<sup>47</sup> *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 339.1; found 339.1.

**Di-tert-butyl ethane-1,2-diylbis((2-(2,2,2-trifluoroacetamido)ethyl)carbamate) (51).** Without purification, a solution of Boc<sub>2</sub>O (25 g, 114.5 mmol, 3.36 equiv.) in DCM (20 mL) was added dropwise over a half hour. Triethylamine (16.1 mL, 114.5 mmol, 3.36 equiv.) was added and the reaction was stirred overnight. The solvent was then evaporated to 150 mL, washed with saturated NaHCO<sub>3</sub>, sodium citrate (5%), and water. The organic layer was dried over sodium sulfate, filtered, and concentrated. *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 539.2; found 529.2.

**Di-tert-butyl ethane-1,2-diylbis((2-aminoethyl)carbamate) (38).** Without purification, **51** was dissolved in methanol. NaOH (1 M) was added to the the solution and the reaction was stirred overnight. The reaction mixture was concentrated by rotary evaporation, dissolved in chloroform, dried over sodium sulfate, filtered through celite, and concentrated. The crude mixture was purified via flash chromatography (SiO<sub>2</sub>;

CH<sub>2</sub>Cl<sub>2</sub>:MeOH 95:5) to yield the pure product in 35.5% yield overall. *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 347.3; found 347.3.

**Scheme 3.8.** Synthesis of pentapropylene glycol diamine<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) TsOH (cat.), toluene, 110 °C, 24 h, 96%; (b) BH<sub>3</sub>-THF, 90 °C, 4 days, 61%; (c) acrylonitrile, NaH, 15-crown-5, 0 °C, 30 min, 94%; (d) BH<sub>3</sub>-THF, 90 °C 24 h, 90%.

**General synthetic procedure for compounds 18-27.** A round-bottom flask, equipped with a stir bar, was charged with **14** or **15** (1 equiv.) and freshly distilled thionyl chloride (16.4 equiv.). A catalytic amount of DMF was added and heated gently under reflux at 70 °C, stirring until homogeneous and then for 1 h. The excess thionyl chloride was distilled off and the last traces of it were removed azeotropically via co-evaporation with DCM (3 x 50 mL). It was left under vacuum (minimally) for 1 h to afford the crude intermediate as a yellow powder. The crude intermediate was dissolved in anhydrous DCM. Anhydrous triethylamine was added to the solution until the pH was 11 and it was cooled to 0 °C. Corresponding diamines **33-43** (0.45 equiv.) was added and the solution was stirred at 0 °C for 2 hours and slowly warmed to room temperature overnight. The solvent was removed by rotary evaporation and the crude mixture was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH, generally from 98:2 to 95:5) to yield **18-27** as a yellow solid.

**General synthetic procedure for compounds 4, 6-8, 12,13, 29-32.** A round-bottom flask, equipped with a stir bar, was charged with one of compounds **18-27** (1 equiv.) and *N*-(4-Aminobutyl)-[1,3,5]triazine-2,4,6-triamine (1.1 equiv.). DIPEA (2.2 equiv.) and anhydrous DMF (25 mL) were added. The solution was heated at 70 °C for 5 hours. The solvent was removed by rotary evaporation and the product purified via flash chromatography (Basic Alumina; DCM:Methanol:NH<sub>4</sub>OH, generally from 95:4.9:0.1 to 90:9.5:0.5) to yield the corresponding compound, **4, 6-8, 12,13,29-32** as a yellow solid.

**General synthetic procedure for compounds 5, 9-11.** A round-bottom flask, equipped with a stir bar, was charged with one of compounds **18-27** (1 equiv.). TFA (30

mL) and anhydrous DCM (70 mL) were added and stirred at room temperature for 2 h. The solvents were removed to yield compounds **5**, **9-11** as a yellow solid in 100% yield.

**9-Chloro-N-methylacridine-2-carboxamide (16).** <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ 8.79 (d, *J* = 1.6 Hz, 1H), 8.40 (d, *J* = 8.6 Hz, 1H), 8.24 – 8.19 (m, 2H), 8.14 (dd, *J* = 9.0, 1.8 Hz, 1H), 7.86 – 7.83 (m, 1H), 7.69 – 7.62 (m, 1H), 6.56 (bs, 1H), 3.13 (d, *J* = 4.8 Hz, 3H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 271.1; found 271.1; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 271.1; found 271.1. Yield = 75%

**9-((4-((4,6-Diamino-1,3,5-triazin-2-yl)amino)butyl)amino)-N-methylacridine-2-carboxamide (2).** <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ 8.91 (s, 1H), 8.33 (d, *J* = 8.6 Hz, 1H), 8.02 (dd, *J* = 9.1, 1.9 Hz, 1H), 7.85 (t, *J* = 7.6 Hz, 2H), 7.75 – 7.68 (m, 1H), 7.42 – 7.35 (m, 1H), 4.03 (t, *J* = 7.2 Hz, 2H), 3.37 – 3.33 (m, 2H), 3.00 (s, 3H), 1.93 (p, *J* = 7.4 Hz, 2H), 1.70 (p, *J* = 7.2 Hz, 2H); *m/z* HRMS (ESI) calculated for [M+H]<sup>+</sup>: 432.2260; found 432.2267; Yield = 70%

**9-Chloro-N-methylacridine-4-carboxamide (17).** <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 11.33 (s, 1H), 8.90 (d, *J* = 6.9 Hz, 1H), 8.41 (d, *J* = 8.3 Hz, 1H), 8.26 (d, *J* = 8.3 Hz, 1H), 8.02 – 7.95 (m, 1H), 7.76 (t, *J* = 7.4 Hz, 1H), 7.65 – 7.55 (m, 2H), 3.19 (d, *J* = 4.7 Hz, 3H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 271.1; found 271.1. *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 271.1; found 271.1; Yield = 78%

**9-((4-((4,6-Diamino-1,3,5-triazin-2-yl)amino)butyl)amino)-N-methylacridine-4-carboxamide (3).** <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.54 (d, *J* = 8.2 Hz, 2H), 8.41 (d, *J* = 8.3 Hz, 1H), 7.99 (d, *J* = 8.2 Hz, 1H), 7.79 – 7.73 (m, 1H), 7.42 (q, *J* = 8.8 Hz, 2H), 6.54 (t, *J* = 5.5 Hz, 1H), 6.14 (s, 2H), 6.00 (s, 2H), 3.89 (t, *J* = 6.5 Hz, 2H), 3.17 (q, *J* = 6.3 Hz, 2H), 3.01 (d, *J* = 4.2 Hz, 3H), 1.79 (p, *J* = 7.6 Hz, 2H), 1.52 (p, *J* = 7.2 Hz, 2H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 432.2; found 432.2. *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 432.2; found 432.2. Yield = 80%

**N,N-(Piperazine-1,4-diylbis(propane-3,1-diyl))bis(9-chloroacridine-2-carboxamide) (18).** *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 679.2 ; found 679.2. Yield= 65%

***N,N'*-(Piperazine-1,4-diylbis(propane-3,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-2-carboxamide) (4).** *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1001.6 ; found 1001.6. Yield = 61%

**Di-*tert*-butyloctane-1,8-diylbis((3-(9-chloroacridine-2-carboxamido)propyl)carbamate) (19).** <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 9.08 (s, 2H), 8.52 – 8.42 (m, 4H), 8.35 (s, 4H), 7.87 (s, 2H), 7.68 (s, 2H), 3.53 (s, 4H), 3.44 (s, 4H), 3.18 (s, 4H), 1.81 (s, 4H), 1.56 (s, 4H), 1.52 (s, 18H), 1.32 (s, 8H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 937.4; found 937.4. Yield = 65%

**Di-*tert*-butyloctane-1,8-diylbis((3-(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-2-carboxamido)propyl)carbamate) (29).** <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.90 (s, 2H), 8.29 (d, *J* = 8.6 Hz, 2H), 8.04 (d, *J* = 8.6 Hz, 2H), 7.84 (s, 4H), 7.73 – 7.66 (m, 2H), 7.40 – 7.33 (m, 2H), 4.01 (t, *J* = 7.0 Hz, 4H), 3.45 (t, *J* = 6.7 Hz, 4H), 3.35 (s, 8H), 3.17 (s, 4H), 1.94 – 1.84 (m, 8H), 1.68 (p, *J* = 7.0 Hz, 4H), 1.49 (s, 4H), 1.43 (s, 18H), 1.26 (s, 8H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1259.7; found 1259.7. Yield = 61%

***N,N'*-(Octane-1,8-diylbis(azanediyl))bis(propane-3,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-2-carboxamide) (5).** <sup>1</sup>H NMR (500 MHz, Deuterium Oxide) δ 8.62 (s, 2H), 8.10 – 8.07 (m, 4H), 7.90 – 7.85 (m, 2H), 7.51 (s, 2H), 7.50 – 7.43 (m, 4H), 3.98 (t, *J* = 6.2 Hz, 4H), 3.56 (t, *J* = 6.6 Hz, 4H), 3.19 (t, *J* = 6.1 Hz, 4H), 3.15 – 3.11 (m, 4H), 3.07 – 3.03 (m, 4H), 2.05 (q, *J* = 6.9 Hz, 4H), 1.91 – 1.85 (m, 4H), 1.73 – 1.68 (m, 4H), 1.66 – 1.61 (m, 4H), 1.41 – 1.34 (m, 8H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1059.6; found 1059.6. Yield = 100%

***N,N'*-(4,8,12,16-Tetraoxanonadecane-1,19-diyl)bis(9-chloroacridine-4-carboxamide) (20).** <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 9.00 (d, *J* = 6.9 Hz, 2H), 8.58 (dd, *J* = 8.7, 1.4 Hz, 2H), 8.41 (d, *J* = 8.7 Hz, 2H), 8.14 (d, *J* = 8.7 Hz, 2H), 7.89 – 7.84 (m, 2H), 7.76 – 7.71 (m, 2H), 7.71 – 7.65 (m, 2H), 3.76 (q, *J* = 6.8 Hz, 4H), 3.66 (t, *J* = 6.3 Hz, 4H), 3.55 (t, *J* = 6.4 Hz, 4H), 3.46 (t, *J* = 6.4 Hz, 4H), 3.41 (t, *J* = 6.4 Hz, 4H), 2.08 (p, *J* = 6.5 Hz, 4H), 1.83 (p, *J* = 6.4 Hz, 4H), 1.76 (p, *J* = 6.4 Hz, 2H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 785.3; found 785.3. Yield = 62%

***N,N'*-(4,8,12,16-Tetraoxanonadecane-1,19-diyl)bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide) (6).** <sup>1</sup>H NMR (500 MHz,

DMSO- $d_6$ )  $\delta$  8.59 (s, 2H), 8.52 (d,  $J = 7.6$  Hz, 2H), 8.39 (d,  $J = 7.2$  Hz, 2H), 7.87 (s, 2H), 7.71 (s, 2H), 7.55 (s, 2H), 7.38 (p,  $J = 7.6$  Hz, 4H), 6.46 (t,  $J = 5.7$  Hz, 2H), 6.08 (s, 4H), 5.94 (s, 4H), 3.85 (t,  $J = 6.3$  Hz, 4H), 3.53 (t,  $J = 5.7$  Hz, 8H), 2.46 – 3.38 (m, 3H), 3.31 (t,  $J = 6.4$  Hz, 2H), 3.26 (t,  $J = 6.4$  Hz, 2H), 3.16 (q,  $J = 6.4$  Hz, 3H), 1.89 (p,  $J = 6.2$  Hz, 4H), 1.81 – 1.73 (m, 4H), 1.67 (p,  $J = 6.4$  Hz, 4H), 1.58 (p,  $J = 6.4$  Hz, 2H), 1.55 – 1.46 (m, 4H);  $m/z$  LRMS (ESI) calculated for  $[M+H]^+$ : 1007.6; found 1007.6. Yield = 60%

***N,N'*-((Ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(9-chloroacridine-2-carboxamide) (21).**  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  9.03 – 8.96 (m, 2H), 8.84 (s, 2H), 8.34 (d,  $J = 8.7$  Hz, 2H), 8.24 (dd,  $J = 9.1, 1.8$  Hz, 2H), 8.20 – 8.12 (m, 4H), 7.95 – 7.90 (m, 2H), 7.79 – 7.72 (m, 2H), 3.65 (d,  $J = 8.6$  Hz, 8H), 3.51 (q,  $J = 5.3$  Hz, 4H);  $^{13}\text{C}$  NMR (125 MHz, DMSO- $d_6$ )  $\delta$  176.86, 165.62, 142.45, 140.77, 133.70, 132.02, 126.79, 126.03, 125.70, 121.55, 120.72, 119.54, 117.51, 117.27, 69.57, 68.94;  $m/z$  LRMS (ESI) calculated for  $[M+H]^+$ : 627.2; found 627.2. Yield = 66%

***N,N'*-((Ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-2-carboxamide) (7).**  $^1\text{H}$  NMR (400 MHz, Methanol- $d_4$ )  $\delta$  8.73 (s, 2H), 8.11 (d,  $J = 8.6$  Hz, 2H), 7.91 (dd,  $J = 9.1, 1.6$  Hz, 2H), 7.69 (dd,  $J = 21.1, 8.7$  Hz, 4H), 7.63 – 7.56 (m, 2H), 7.29 – 7.20 (m, 2H), 3.82 (t,  $J = 7.1$  Hz, 4H), 3.74 (d,  $J = 5.5$  Hz, 8H), 3.64 (t,  $J = 5.2$  Hz, 4H), 3.28 (d,  $J = 6.7$  Hz, 4H), 1.81 (p,  $J = 7.5$  Hz, 4H), 1.66 – 1.59 (m, 4H);  $m/z$  LRMS (ESI) calculated for  $[M+H]^+$ : 949.5; found 949.5. Yield = 61%

***N,N'*-(3,6,9,12,15-Pentaoxaheptadecane-1,17-diyl)bis(9-chloroacridine-4-carboxamide) (22).**  $^1\text{H}$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  8.88 (dd,  $J = 9.5, 1.4$  Hz, 2H), 8.41 (dd,  $J = 9.7, 1.6$ , 2H), 8.22 (d,  $J = 8.7$  Hz, 2H), 8.05 (d,  $J = 8.6$  Hz, 2H), 7.76 (t,  $J = 9.7$  Hz, 2H), 7.62 (t,  $J = 7.1$  Hz, 2H), 7.55 (t,  $J = 7.5$  Hz, 2H), 3.85 (d,  $J = 4.9$  Hz, 4H), 3.83 – 3.79 (m, 4H), 3.76 – 3.73 (m, 4H), 3.72 – 3.68 (m, 4H), 3.62 – 3.58 (m, 4H), 3.55 – 3.51 (m, 4H);  $^{13}\text{C}$  NMR (125 MHz, Chloroform- $d$ )  $\delta$  165.36, 146.99, 146.02, 142.74, 135.50, 131.49, 129.35, 128.64, 128.42, 127.47, 126.46, 124.51, 124.26, 123.57, 70.80, 70.68, 70.58, 70.42, 70.22, 39.76;  $m/z$  LRMS (ESI) calculated for  $[M+H]^+$ : 759.2; found 759.2. Yield = 70%

***N,N'*-(3,6,9,12,15-Pentaoxaheptadecane-1,17-diyl)bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide) (8).**  $^1\text{H}$  NMR (500 MHz,

DMSO-*d*<sub>6</sub>) δ 8.60 (dd, *J* = 7.1, 1.1 Hz, 2H), 8.56 – 8.51 (m, 2H), 8.39 (d, *J* = 9.1 Hz, 2H), 7.92 (d, *J* = 8.6 Hz, 2H), 7.75 – 7.68 (m, 2H), 7.53 (t, *J* = 5.5 Hz, 2H), 7.41 (dd, *J* = 8.6, 7.3 Hz, 2H), 7.39 – 7.35 (m, 2H), 6.42 (t, *J* = 5.7 Hz, 2H), 6.02 (bs, 4H), 5.88 (bs, 4H), 3.85 (q, *J* = 6.5 Hz, 4H), 3.69 – 3.64 (m, 4H), 3.64 – 3.61 (m, 8H), 3.58 – 3.55 (m, 4H), 3.47 – 3.43 (m, 4H), 3.39 – 3.35 (m, 3H), 3.14 (q, *J* = 6.5 Hz, 4H), 1.76 (p, *J* = 7.5 Hz, 4H), 1.49 (p, *J* = 6.9 Hz, 4H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1081.6; found 1081.6. Yield = 65%

**Di-*tert*-butylpropane-1,3-diylbis((3-(9-chloroacridine-4-carboxamido)propyl) carbamate) (23).** <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 9.00 (s, 2H), 8.36 (d, *J* = 14.0 Hz, 4H), 8.27 (s, 2H), 8.20 (d, *J* = 16.2 Hz, 4H), 7.80 (t, *J* = 7.4, 2H), 7.62 (t, *J* = 6.2, 2H), 3.54 (s, 4H), 3.45 (s, 4H), 3.23 (s, 4H), 1.89 – 1.81 (m, 6H), 1.52 (s, 18H); <sup>13</sup>C NMR (125 MHz, Chloroform-*d*) δ 149.56, 149.23, 132.76, 131.12, 130.12, 129.75, 128.54, 127.13, 124.59, 124.33, 123.26, 44.94, 43.54, 36.14, 28.52, 27.87, 27.63; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 867.3; found 867.3. Yield = 60%

**Di-*tert*-butylpropane-1,3-diylbis((3-(9-((4-(4,6-diamino-1,3,5-triazin-2-yl)amino) butyl)amino)acridine-4-carboxamido)propyl)carbamate) (30).** <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.60 (s, 2H), 8.37 (d, *J* = 8.5 Hz, 2H), 8.20 (d, *J* = 8.5 Hz, 2H), 7.86 (d, *J* = 54.6 Hz, 2H), 7.63 (t, *J* = 7.1 Hz, 2H), 7.38 – 7.32 (m, 2H), 7.32 – 7.27 (m, 2H), 3.82 (s, 4H), 3.49 (t, *J* = 6.2 Hz, 4H), 3.35 (s, 4H), 3.27 (t, *J* = 6.8 Hz, 4H), 3.19 (s, 4H), 1.88 (s, 1H), 1.84 – 1.75 (m, 4H), 1.59 (p, *J* = 7.1, 6.7 Hz, 17H), 1.34 (s, 4H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1189.7; found 1189.7. Yield = 55%

***N,N*-((Propane-1,3-diylbis(azanediyl))bis(propane-3,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide) (9).** <sup>1</sup>H NMR (500 MHz, Deuterium Oxide) δ 8.34 (d, *J* = 8.3 Hz, 1H), 8.17 (d, *J* = 8.3 Hz, 1H), 8.11 (d, *J* = 7.4 Hz, 1H), 7.78 (t, *J* = 7.7 Hz, 1H), 7.52 (d, *J* = 8.3 Hz, 1H), 7.47 – 7.38 (m, 2H), 4.07 (t, *J* = 6.8 Hz, 2H), 3.58 (t, *J* = 6.7 Hz, 2H), 3.36 – 3.31 (m, 2H), 3.26 (t, *J* = 5.9 Hz, 6H), 2.31 (p, *J* = 7.9, 7.5 Hz, 1H), 2.14 – 2.08 (m, 2H), 1.94 (p, *J* = 7.2 Hz, 3H), 1.68 (q, *J* = 6.5, 5.7 Hz, 3H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 989.6; found 989.6. Yield = 100%

**Di-*tert*-butyl(((*tert*-butoxycarbonyl)azanediyl)bis(ethane-2,1-diyl))bis((2-(9-chloroacridine-4-carboxamido)ethyl)carbamate) (24).** *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 968.4; found 968.4. Yield = 60%

**Di-*tert*-butyl (((*tert*-butoxycarbonyl)azanediyl)bis(ethane-2,1-diyl))bis((2-(9-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)ethyl)carbamate) (31).** *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1290.7; found 1290.7. Yield = 58%

***N,N*-(((Azanediyl)bis(ethane-2,1-diyl))bis(azanediyl))bis(ethane-2,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide) (10).** <sup>1</sup>H NMR (500 MHz, Deuterium Oxide) δ 8.23 (d, *J* = 8.5 Hz, 2H), 8.09 (d, *J* = 7.4 Hz, 2H), 7.99 (d, *J* = 8.6 Hz, 2H), 7.65 (t, *J* = 7.7 Hz, 2H), 7.43 (d, *J* = 8.4 Hz, 2H), 7.36 (t, *J* = 7.9 Hz, 2H), 7.28 (s, 2H), 3.88 (t, *J* = 6.7 Hz, 3H), 3.76 (t, *J* = 5.4 Hz, 3H), 3.52 (s, 4H), 3.37 (t, *J* = 5.5 Hz, 3H), 3.13 (t, *J* = 6.1 Hz, 3H), 1.79 (p, *J* = 6.9 Hz, 3H), 1.56 – 1.50 (m, 3H). *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 990.6; found 990.6. Yield = 100%

**Di-*tert*-butylethane-1,2-diylbis((2-(9-chloroacridine-4-carboxamido)ethyl)carbamate) (25).** *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 825.3; found 825.3. Yield = 62%

**Di-*tert*-butylethane-1,2-diylbis((2-(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)ethyl)carbamate) (32).** *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1147.6; found 1147.6. Yield = 57%

***N,N*-((Ethane-1,2-diyl)bis(azanediyl))bis(ethane-2,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide) (11).** <sup>1</sup>H NMR (500 MHz, Deuterium Oxide) δ 8.18 (d, *J* = 8.6 Hz, 2H), 8.07 (d, *J* = 7.2 Hz, 2H), 8.05 – 8.01 (m, 2H), 7.67 (t, *J* = 7.5 Hz, 2H), 7.38 (d, *J* = 8.4 Hz, 2H), 7.34 (s, 2H), 7.29 (t, *J* = 8.0 Hz, 2H), 3.95 (s, 4H), 3.83 (s, 4H), 3.73 (s, 4H), 3.50 (s, 4H), 3.21 (s, 4H), 1.86 (s, 4H), 1.61 (s, 4H); <sup>13</sup>C NMR (125 MHz, Deuterium Oxide) δ 170.11, 163.32, 163.03, 162.75, 162.47, 159.99, 159.26, 157.58, 156.52, 119.95, 118.19, 117.63, 115.30, 112.98, 48.85, 48.65, 43.81, 40.32, 36.56, 26.06, 24.93; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 947.5; found 947.5. Yield = 100%



***N,N'*-(((Oxybis(ethane-2,1-diyl))bis(oxy))bis(propane-3,1-diyl))bis(9-chloroacridine-2-carboxamide) (26).** <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 8.76 (s, 2H), 8.24 (d, *J* = 8.6 Hz, 2H), 8.15 – 8.09 (m, 4H), 8.07 (d, *J* = 9.1 Hz, 2H), 7.77 – 7.69 (m, 4H), 7.54 (t, *J* = 7.6 Hz, 2H), 3.65 – 3.56 (m, 12H), 3.52 – 3.48 (m, 4H), 1.89 (p, *J* = 5.7 Hz, 4H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 699.2; found 699.2. Yield = 70%

***N,N'*-(((Oxybis(ethane-2,1-diyl))bis(oxy))bis(propane-3,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-2-carboxamide) (12).** <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.75 (s, 2H), 8.15 (d, *J* = 8.6 Hz, 2H), 7.97 – 7.91 (m, 2H), 7.79 – 7.70 (m, 4H), 7.65 – 7.58 (m, 2H), 7.31 – 7.24 (m, 2H), 3.86 (t, *J* = 7.1 Hz, 4H), 3.66 – 3.62 (m, 4H), 3.59 – 3.55 (m, 8H), 3.52 (t, *J* = 6.8 Hz, 4H), 3.32 – 3.27 (m, 4H), 1.91 – 1.86 (m, 4H), 1.86 – 1.80 (m, 4H), 1.64 (p, *J* = 6.9 Hz, 4H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1021.5; found 1021.5. Yield = 66%

***N,N'*-(3,6,9,12,15-Pentaoxaheptadecane-1,17-diyl)bis(9-chloroacridine-2-carboxamide) (27).** <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 8.89 (s, 2H), 8.39 – 8.15 (m, 8H), 7.88 – 7.82 (m, 2H), 7.66 – 7.61 (m, 2H), 3.75 (s, 8H), 3.66 (d, *J* = 5.3 Hz, 4H), 3.63 (d, *J* = 5.4 Hz, 4H), 3.61 – 3.58 (m, 4H), 3.55 – 3.51 (m, 4H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 759.2; found 759.2. Yield = 65%

***N,N'*-(3,6,9,12,15-Pentaoxaheptadecane-1,17-diyl)bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-2-carboxamide) (13).** <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.82 (s, 2H), 8.22 (d, *J* = 8.7 Hz, 2H), 8.03 – 7.96 (m, 2H), 7.79 (s, 4H), 7.66 (q, *J* = 8.4, 7.7 Hz, 2H), 7.36 – 7.28 (m, 2H), 3.91 (t, *J* = 7.2 Hz, 4H), 3.66 (t, *J* = 5.0 Hz, 4H), 3.64 – 3.60 (m, 4H), 3.60 – 3.57 (m, 4H), 3.57 – 3.54 (m, 4H), 3.52 – 3.51 (m, 4H), 3.48 – 3.47 (m, 4H), 3.30 – 3.22 (m, 4H), 1.86 (p, *J* = 7.6 Hz, 4H), 1.65 (p, *J* = 7.0 Hz, 4H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1081.6; found 1081.6. Yield = 60%.

### 3.7 Acknowledgements

I would like to thank Professor Anne M. Baranger for the collaboration on this project, Dr. Yuan Fu, Lien Nguyen and Kali A. Miller for their assistance with MBNL1 expression and purification, optical melting and fluorescence studies, cell preparation and splicing assays and the synthetic procedures. I thank Dr. Sivaguru Mayandi and Dr. John P. Eichorst for their help with confocal microscopy imaging.

### 3.8 References

- (1) Arambula, J. F., Ramisetty, S. R., Baranger, A. M., and Zimmerman, S. C. (2009) A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding, *Proc. Natl. Acad. Sci. U. S. A.* 106, 16068-16073.
- (2) Krishnamurthy, V. M., Estroff, L. A. and Whitesides, G. M. (2006) Multivalency in Ligand Design, in *Fragment-based Approaches in Drug Discovery* (eds W. Jahnke and D. A. Erlanson), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG. doi: 10.1002/3527608761.ch2.
- (3) Mammen, M., Choi, S. K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors, *Angew. Chem. Int. Ed.* 37, 2755-2794.
- (4) Lees, W. J., Spaltenstein, A., Kingery-Wood, J. E., and Whitesides, G. M. (1994) Polyacrylamides bearing pendant alpha-sialoside groups strongly inhibit agglutination of erythrocytes by influenza A virus: multivalency and steric stabilization of particulate biological systems, *J. Med. Chem.* 37, 3419-3433.
- (5) Jervis, P. J., Moulis, M., Jukes, J. P., Ghadbane, H., Cox, L. R., Cerundolo, V., and Besra, G. S. (2012) Towards multivalent CD1d ligands: synthesis and biological activity of homodimeric alpha-galactosyl ceramide analogues, *Carbohydr. Res.* 356, 152-162.
- (6) Sucheck, S. J., Wong, A. L., Koeller, K. M., Boehr, D. D., Draker, K.-a., Sears, P., Wright, G. D., and Wong, C.-H. (2000) Design of Bifunctional Antibiotics that Target Bacterial rRNA and Inhibit Resistance-Causing Enzymes, *J. Am. Chem. Soc.* 122, 5230-5231.
- (7) Tamiz, A. P., Zhang, J., Zhang, M., Wang, C. Z., Johnson, K. M., and Kozikowski, A. P. (2000) Application of the Bivalent Ligand Approach to the Design of Novel Dimeric Serotonin Reuptake Inhibitors, *J. Am. Chem. Soc.* 122, 5393-5394.
- (8) Gestwicki, J. E., Cairo, C. W., Strong, L. E., Oetjen, K. a., and Kiessling, L. L. (2002) Influencing receptor-ligand binding mechanisms with multivalent ligand architecture., *J. Am. Chem. Soc.* 124, 14922-14933.
- (9) Agnelli, F., Sucheck, S. J., Marby, K. A., Rabuka, D., Yao, S. L., Sears, P. S., Liang, F. S., and Wong, C. H. (2004) Dimeric aminoglycosides as antibiotics, *Angew. Chem. Int. Ed. Engl.* 43, 1562-1566.

- (10) Lee, M. M., Pushechnikov, A., and Disney, M. D. (2009) Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2., *ACS Chem. Biol.* *4*, 345-355.
- (11) Childs-Disney, J. L., Tsitovich, P. B., and Disney, M. D. (2011) Using modularly assembled ligands to bind RNA internal loops separated by different distances, *ChemBioChem* *12*, 2143-2146.
- (12) Shonberg, J., Scammells, P. J., and Capuano, B. (2011) Design strategies for bivalent ligands targeting GPCRs, *ChemMedChem* *6*, 963-974.
- (13) Childs-Disney, J. L., Hoskins, J., Rzuczek, S. G., Thornton, C. A., and Disney, M. D. (2012) Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive, *ACS Chem. Biol.* *7*, 856-862.
- (14) Kitov, P. I., and Bundle, D. R. (2003) On the nature of the multivalency effect: a thermodynamic model, *J. Am. Chem. Soc.* *125*, 16271-16284.
- (15) Zhang, Y., Gilliam, A., Maitra, R., Damaj, M. I., Tajuba, J. M., Seltzman, H. H., and Thomas, B. F. (2010) Synthesis and biological evaluation of bivalent ligands for the cannabinoid 1 receptor, *J. Med. Chem.* *53*, 7048-7060.
- (16) Birnkammer, T., Spickenreither, A., Brunskole, I., Lopuch, M., Kagermeier, N., Bernhardt, G., Dove, S., Seifert, R., Elz, S., and Buschauer, A. (2012) The bivalent ligand approach leads to highly potent and selective acylguanidine-type histamine H(2) receptor agonists, *J. Med. Chem.* *55*, 1147-1160.
- (17) Kuhhorn, J., Gotz, A., Hubner, H., Thompson, D., Whistler, J., and Gmeiner, P. (2011) Development of a bivalent dopamine D(2) receptor agonist, *J. Med. Chem.* *54*, 7911-7919.
- (18) Kuhhorn, J., Hubner, H., and Gmeiner, P. (2011) Bivalent dopamine D2 receptor ligands: synthesis and binding properties, *J. Med. Chem.* *54*, 4896-4903.
- (19) Peng, Y., Sun, H., Lu, J., Liu, L., Cai, Q., Shen, R., Yang, C. Y., Yi, H., and Wang, S. (2012) Bivalent Smac mimetics with a diazabicyclic core as highly potent antagonists of XIAP and clAP1/2 and novel anticancer agents, *J. Med. Chem.* *55*, 106-114.
- (20) Liu, J., Brahimi, F., Saragovi, H. U., and Burgess, K. (2010) Bivalent diketopiperazine-based tropomyosin receptor kinase C (TrkC) antagonists, *J. Med. Chem.* *53*, 5044-5048.
- (21) Wakelin, G. (1986) G. Wakelin, *Chemotherapy* *6*, 275-340.

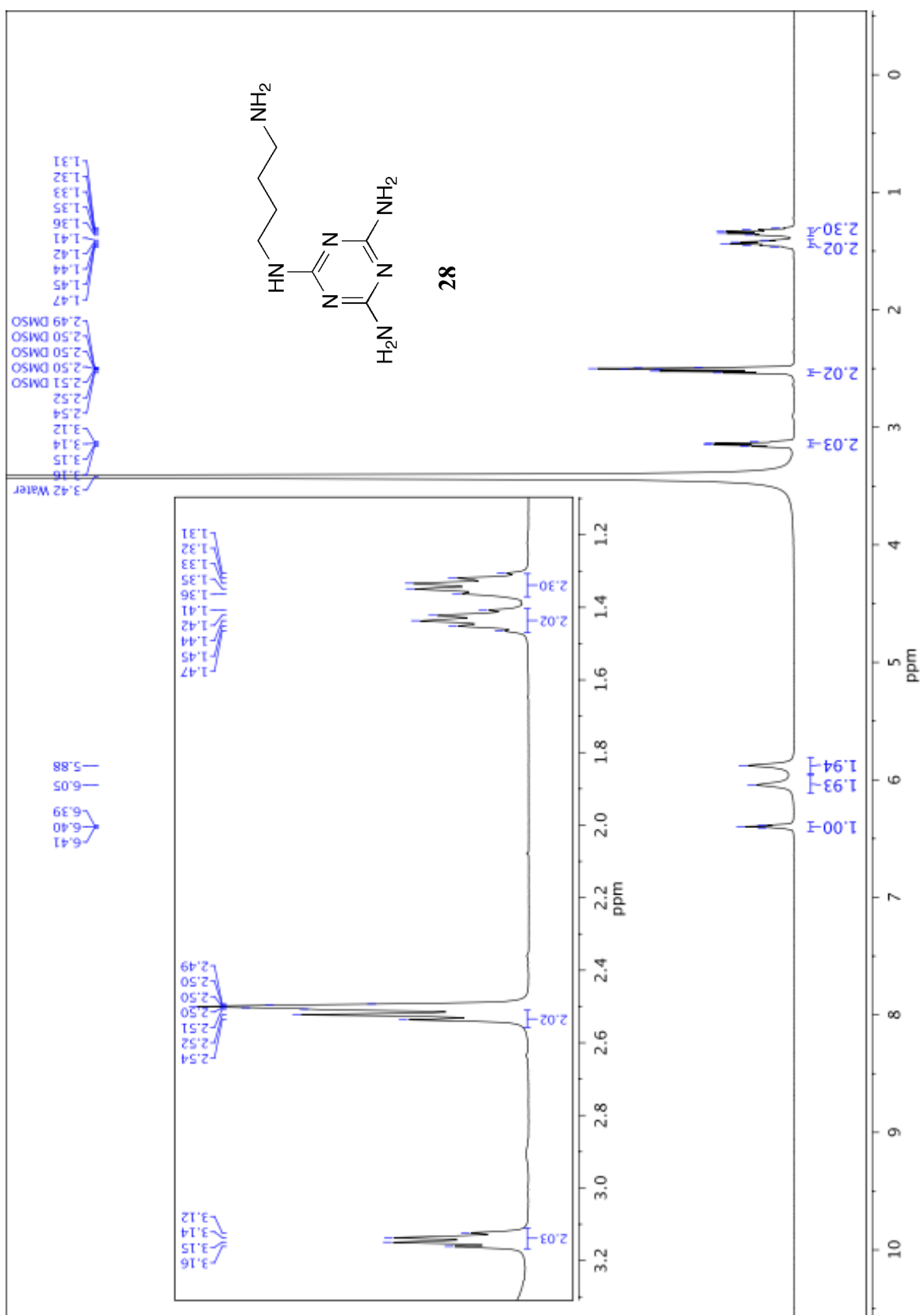
- (22) Zimmerman, S. C., Lamberson, C. R., Cory, M., and Fairley, T. A. (1989) Topologically Constrained Bifunctional Intercalators - DNA Intercalation by a Macrocyclic Bisacridine, *J. Am. Chem. Soc.* *111*, 6805-6809.
- (23) Veal, J. M., Li, Y., Zimmerman, S. C., Lamberson, C. R., Cory, M., Zon, G., and Wilson, W. D. (1990) Interaction of a macrocyclic bisacridine with DNA, *Biochemistry* *29*, 10918-10927.
- (24) Denny, W. a. (2002) Acridine derivatives as chemotherapeutic agents., *Curr. Med. Chem.* *9*, 1655-1665.
- (25) Fechter, E. J., Olenyuk, B., and Dervan, P. B. (2004) Design of a sequence-specific DNA bisintercalator, *Angew. Chem. Int. Ed. Engl.* *43*, 3591-3594.
- (26) Goodell, J. R., Madhok, A. A., Hiasa, H., and Ferguson, D. M. (2006) Synthesis and evaluation of acridine- and acridone-based anti-herpes agents with topoisomerase activity, *Bioorg. Med. Chem.* *14*, 5467-5480.
- (27) Antonini, I., Polucci, P., Magnano, A., Gatto, B., Palumbo, M., Menta, E., Pescalli, N., and Martelli, S. (2003) Design, synthesis, and biological properties of new bis(acridine-4-carboxamides) as anticancer agents., *J. Med. Chem.* *46*, 3109-3115.
- (28) Caffrey, C. R., Steverding, D., Swenerton, R. K., Kelly, B., Walshe, D., Debnath, A., Zhou, Y. M., Doyle, P. S., Fafarman, A. T., Zorn, J. A., Land, K. M., Beauchene, J., Schreiber, K., Moll, H., Ponte-Sucre, A., Schirmeister, T., Saravanamuthu, A., Fairlamb, A. H., Cohen, F. E., McKerrow, J. H., Weisman, J. L., and May, B. C. (2007) Bis-acridines as lead antiparasitic agents: structure-activity analysis of a discrete compound library in vitro, *Antimicrob. Agents Chemother.* *51*, 2164-2172.
- (29) Bobrovnik, S. A. (2007) The influence of rigid or flexible linkage between two ligands on the effective affinity and avidity for reversible interactions with bivalent receptors, *Journal of molecular recognition : JMR* *20*, 253-262.
- (30) Fischer, W., Brissault, B., Prevost, S., Kopaczynska, M., Andreou, I., Janosch, A., Gradzielski, M., and Haag, R. (2010) Synthesis of linear polyamines with different amine spacings and their ability to form dsDNA/siRNA complexes suitable for transfection, *Macromol. Biosci.* *10*, 1073-1083.
- (31) Wong, C. H., Richardson, S. L., Ho, Y. J., Lucas, A. M., Tuccinardi, T., Baranger, A. M., and Zimmerman, S. C. (2012) Investigating the Binding Mode of an Inhibitor of the MBNL1RNA Complex in Myotonic Dystrophy Type 1 (DM1) Leads to the

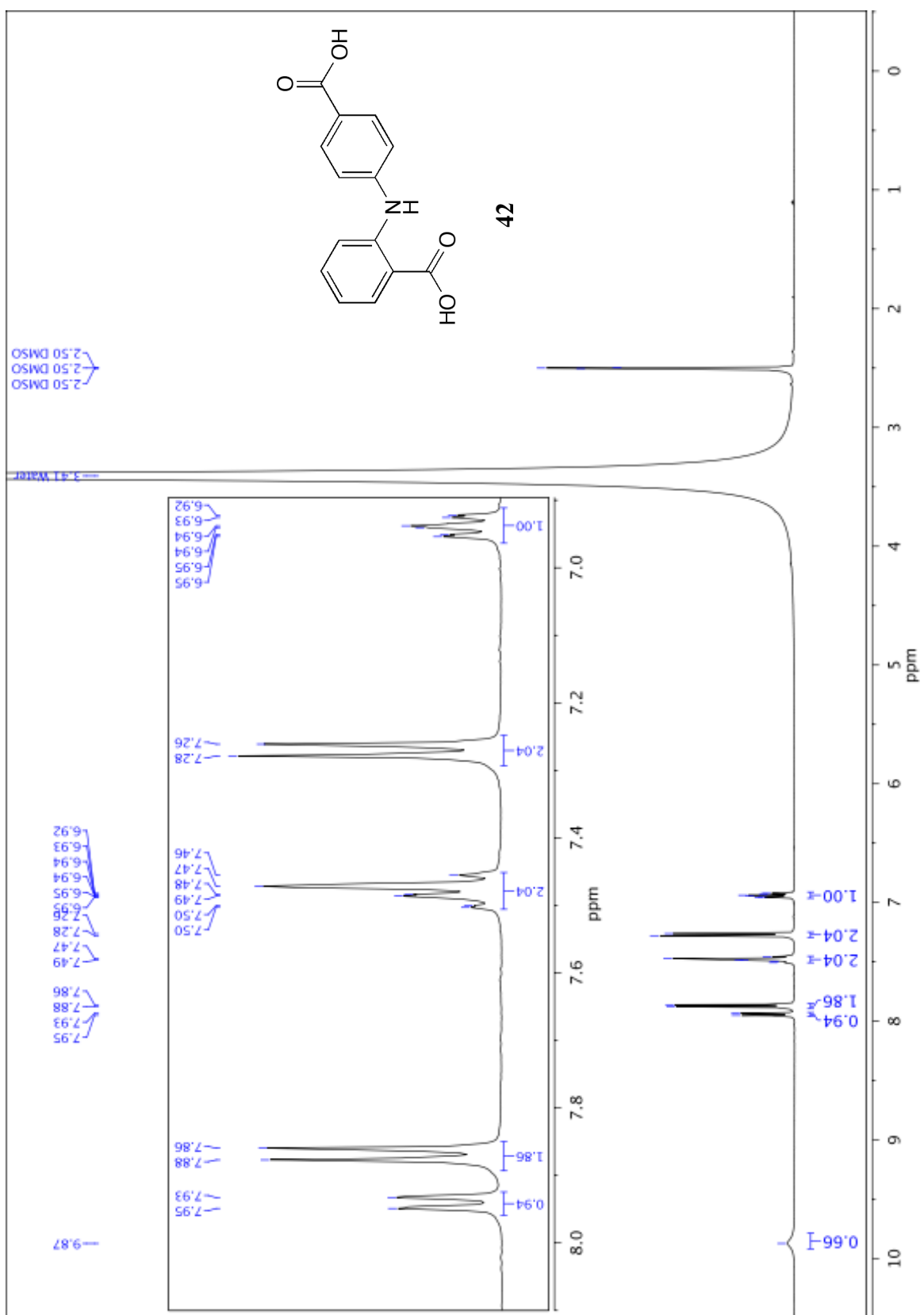
Unexpected Discovery of a DNA-Selective Binder, *ChemBioChem*, DOI: 10.1002/cbic.201200602.

- (32) Laurent, F. X., Sureau, A., Klein, A. F., Trouslard, F., Gasnier, E., Furling, D., and Marie, J. (2012) New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats, *Nucleic Acids Res.* *40*, 3159-3171.
- (33) Fu, Y., Ramisetty, S. R., Hussain, N., and Baranger, A. M. (2012) MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation, *ChemBioChem* *13*, 112-119.
- (34) Alessi, M. L., Norman, A. I., Knowlton, S. E., Ho, D. L., and Greer, S. C. (2005) Helical and coil conformations of poly(ethylene glycol) in isobutyric acid and water, *Macromolecules* *38*, 9333-9340.
- (35) Vamosi, G., Gohlke, C., and Clegg, R. M. (1996) Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: multiple conformers of single-stranded and double-stranded dye-DNA complexes, *Biophys. J.* *71*, 972-994.
- (36) Qu, P., Chen, X. D., Zhou, X. X., Li, X., and Zhao, X. S. (2009) Fluorescence quenching of TMR by guanosine in oligonucleotides, *Sci. China, Ser. B* *52*, 1653-1659.
- (37) Feng, B. Y., and Shoichet, B. K. (2006) A detergent-based assay for the detection of promiscuous inhibitors, *Nat. Protoc.* *1*, 550-553.
- (38) Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., and Swanson, M. S. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs, *Nucleic Acids Res.* *35*, 5474-5486.
- (39) Ho, T. H., Savkur, R. S., Poulos, M. G., Mancini, M. A., Swanson, M. S., and Cooper, T. A. (2005) Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy, *J. Cell Sci.* *118*, 2923-2933.
- (40) Cuenca, F., Moore, M. J., Johnson, K., Guyen, B., De Cian, A., and Neidle, S. (2009) Design, synthesis and evaluation of 4,5-di-substituted acridone ligands with high G-quadruplex affinity and selectivity, together with low toxicity to normal cells, *Bioorg. Med. Chem. Lett.* *19*, 5109-5113.
- (41) Mei, X., August, A. T., and Wolf, C. (2006) Regioselective copper-catalyzed amination of chlorobenzoic acids: synthesis and solid-state structures of N-aryl anthranilic acid derivatives., *J. Org. Chem.* *71*, 142-149.

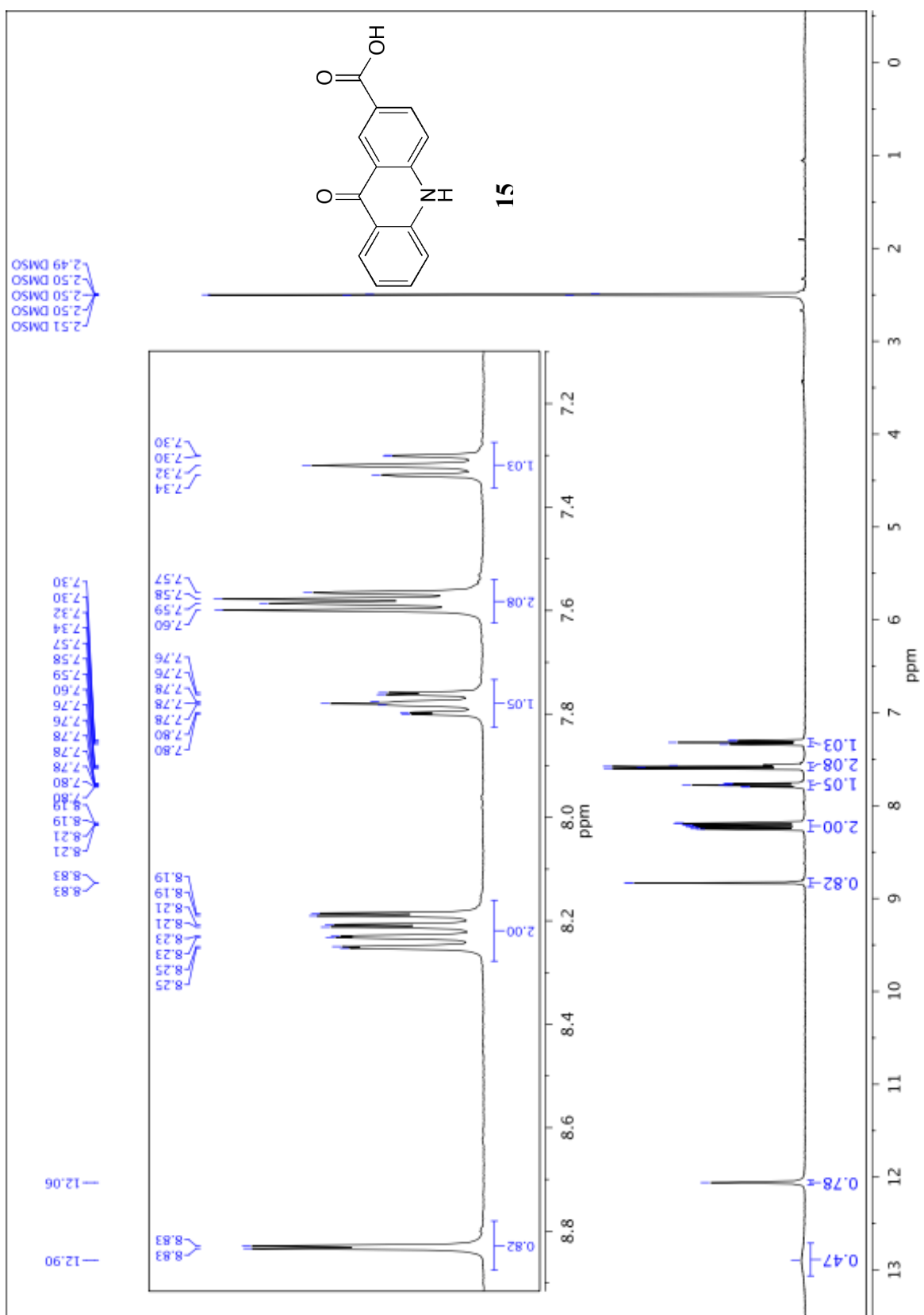
- (42) Kauffman, M. (1990) 4-(2=Carboxyphenyl)aminobenzenealkanoic, *J. Pharm. Sci.* 79, 173-178.
- (43) Ilies, M. A., Seitz, W. a., Johnson, B. H., Ezell, E. L., Miller, A. L., Thompson, E. B., and Balaban, A. T. (2006) Lipophilic pyrylium salts in the synthesis of efficient pyridinium-based cationic lipids, gemini surfactants, and lipophilic oligomers for gene delivery., *J. Med. Chem.* 49, 3872-3887.
- (44) Carta, F., Temperini, C., Innocenti, A., Scozzafava, A., Kaila, K., and Supuran, C. T. (2010) Polyamines inhibit carbonic anhydrases by anchoring to the zinc-coordinated water molecule, *J. Med. Chem.* 53, 5511-5522.
- (45) LaFrate, A. L., Carlson, K. E., and Katzenellenbogen, J. A. (2009) Steroidal bivalent ligands for the estrogen receptor: design, synthesis, characterization and binding affinities, *Bioorg. Med. Chem.* 17, 3528-3535.
- (46) Srinivasachari, S., Liu, Y., Zhang, G., Prevet, L., and Reineke, T. M. (2006) Trehalose click polymers inhibit nanoparticle aggregation and promote pDNA delivery in serum., *J. Am. Chem. Soc.* 128, 8176-8184.
- (47) Schaffert, D., Badgular, N., and Wagner, E. (2011) Novel Fmoc-polyamino acids for solid-phase synthesis of defined polyamidoamines., *Org. Lett.* 13, 1586-1589.

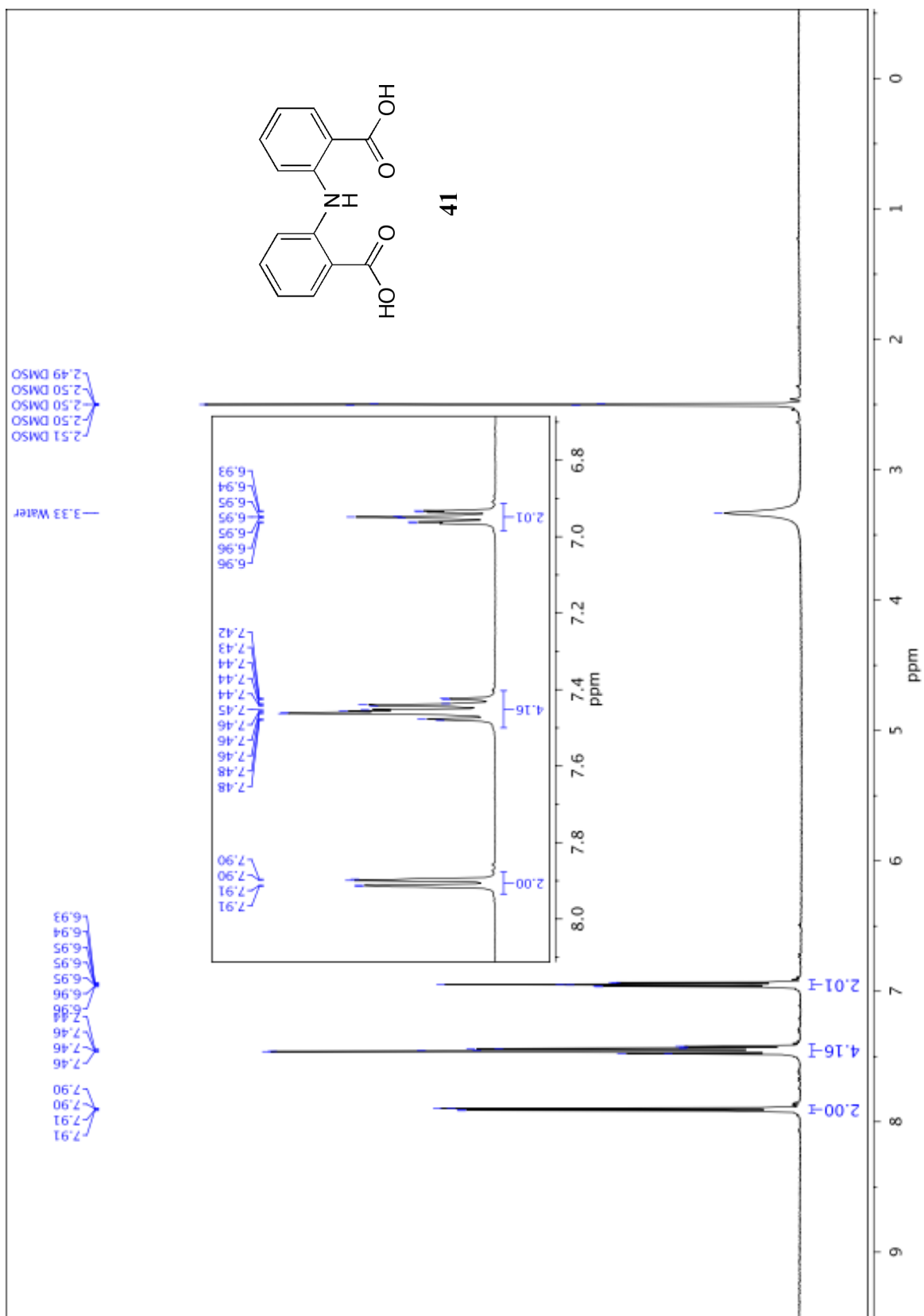
### 3.9 NMR Spectra

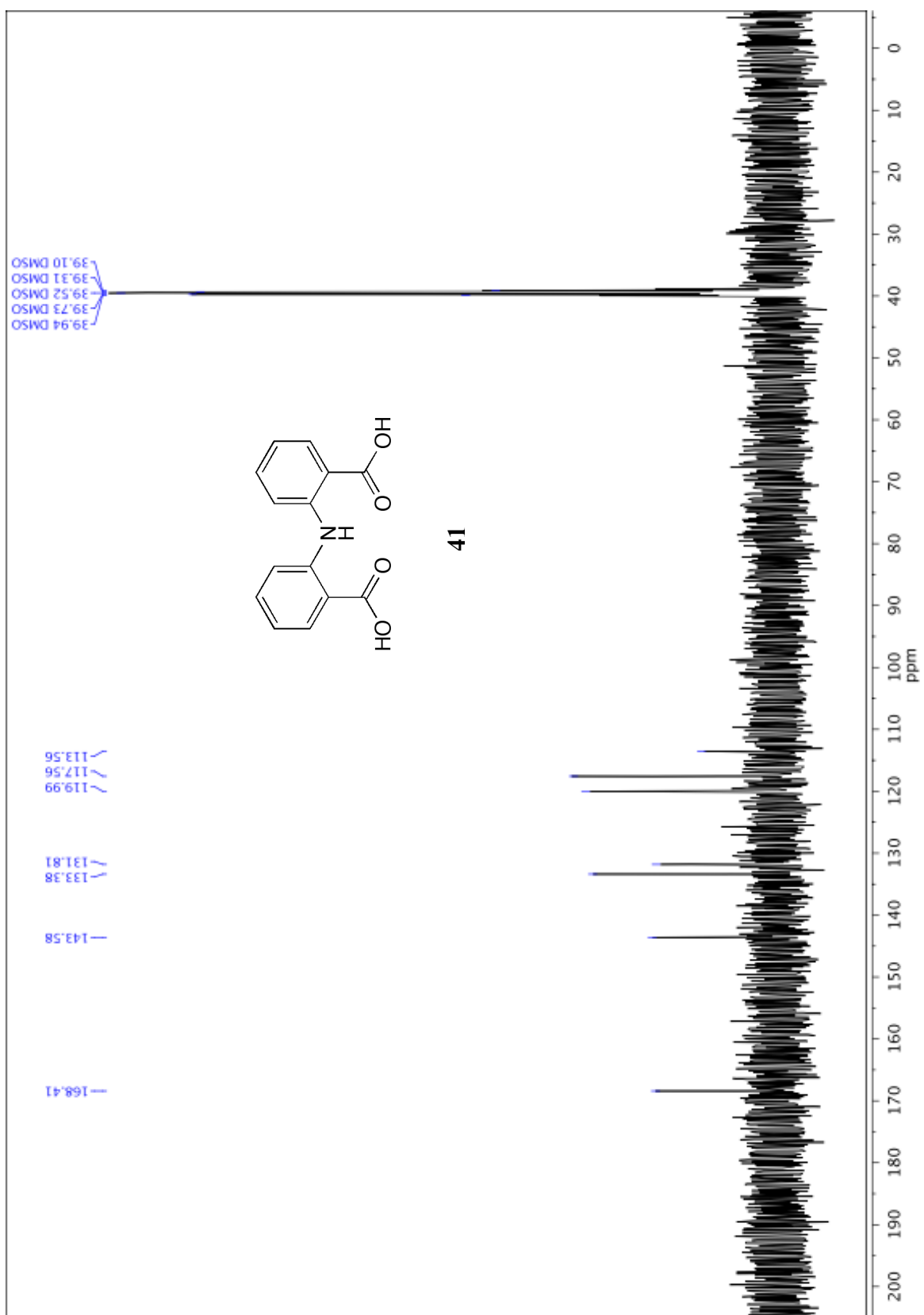


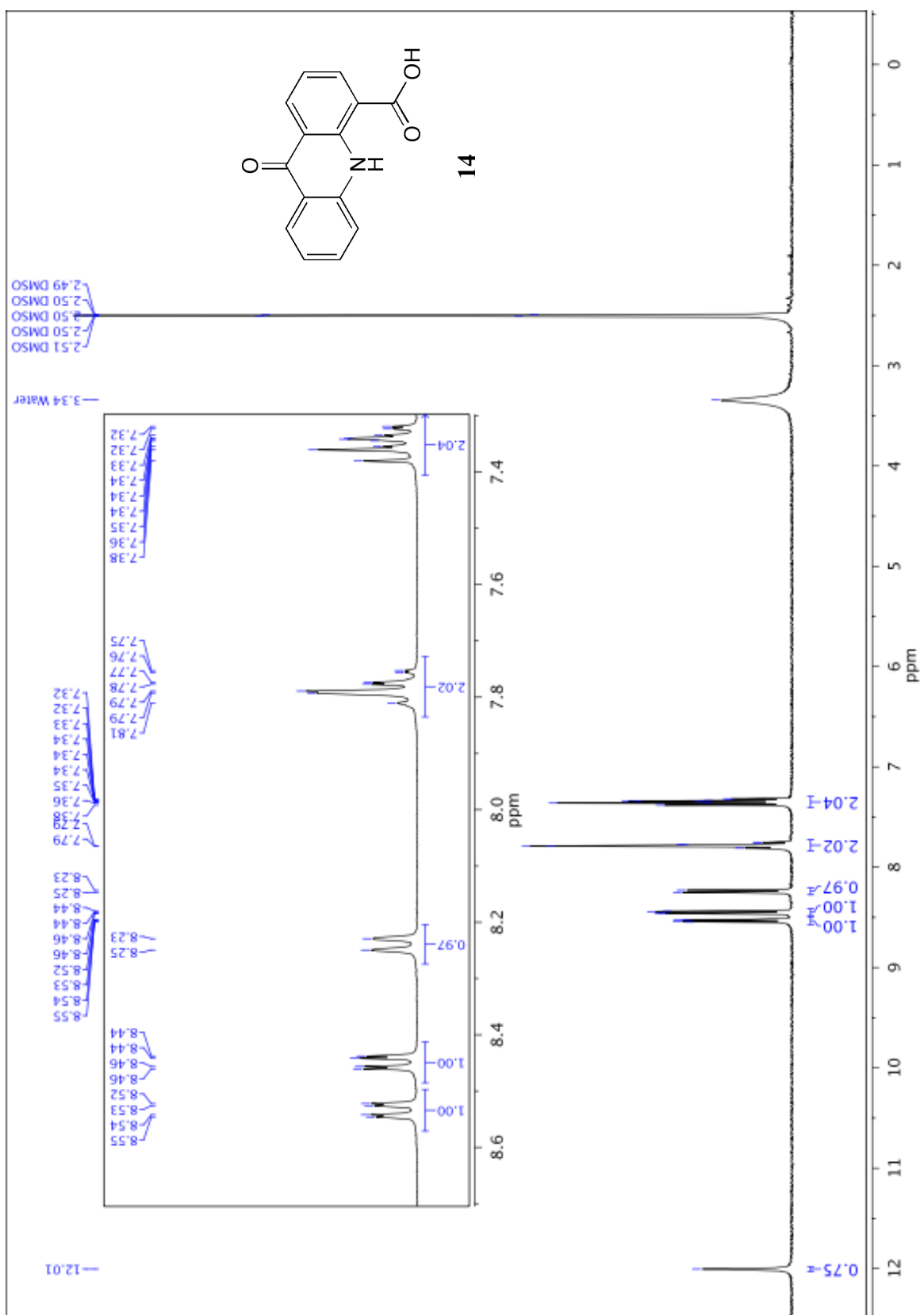


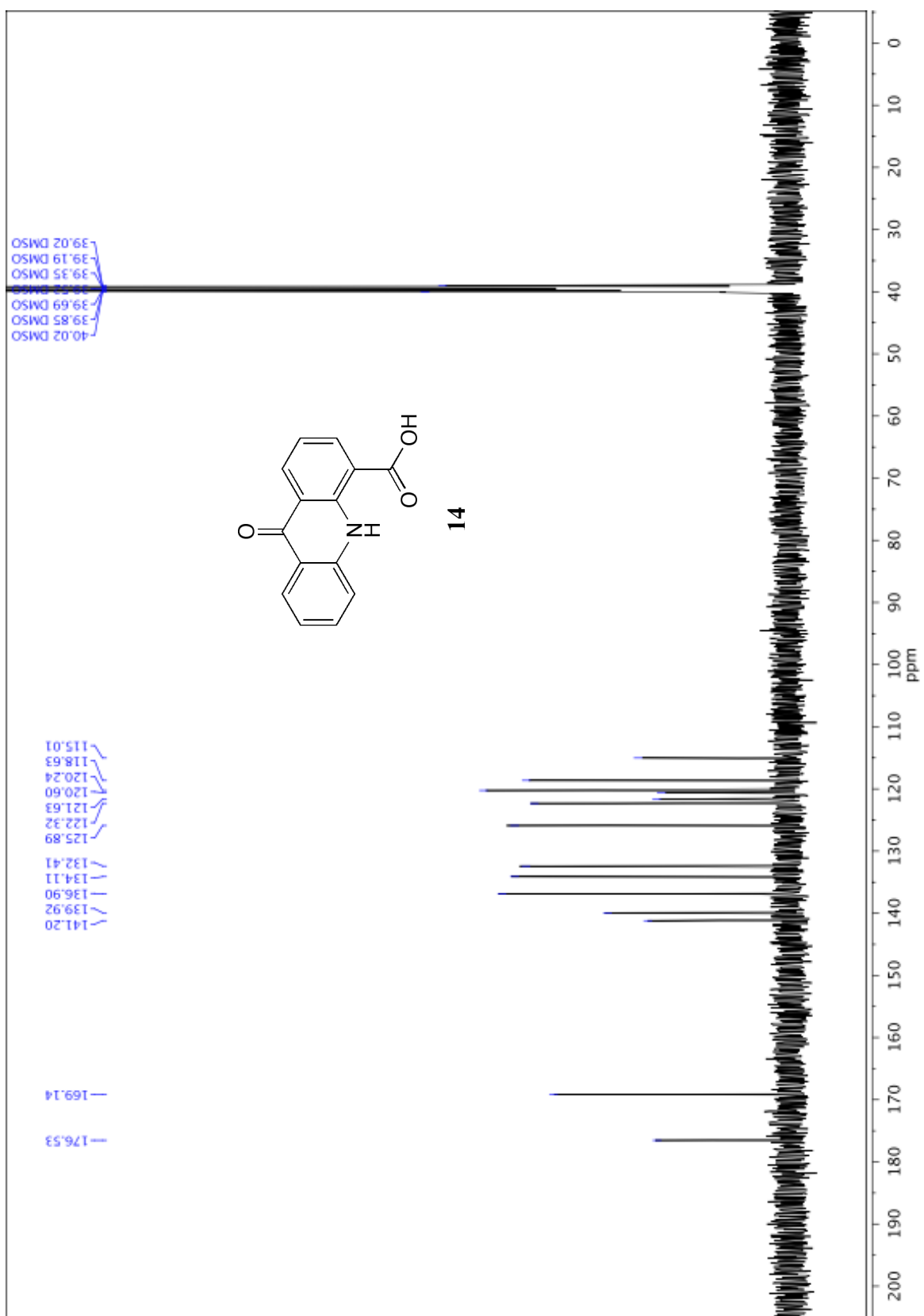


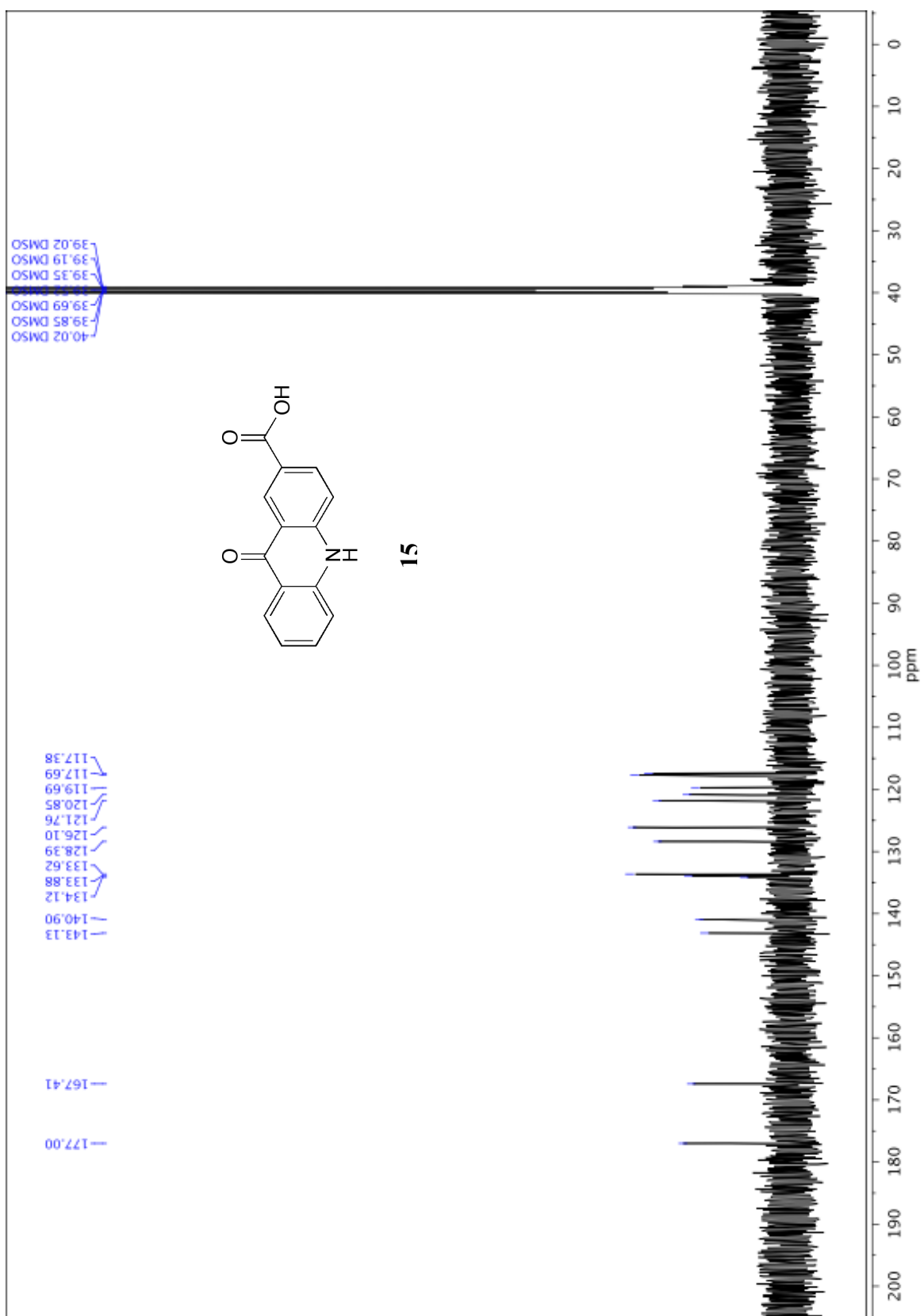


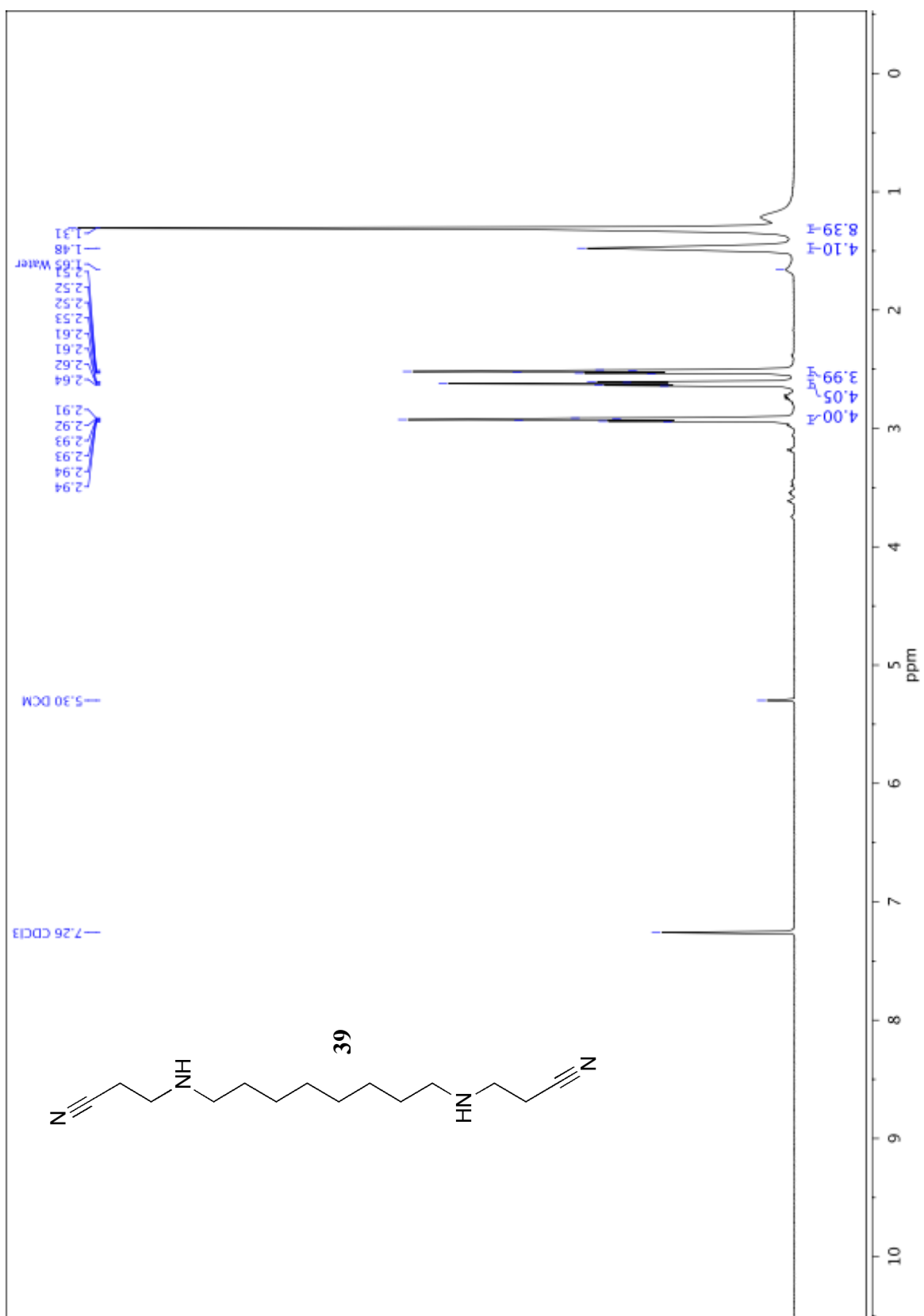


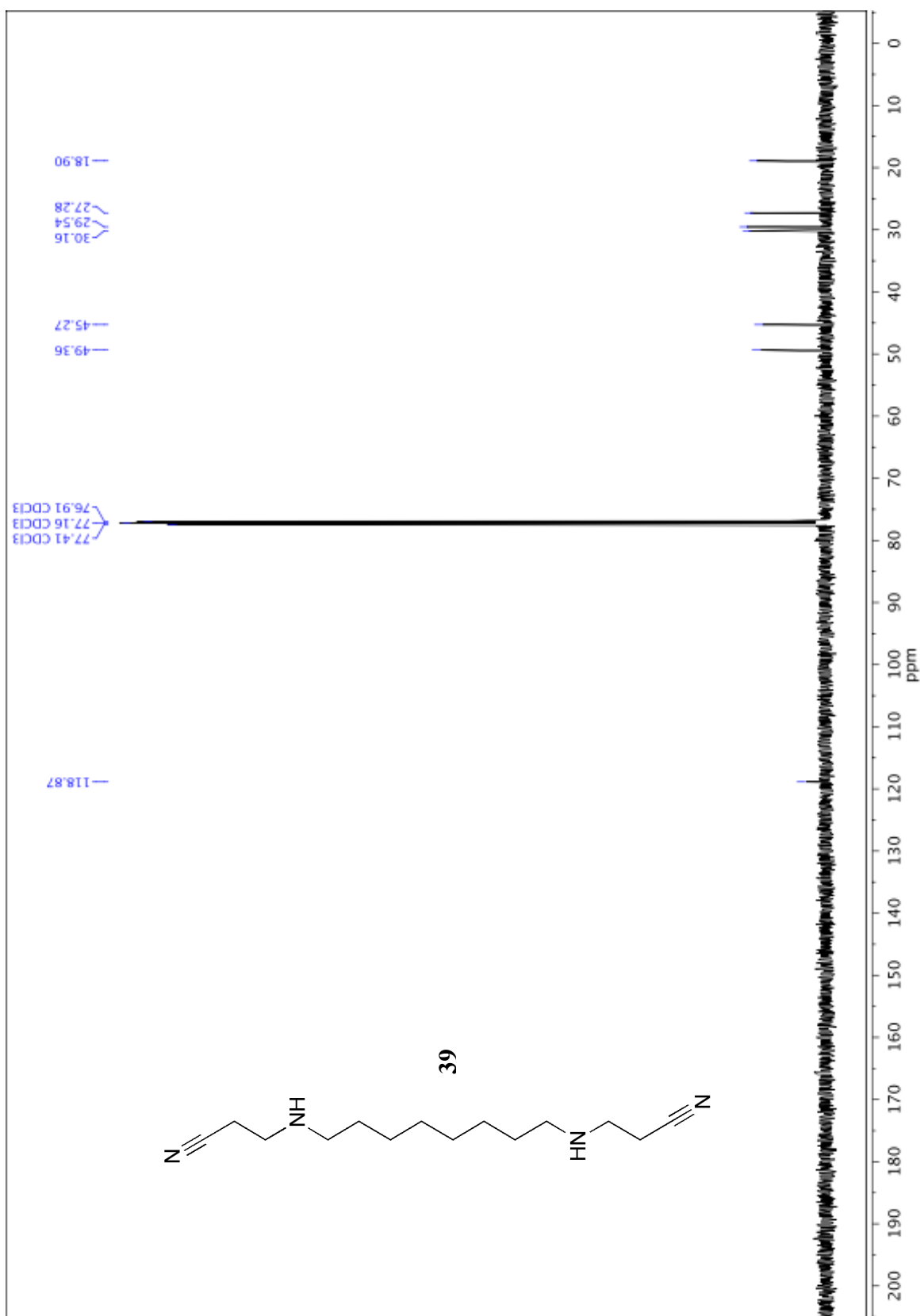




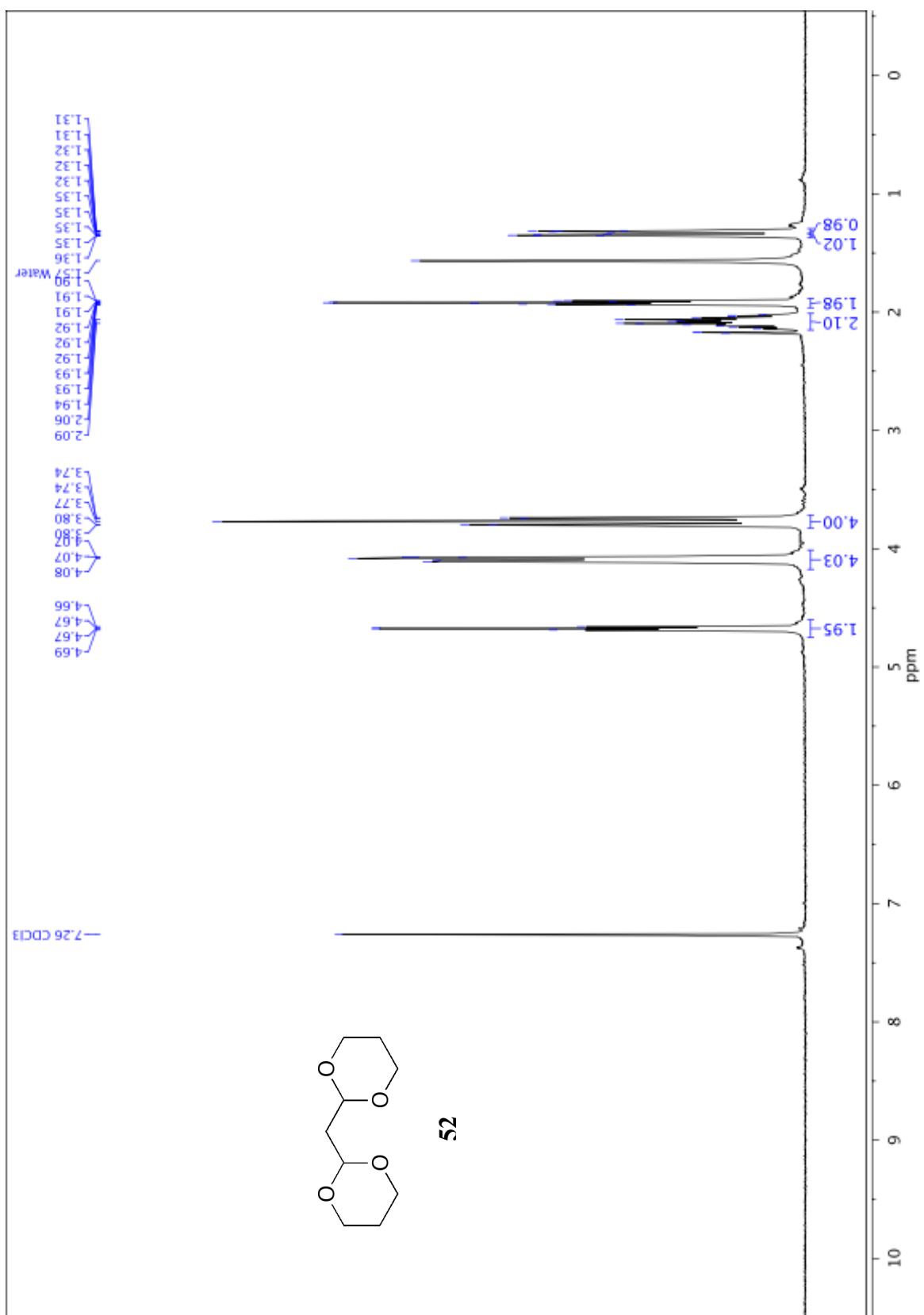


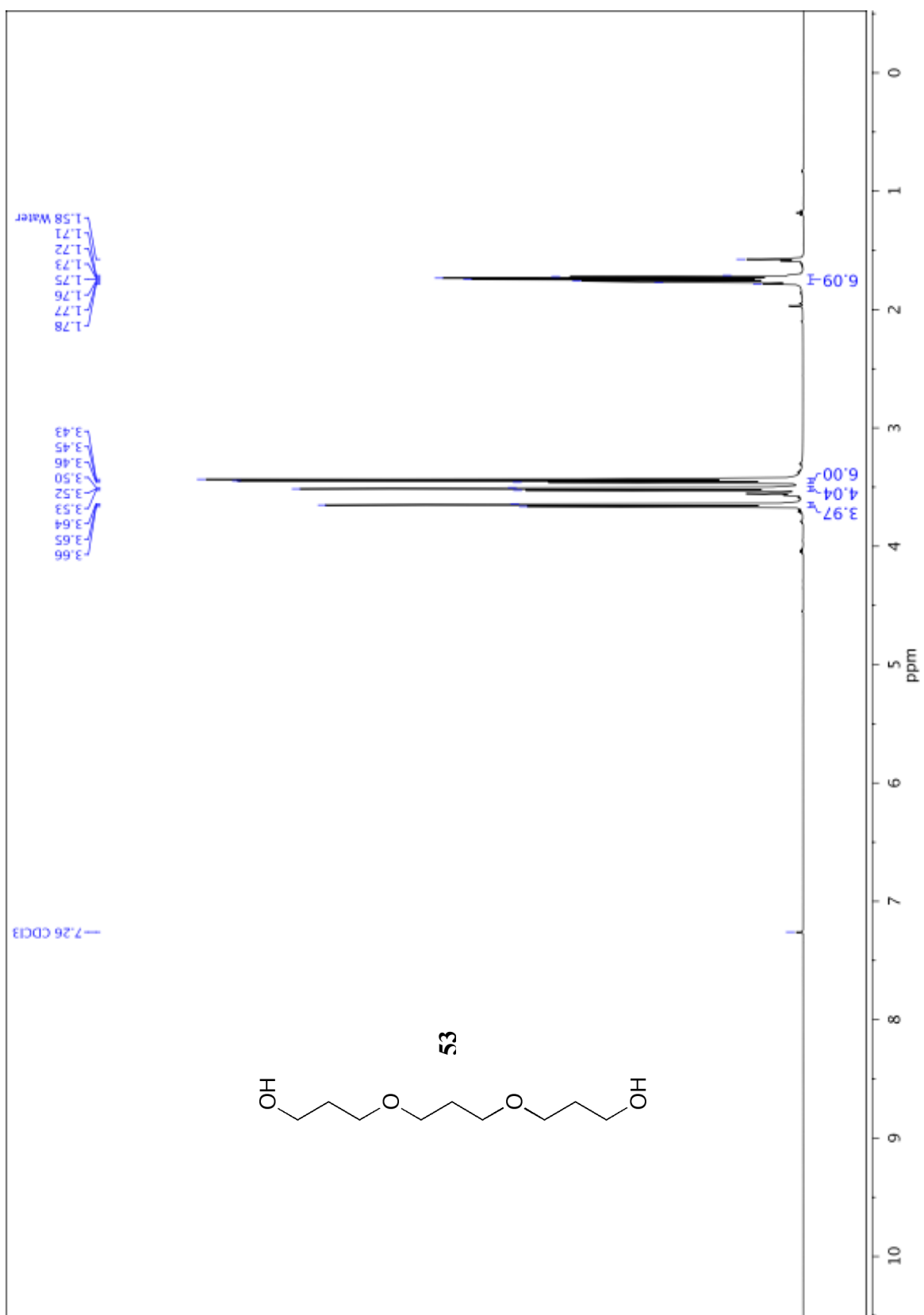


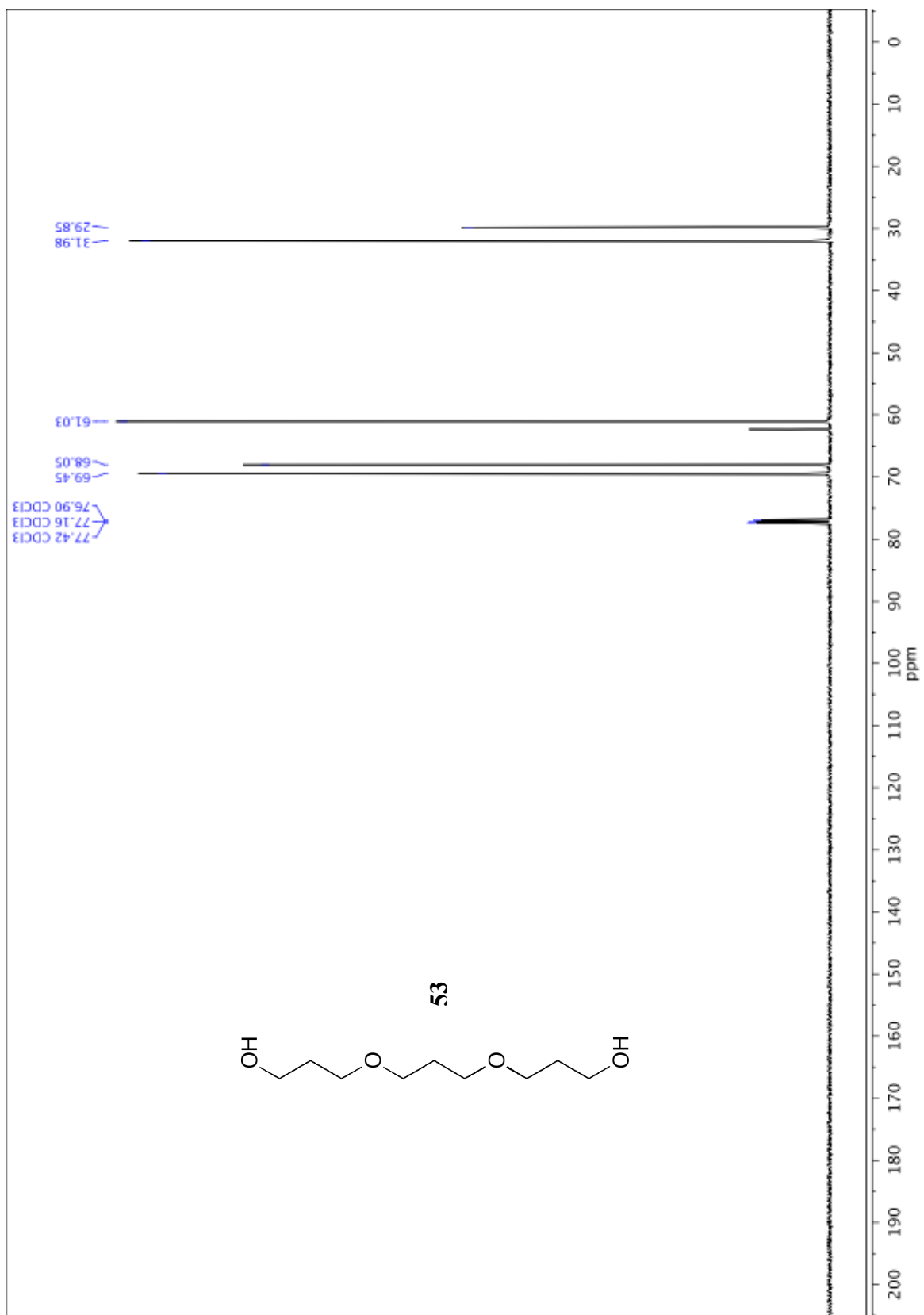


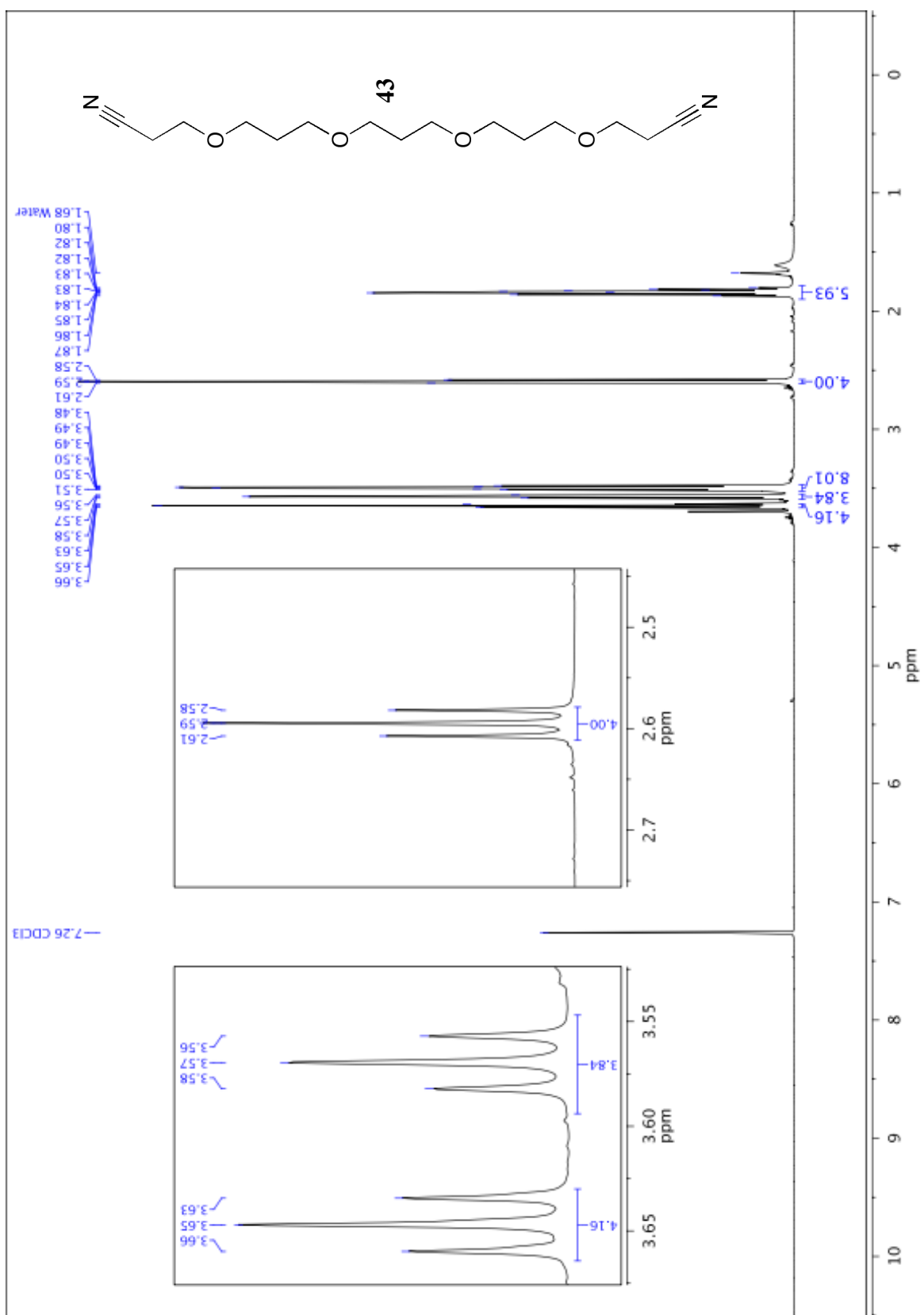


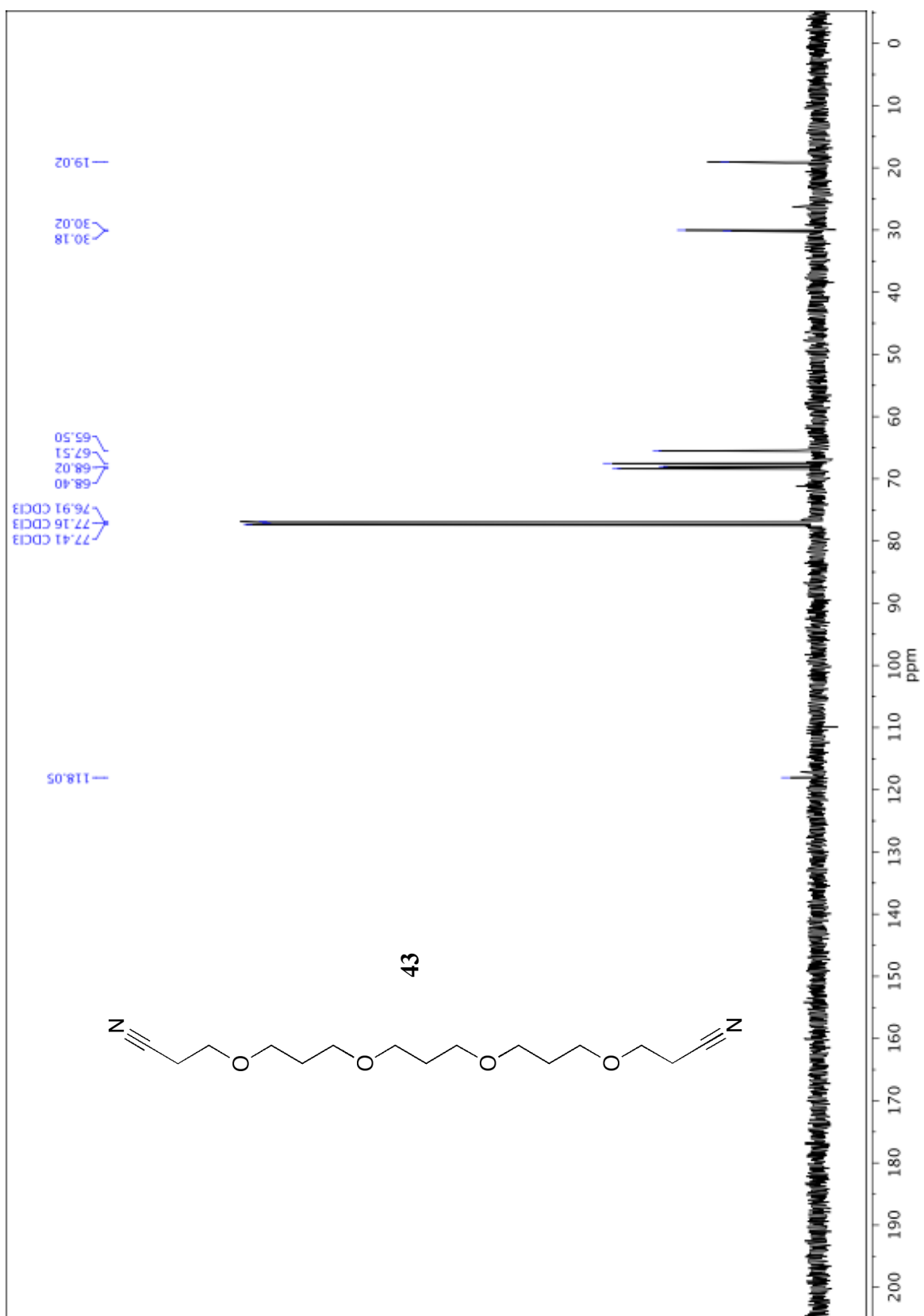


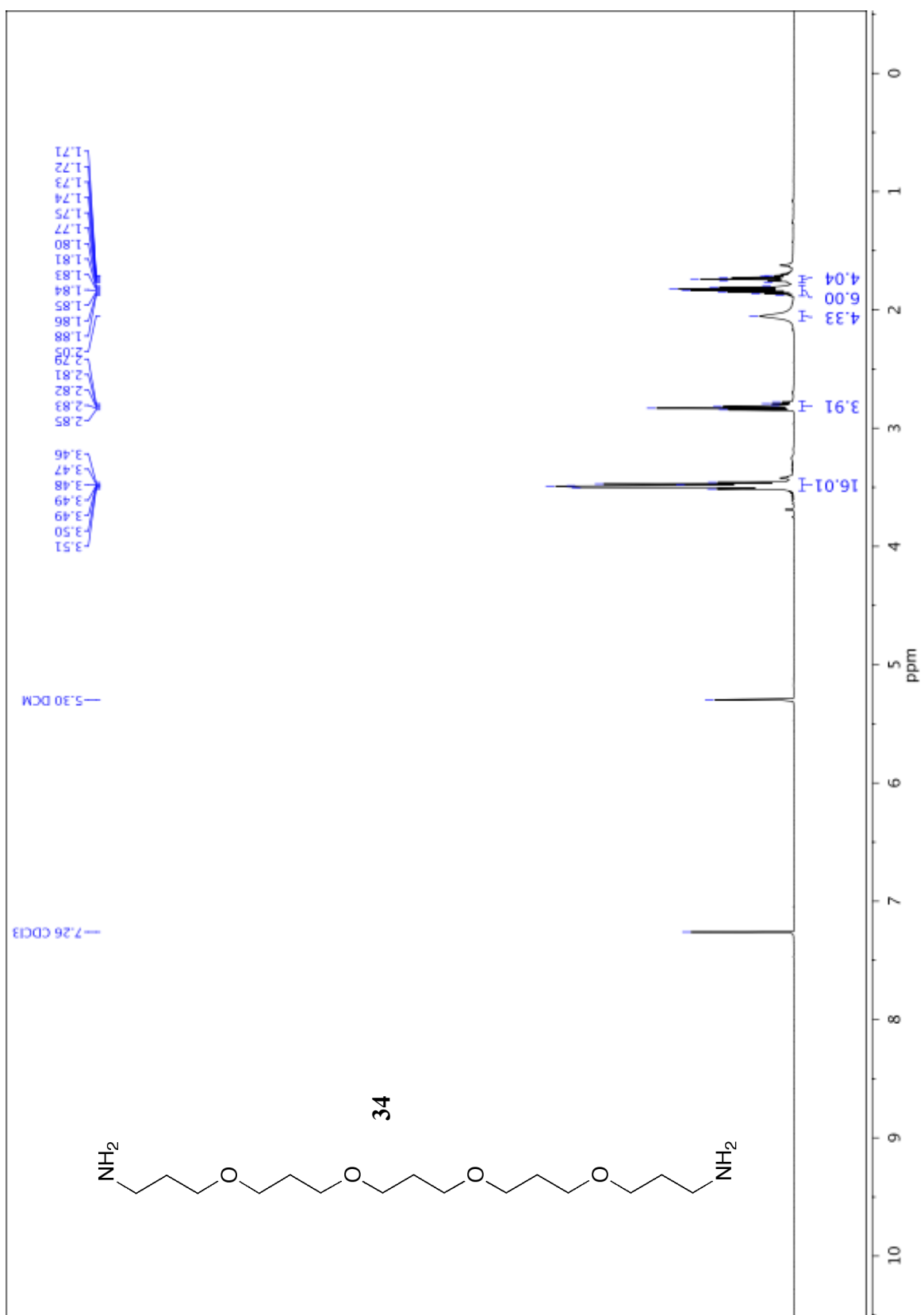


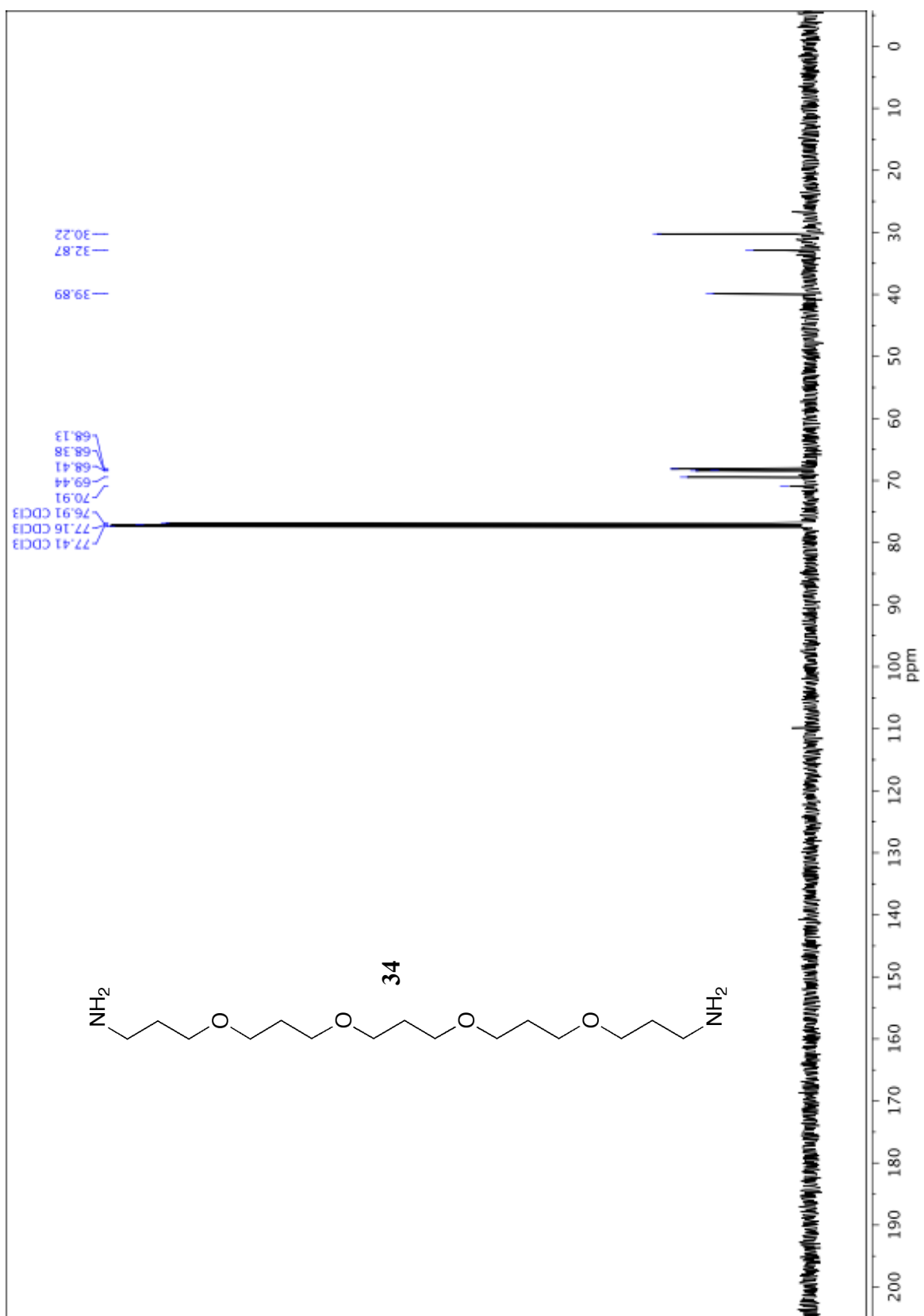


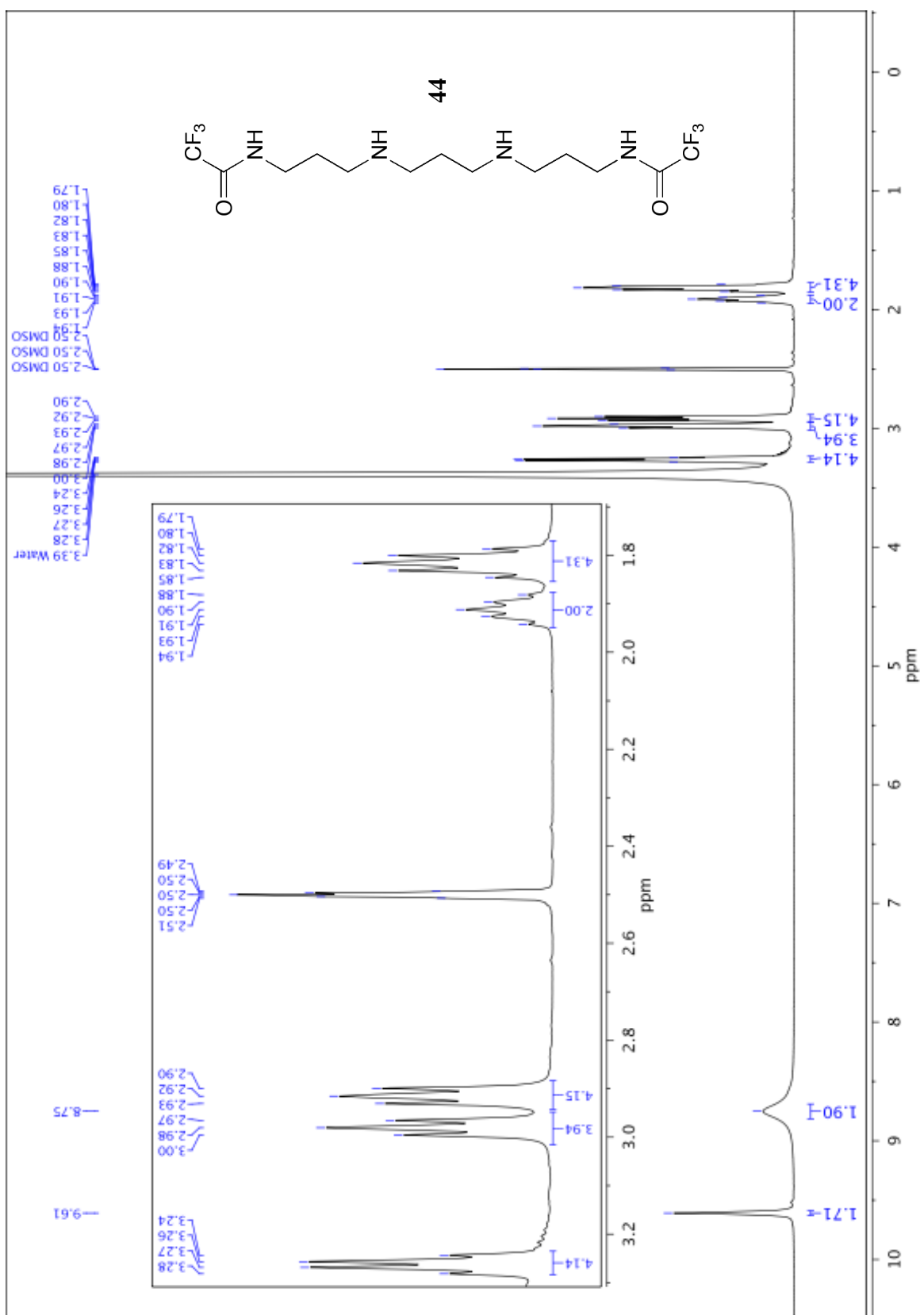




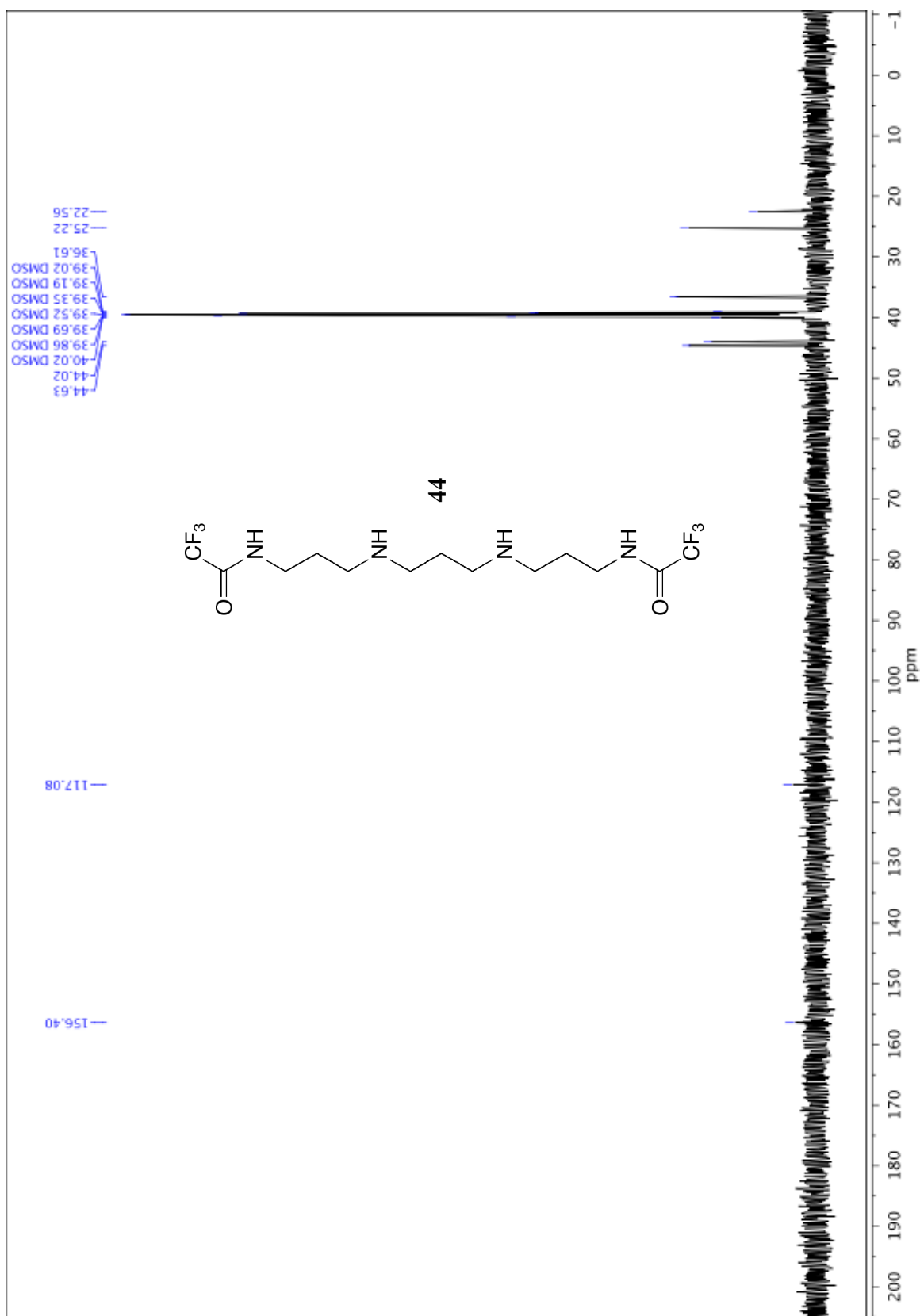


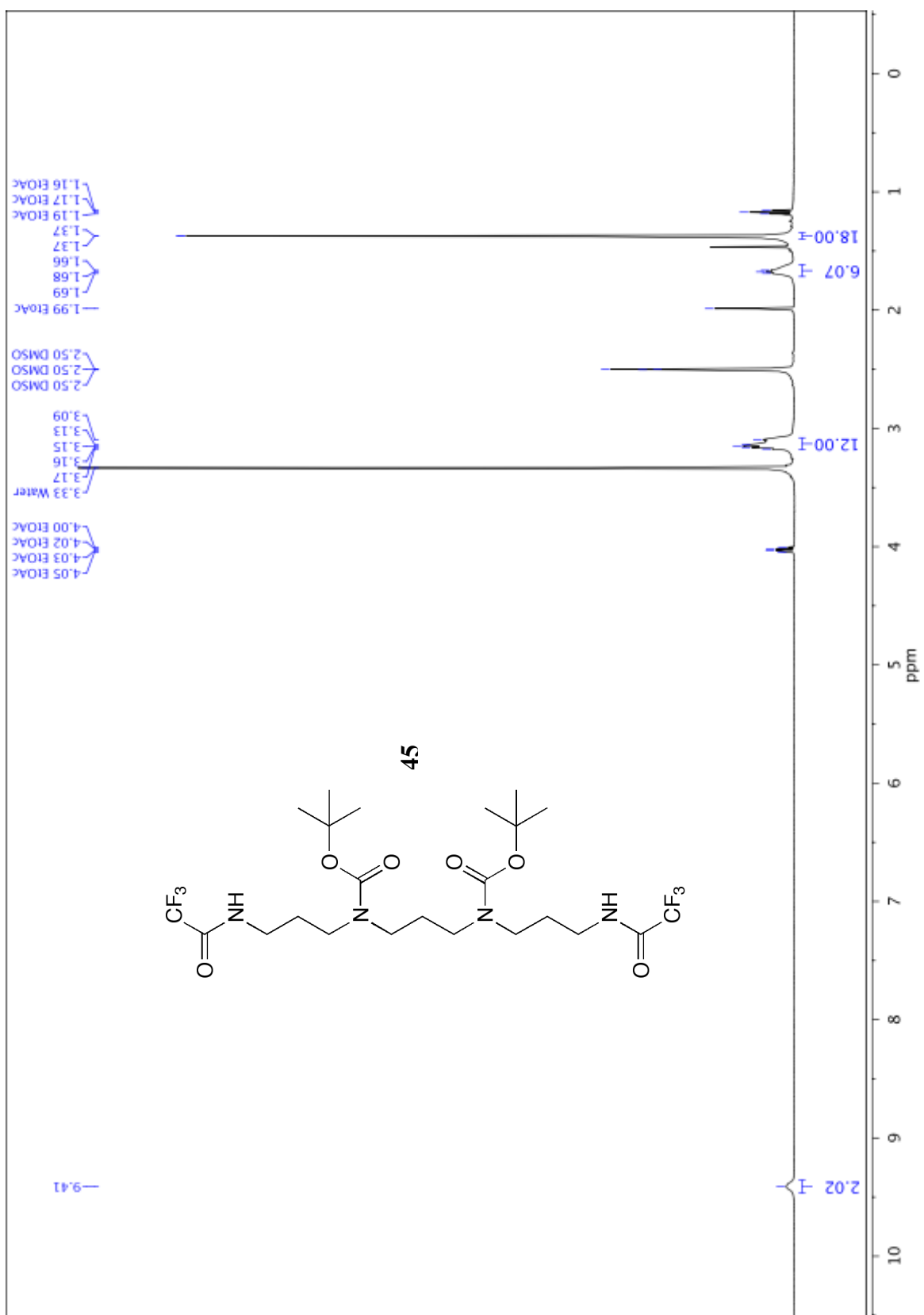


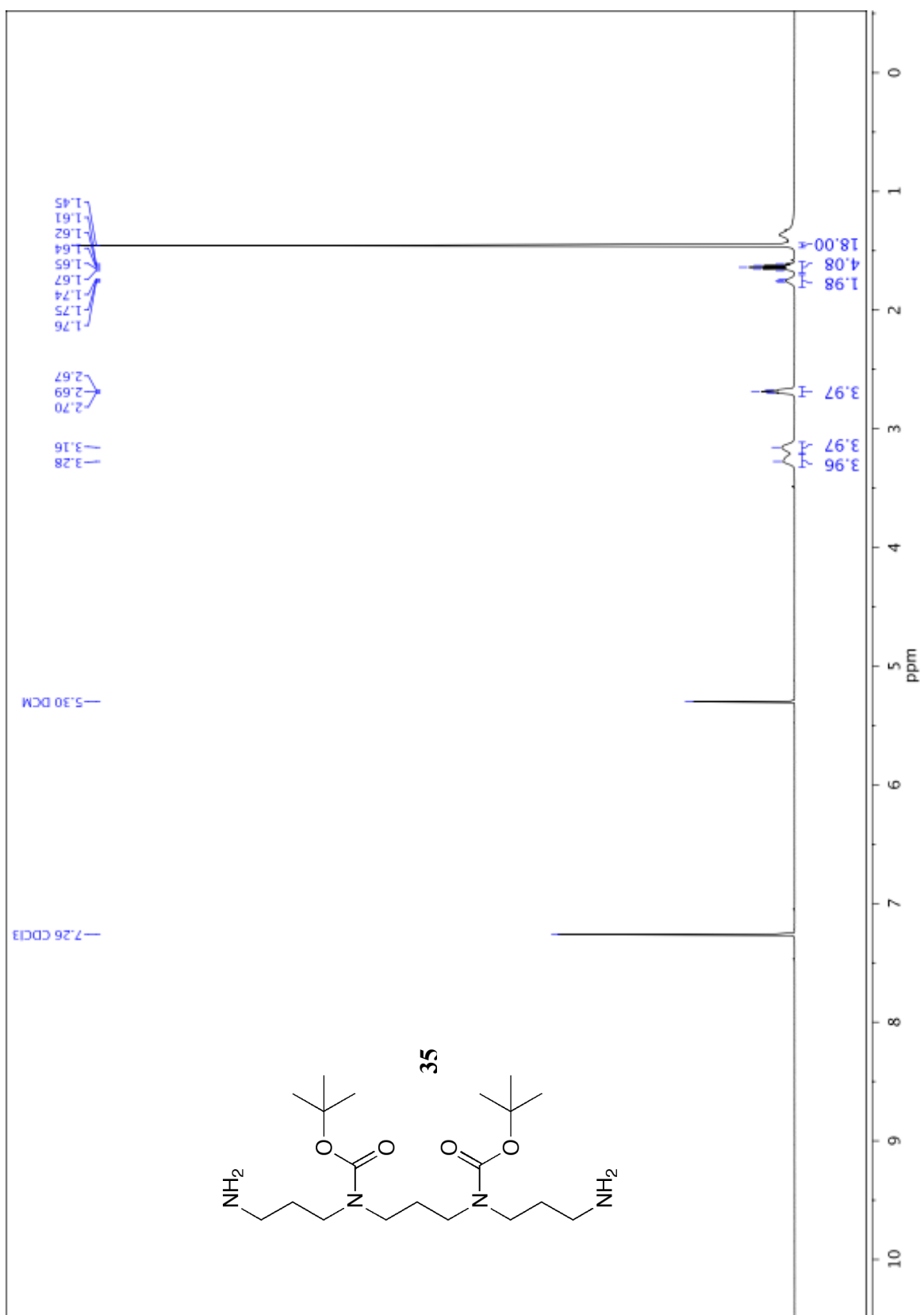


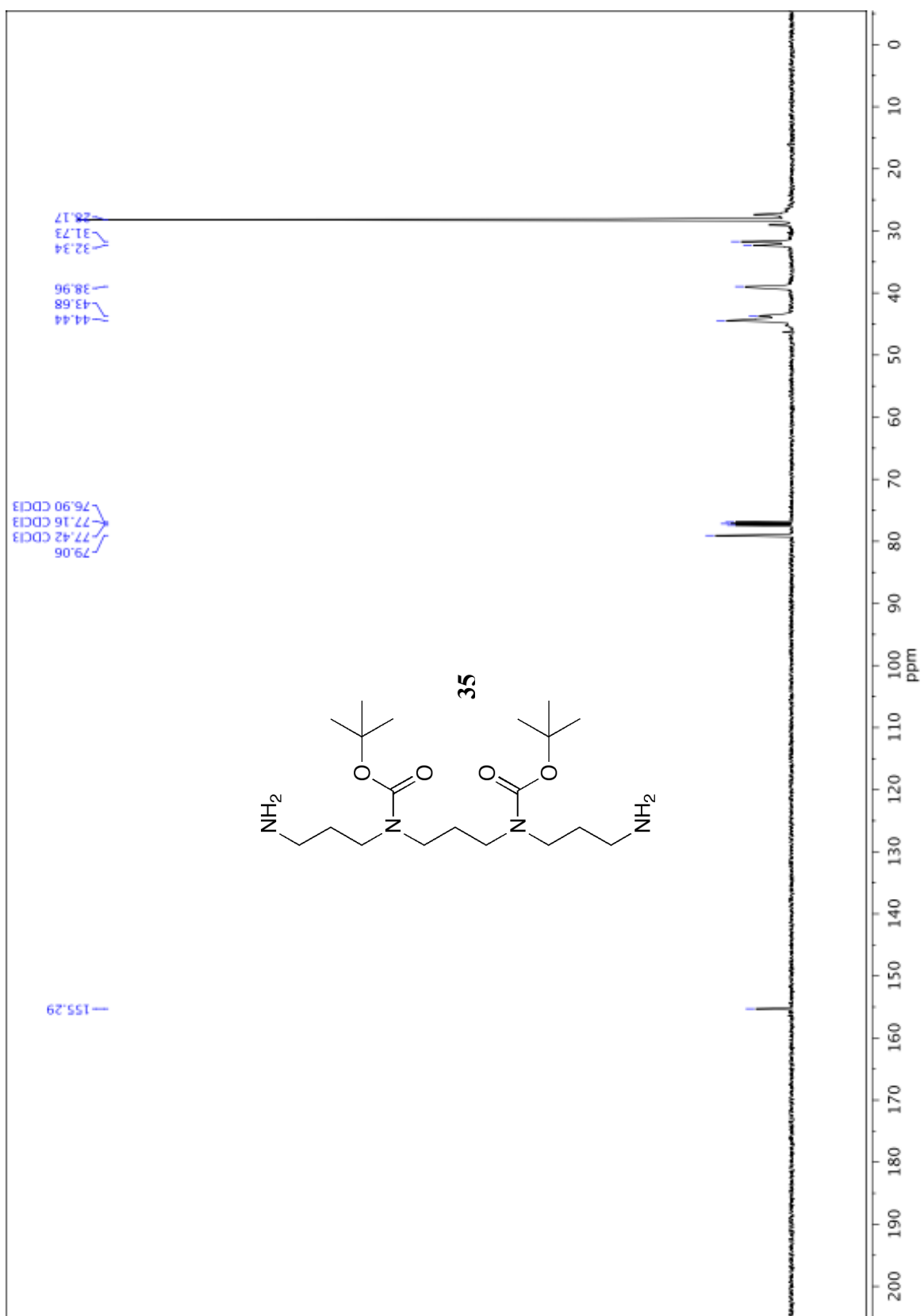








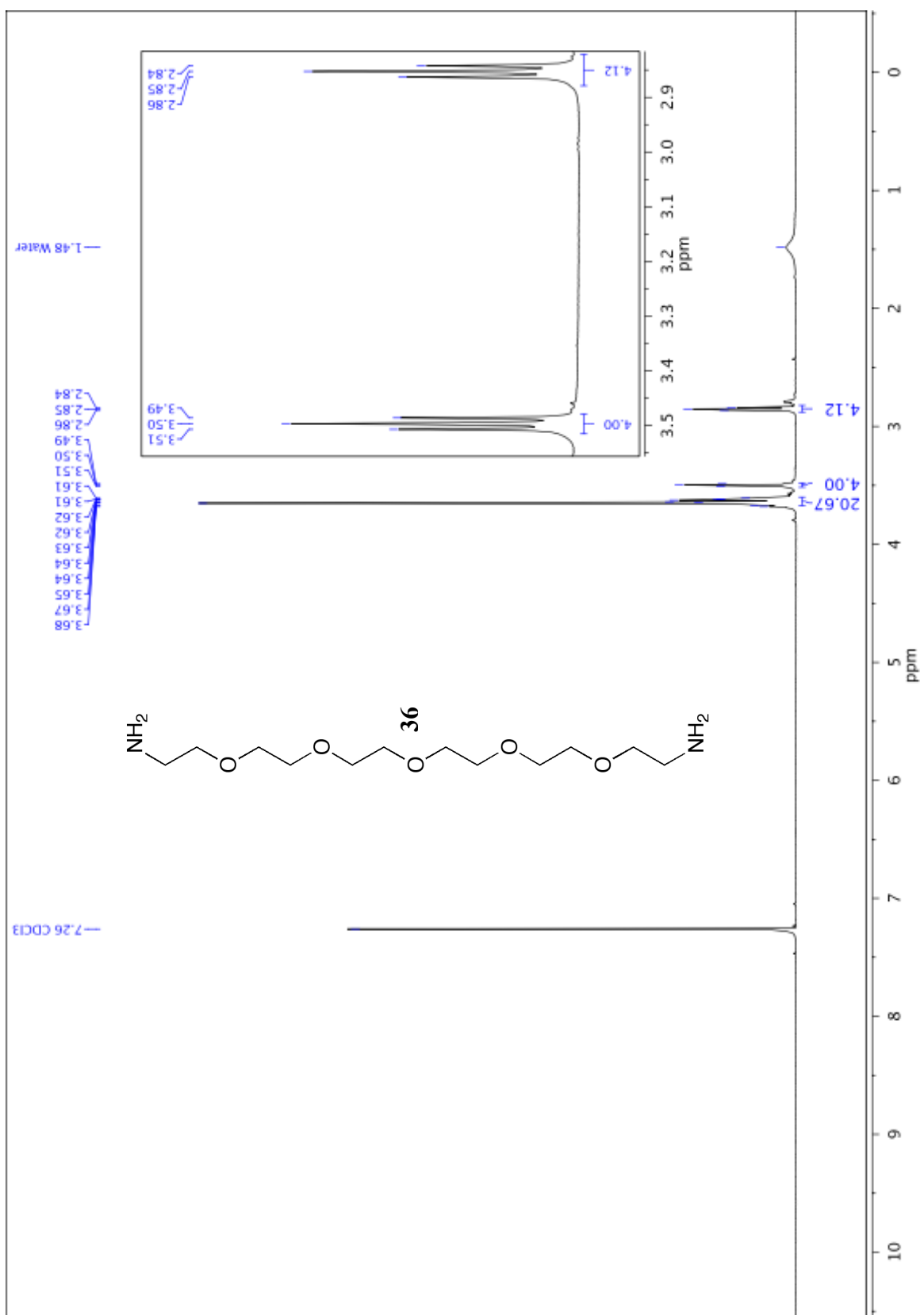




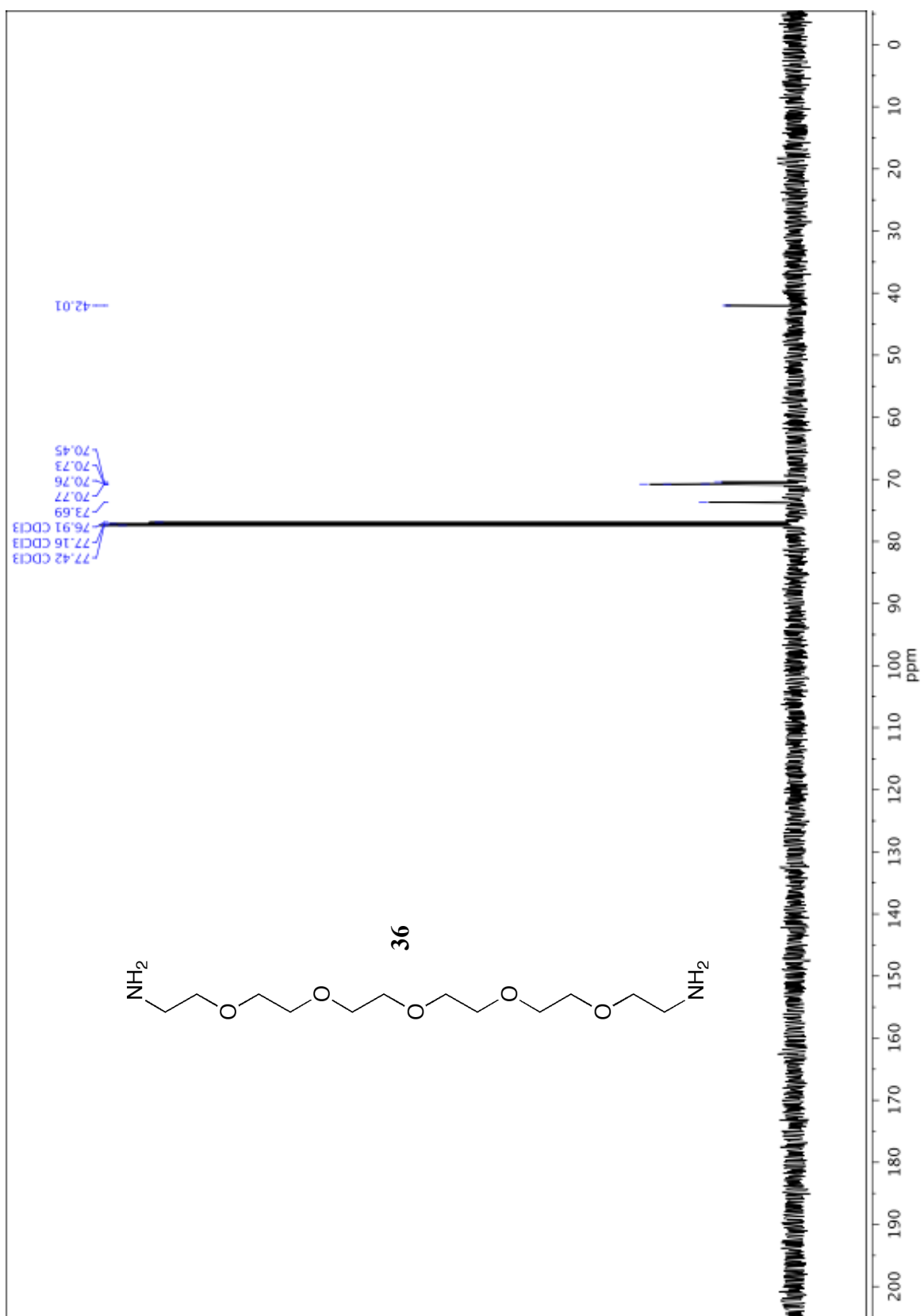


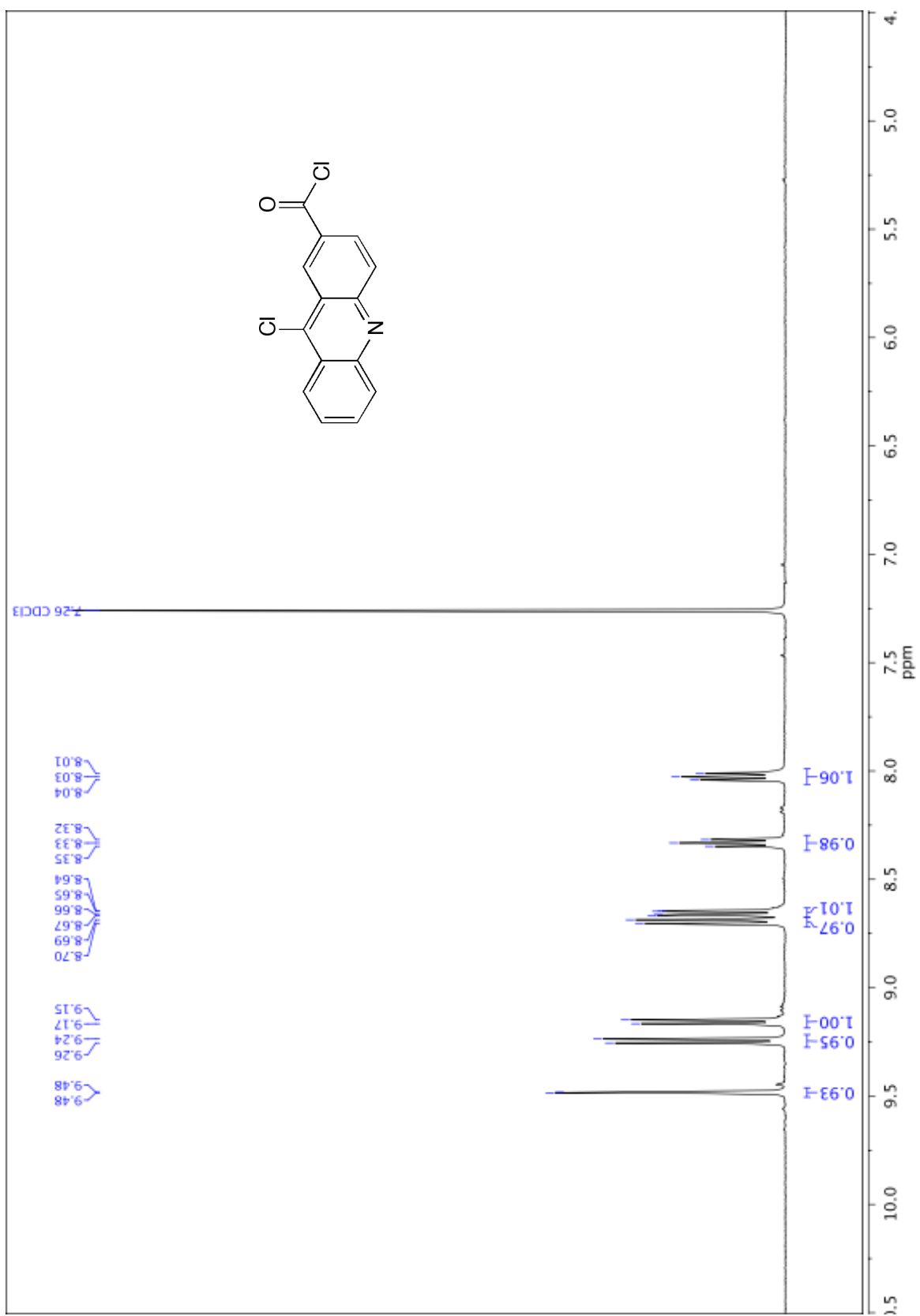


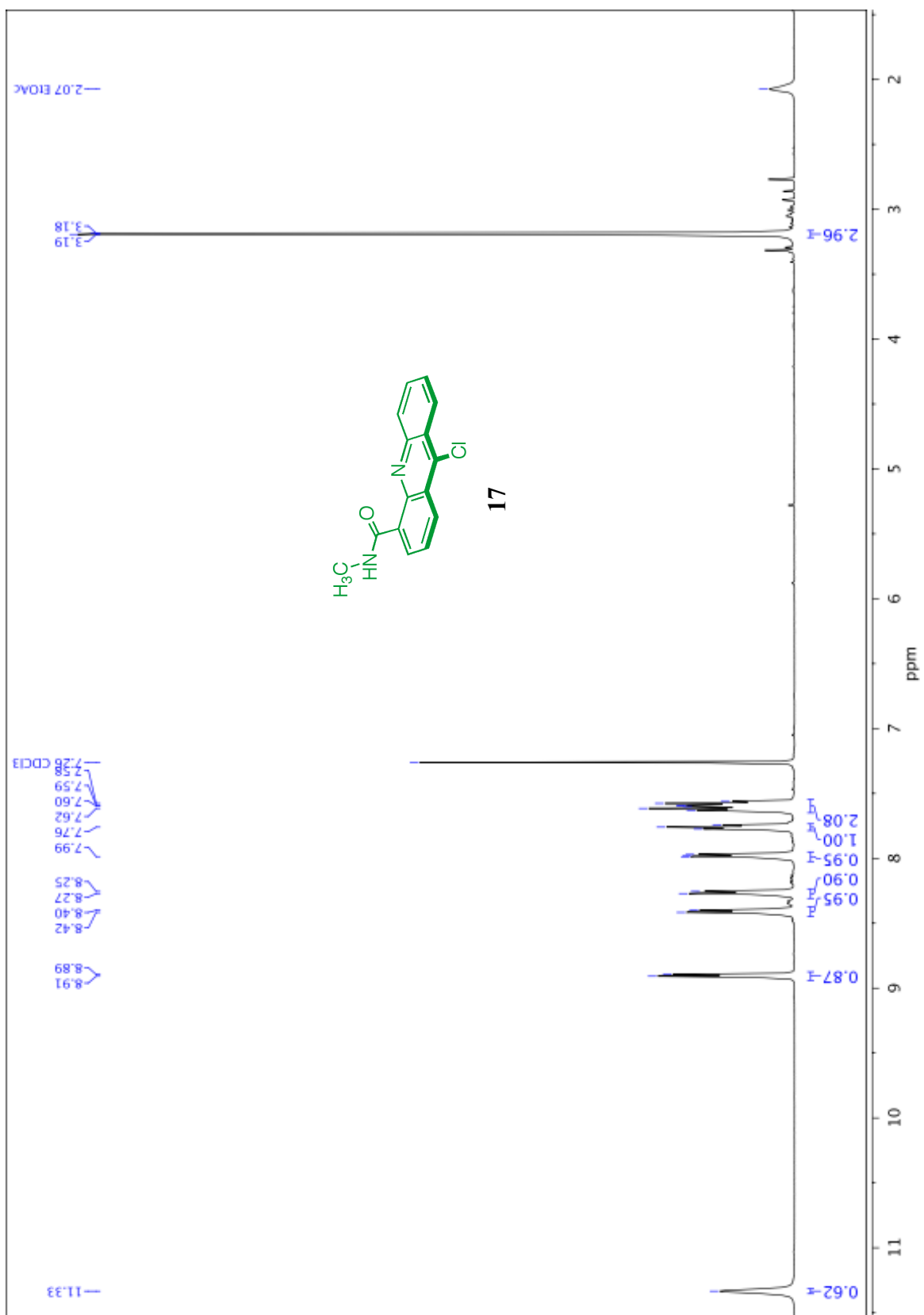


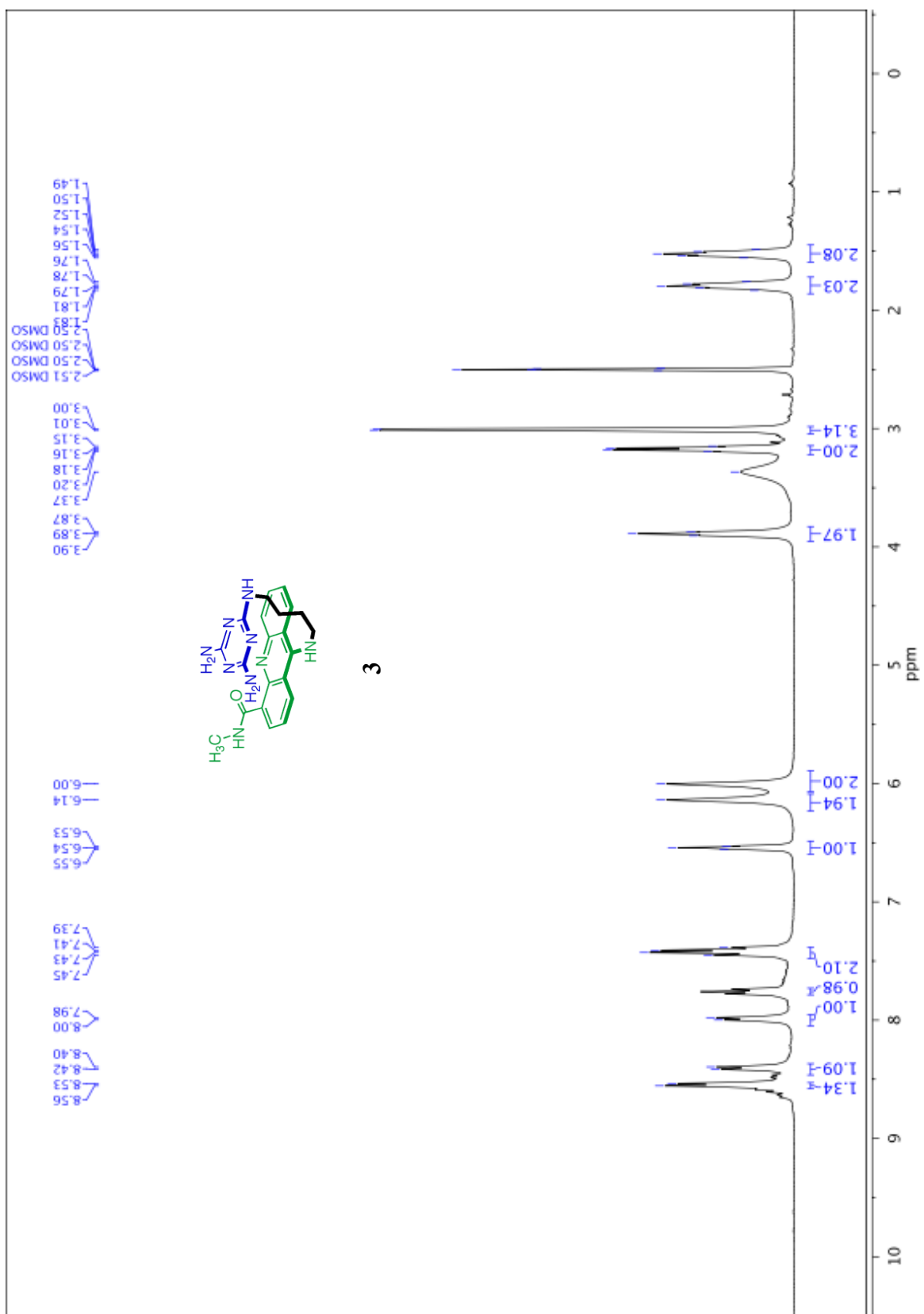


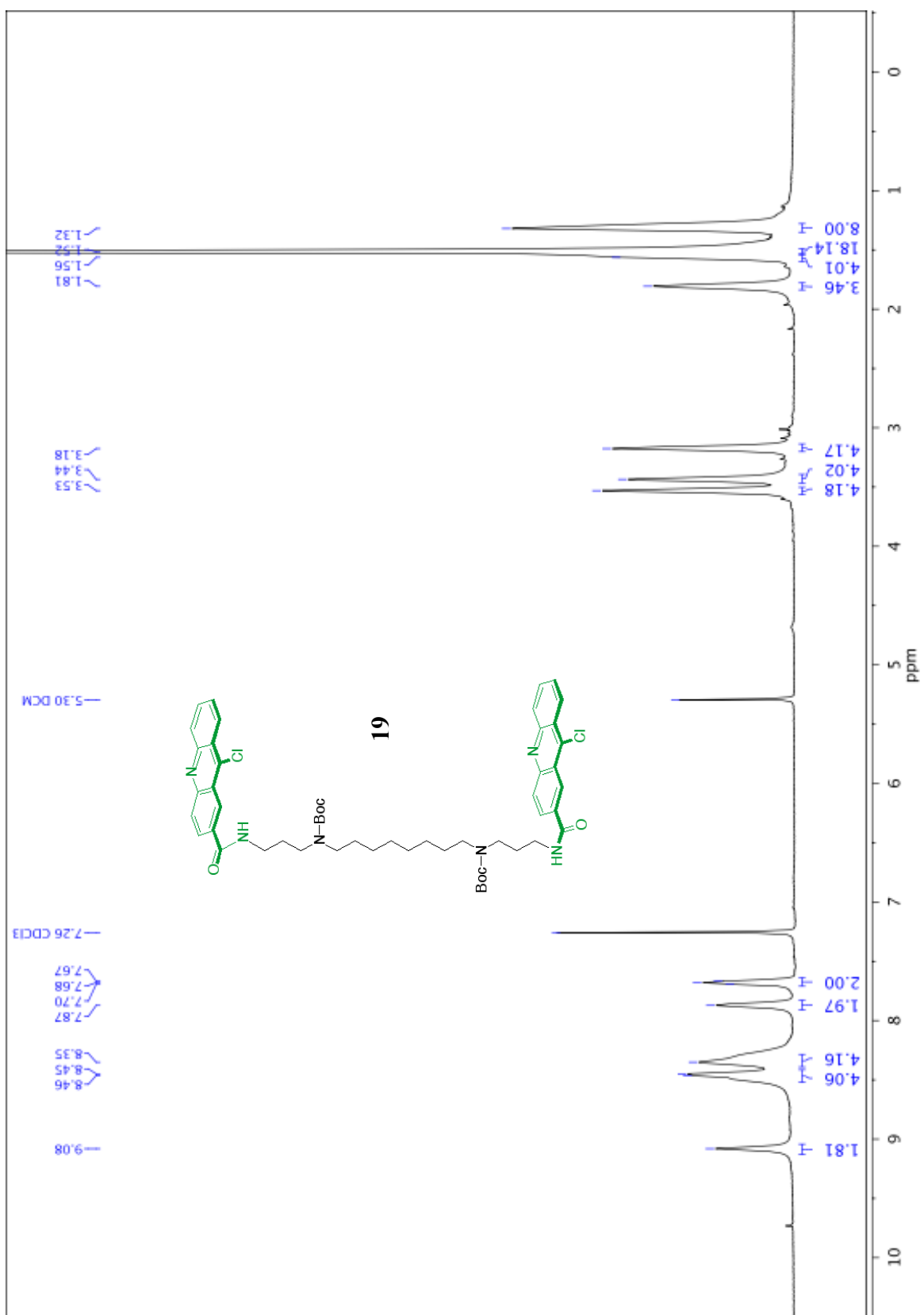




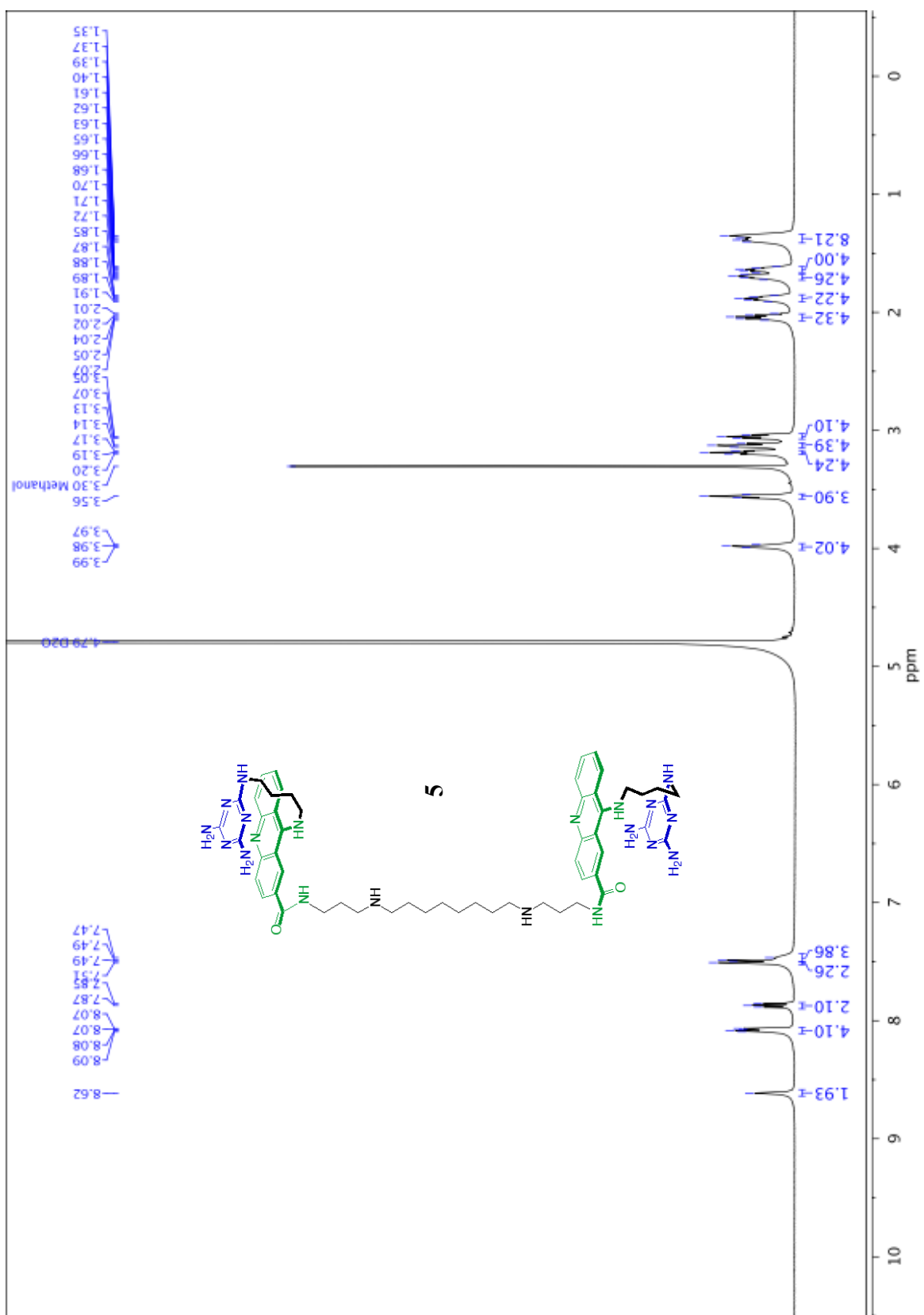






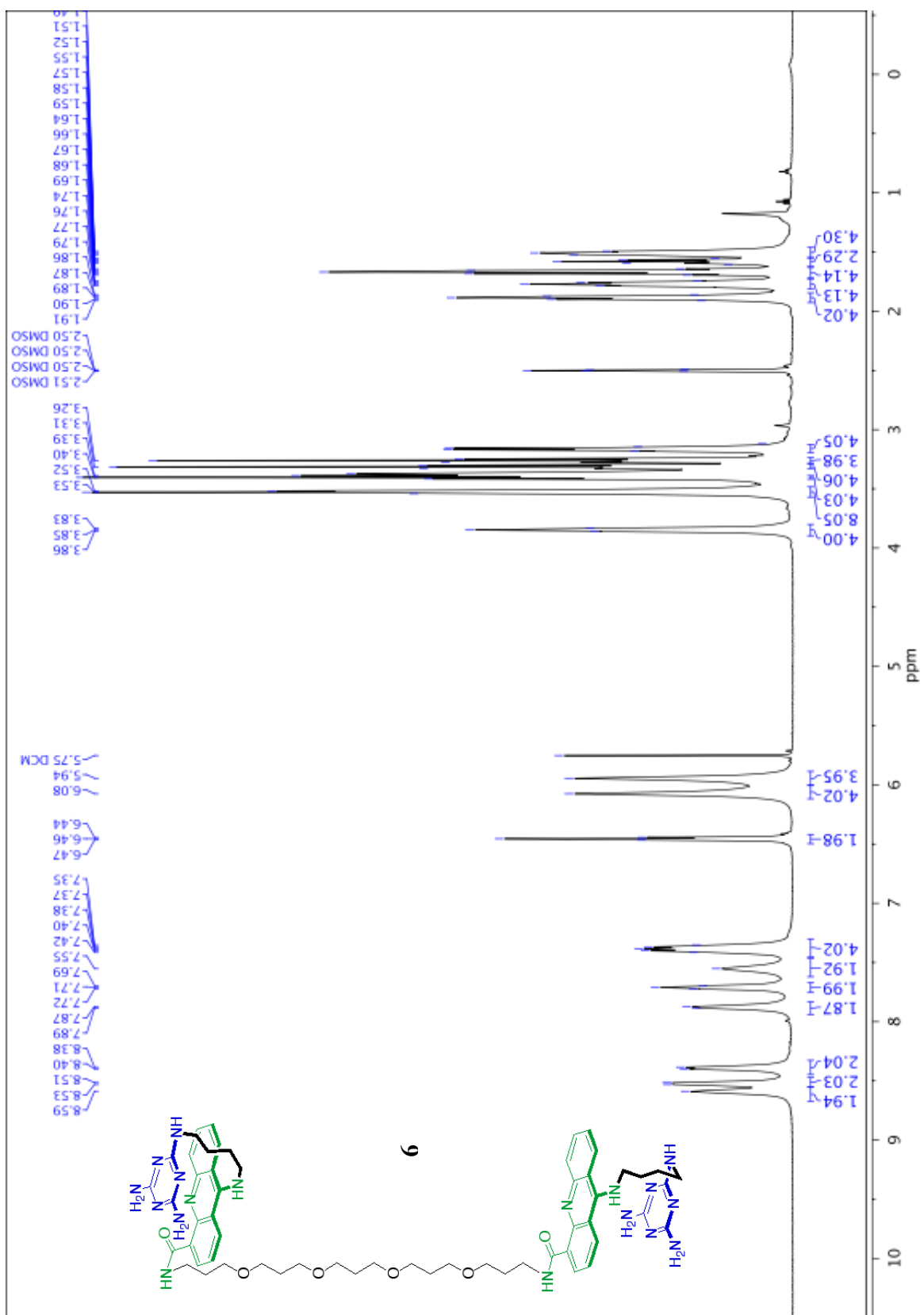




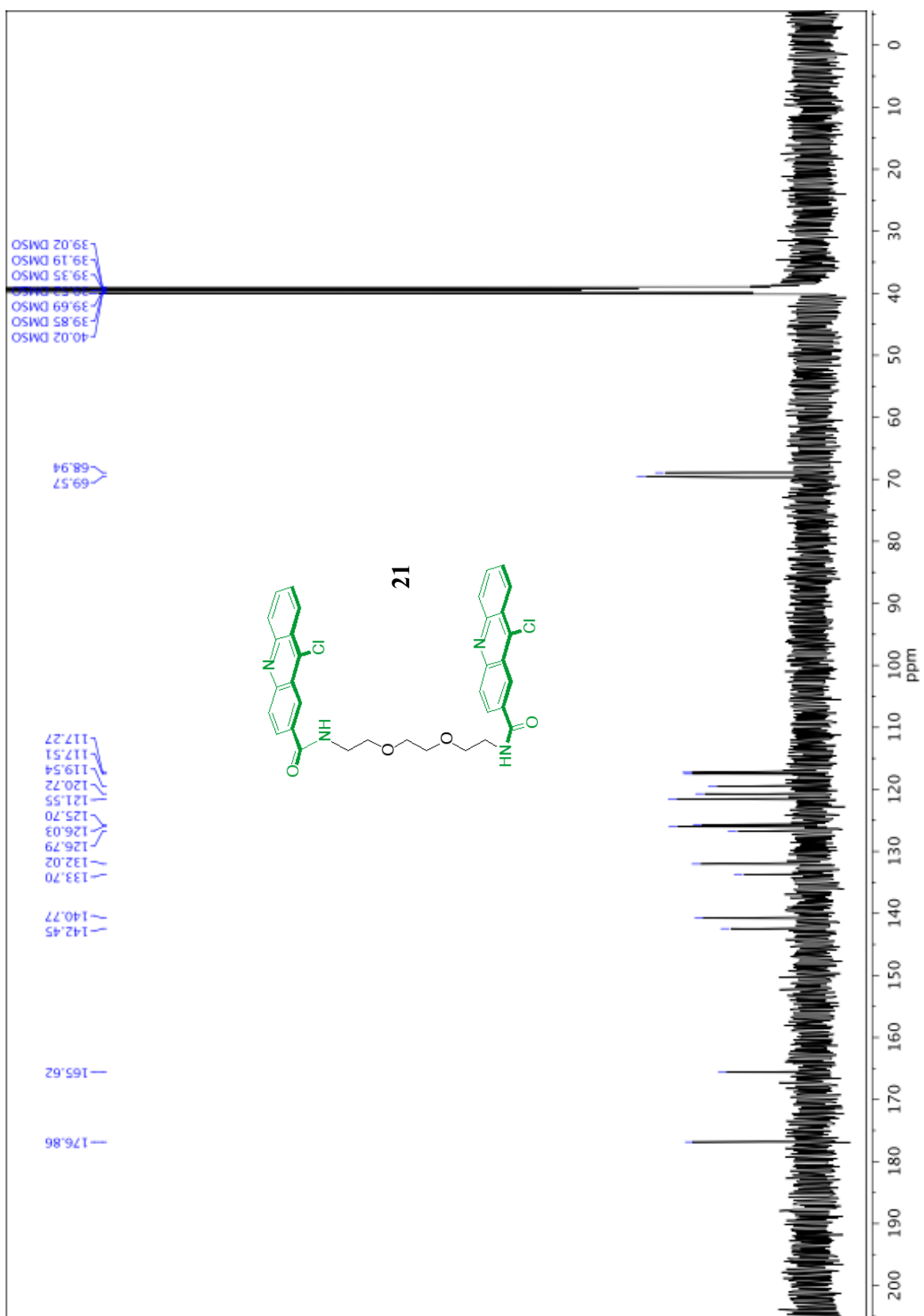






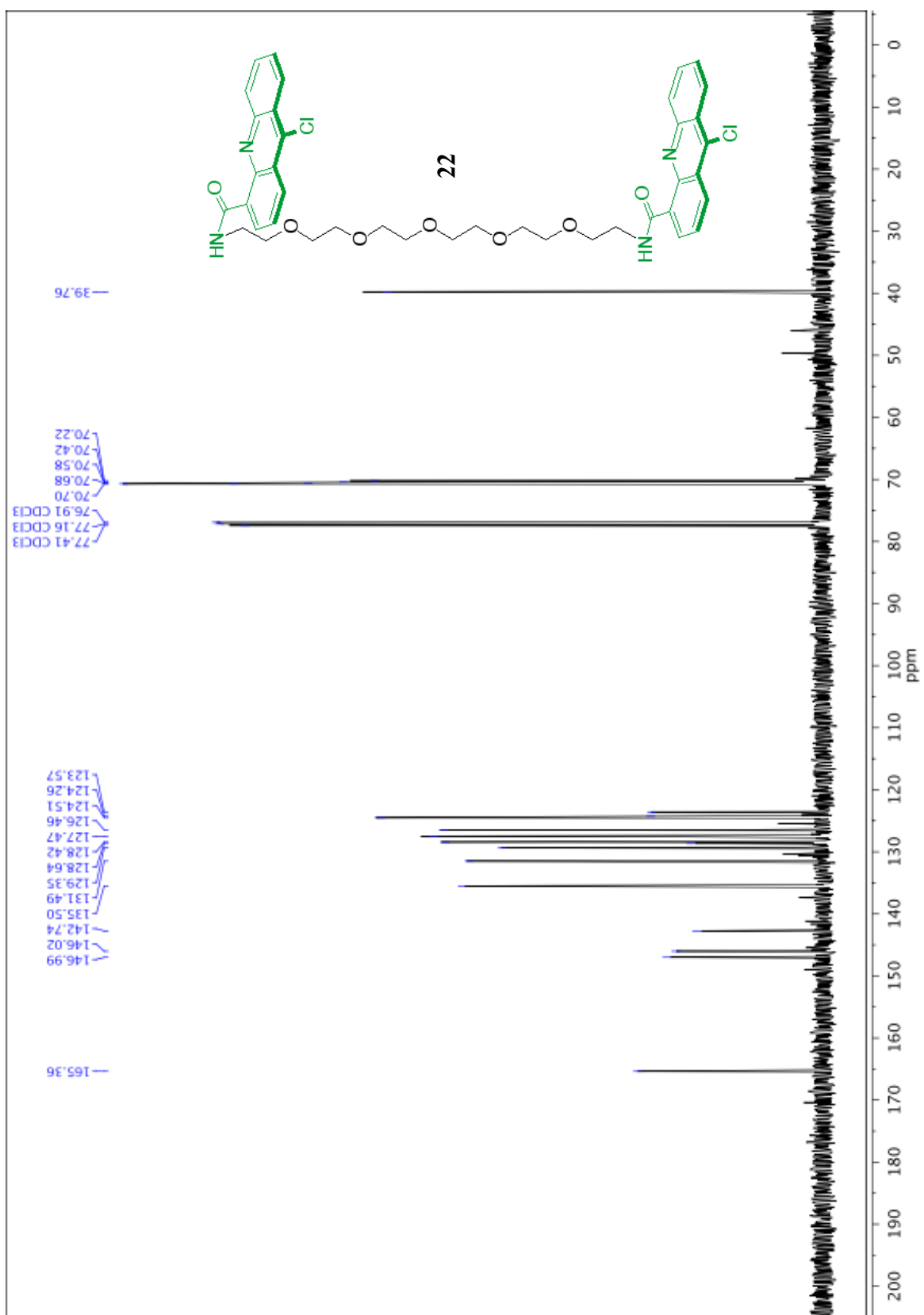




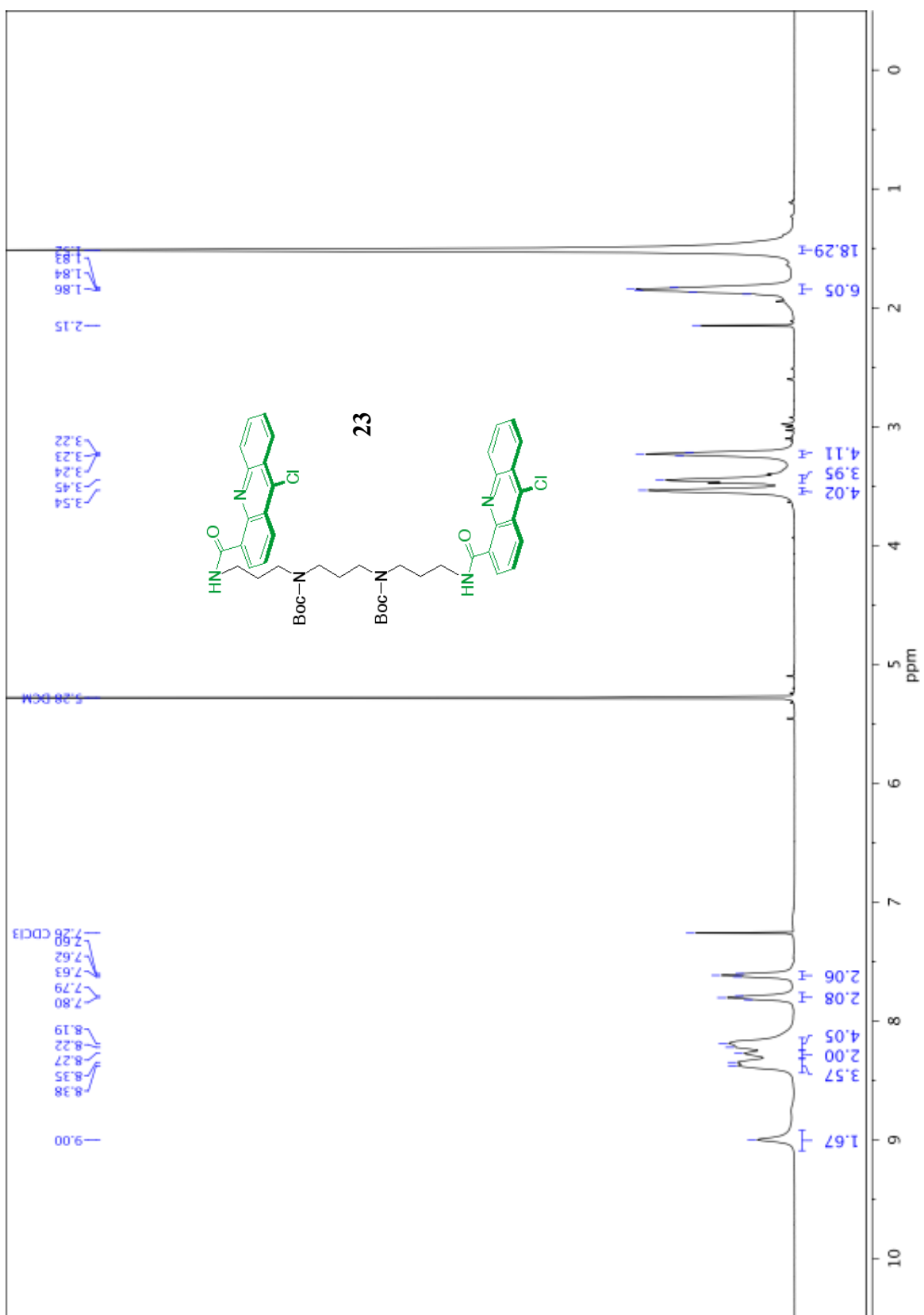




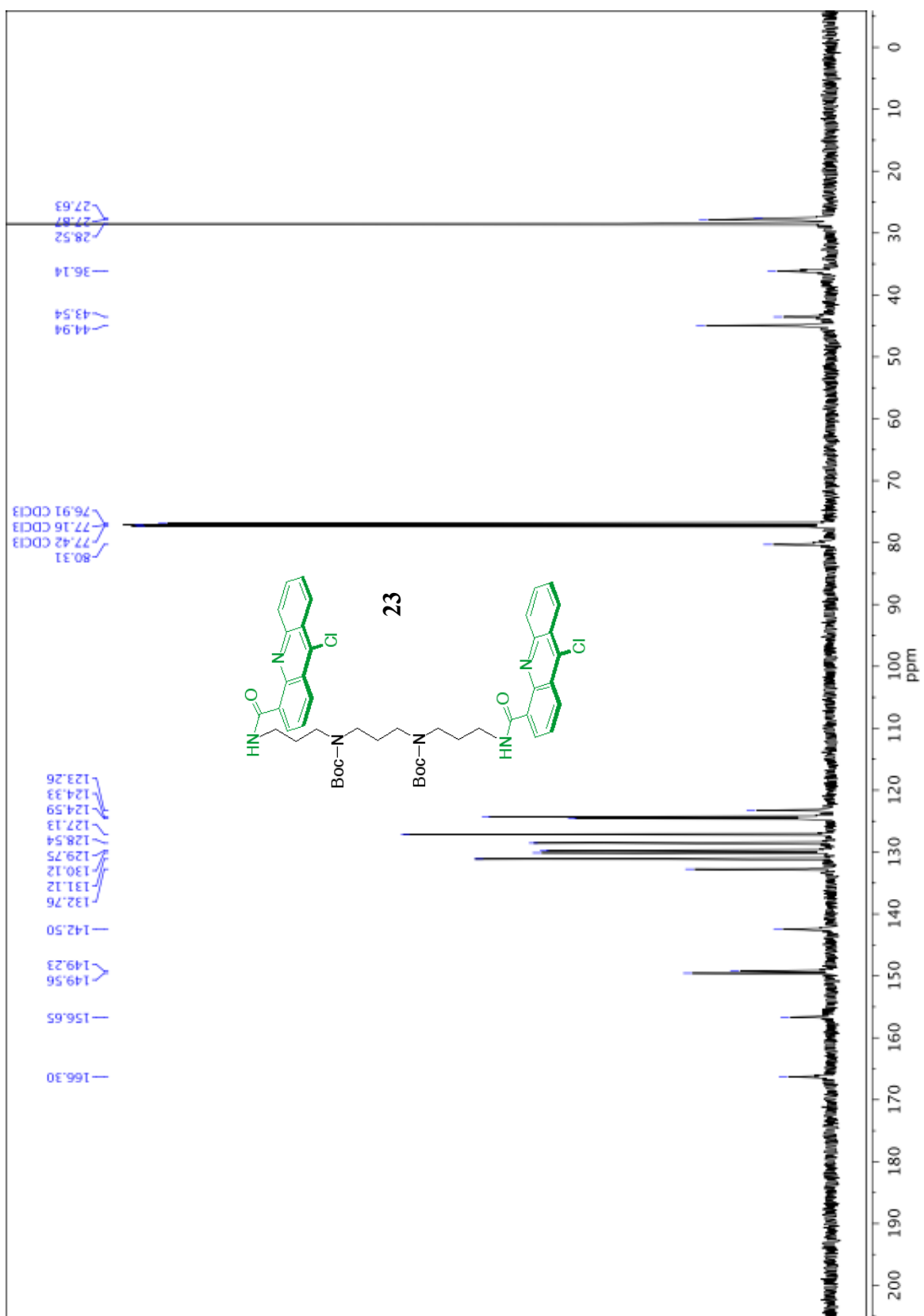






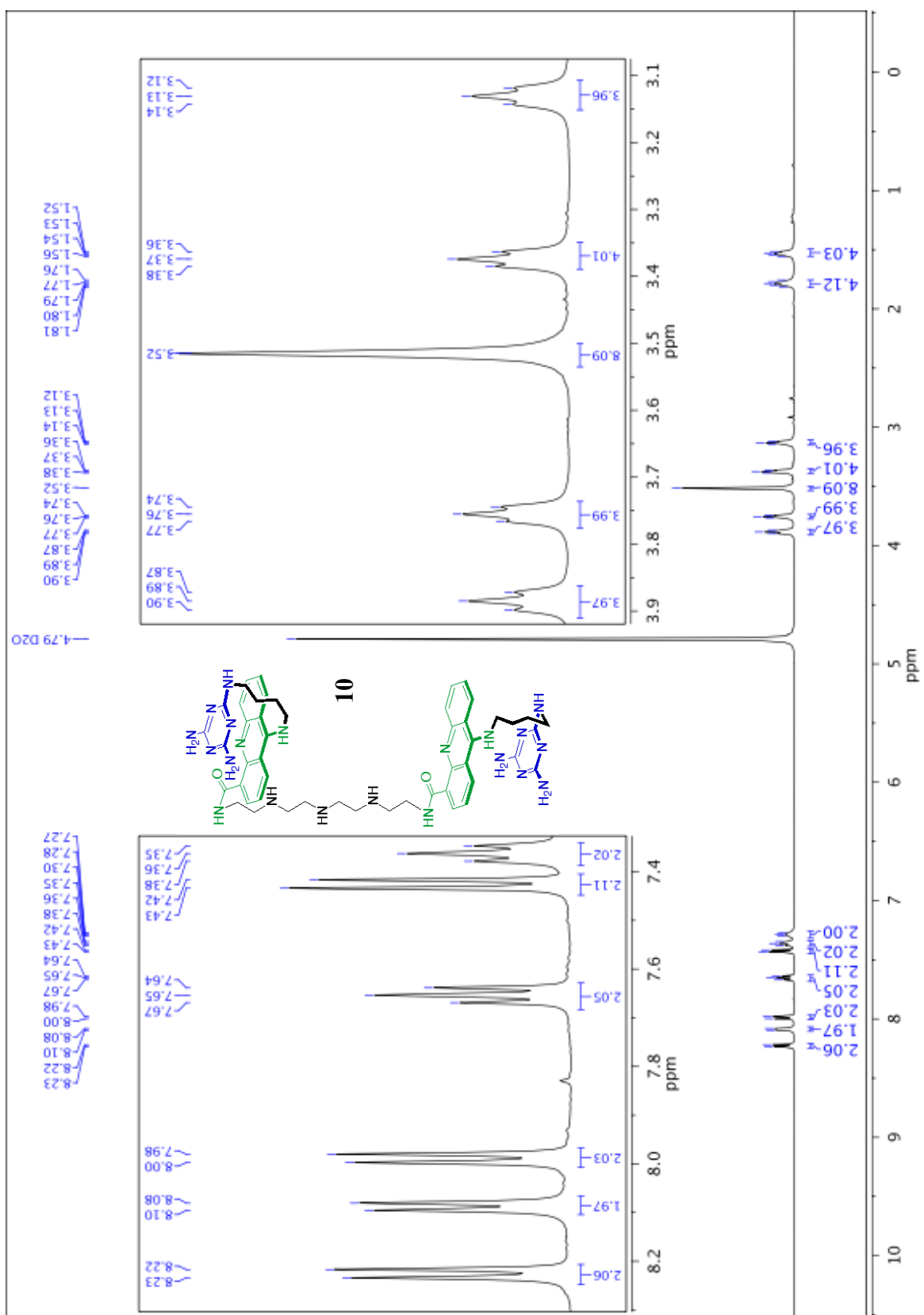


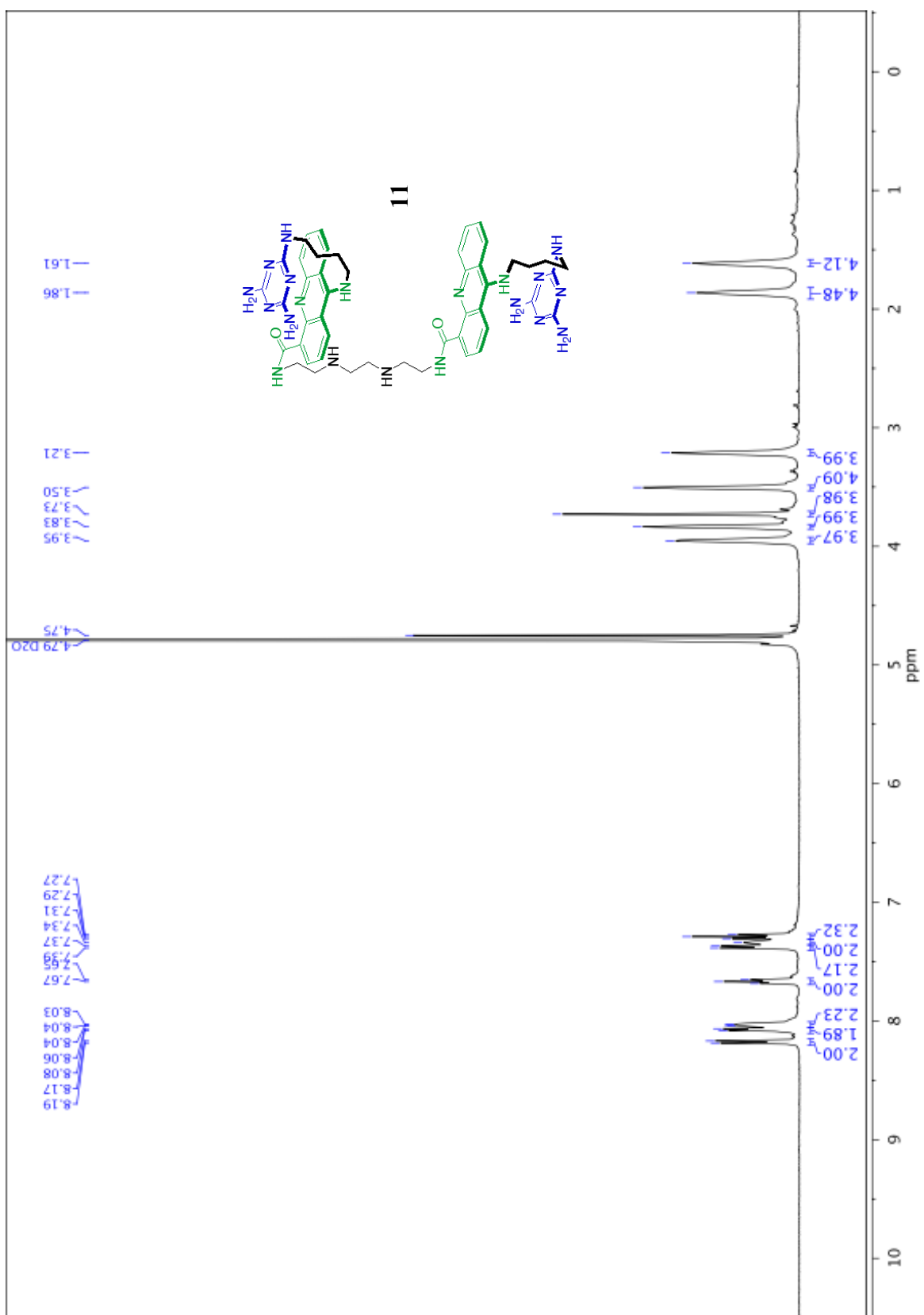


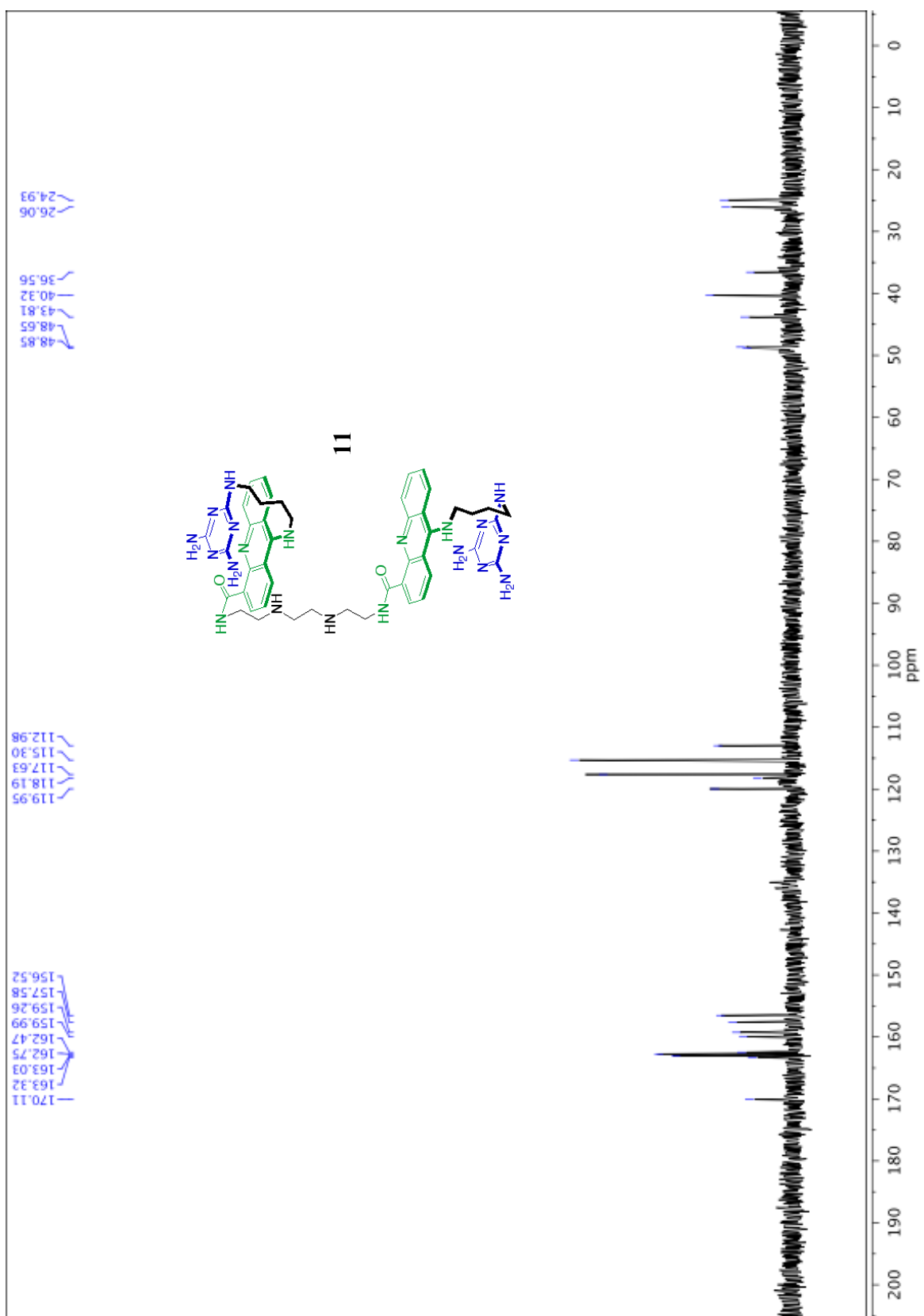


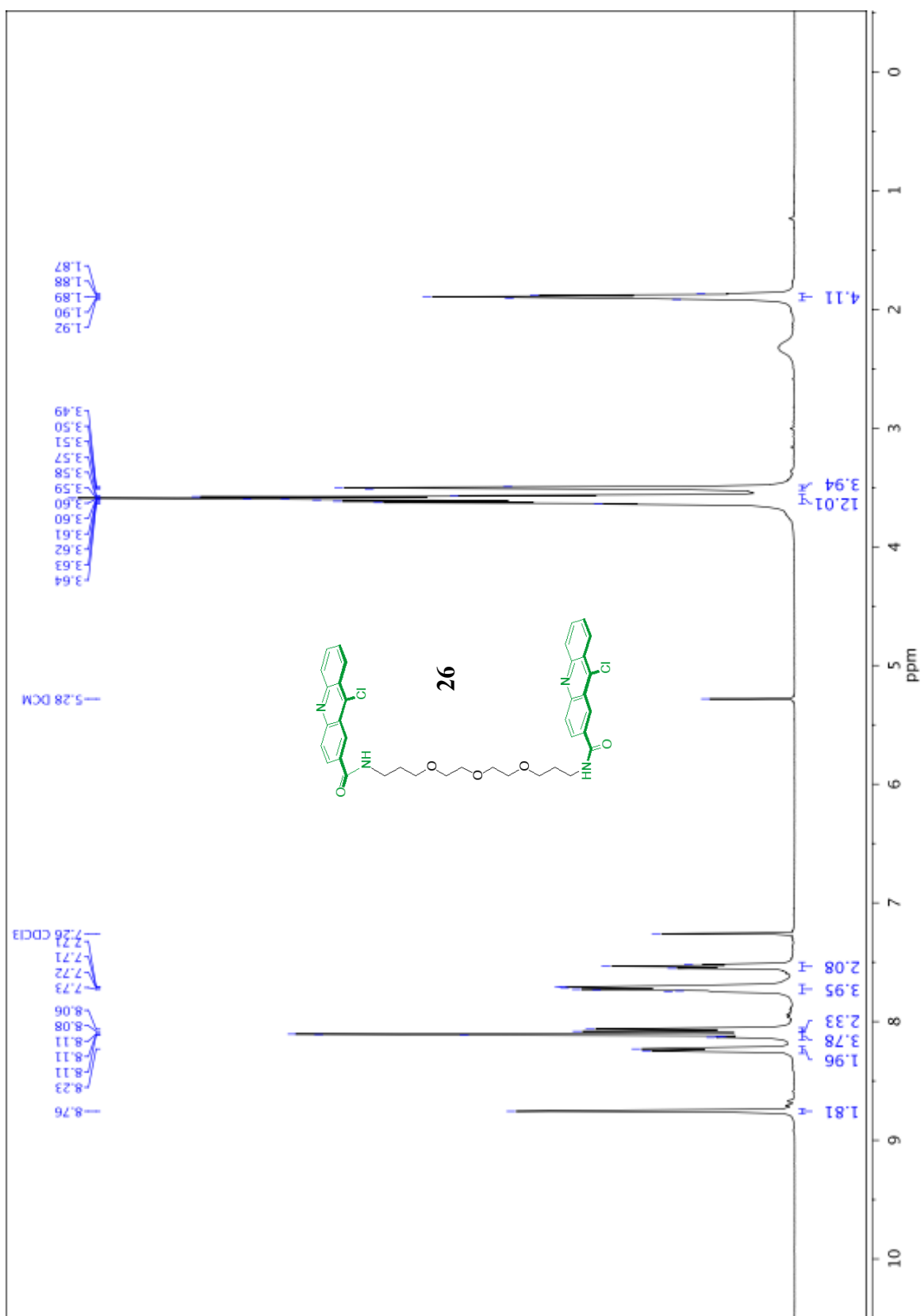


















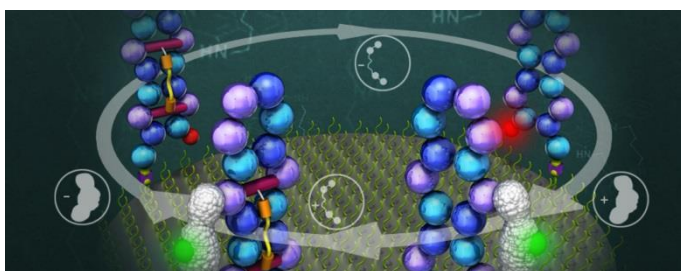


# Chapter 4.<sup>2</sup>

## Single Molecule Study of the CUG Repeat·MBNL1 Interaction and its Inhibition by Small Molecules

### 4.1 Abstract

Effective drug discovery and optimization can be accelerated by techniques capable of deconvoluting the complexities often present in targeted biological systems. We report a single-molecule approach to study the binding of an alternative splicing regulator, muscleblind-like 1 protein (MBNL1), to (CUG)<sub>n=4,6</sub> and the effect of small molecules on this interaction.



Expanded CUG repeats (CUG<sup>exp</sup>) are the causative agent of myotonic dystrophy type 1 by sequestering MBNL1. MBNL1 is able to bind to the (CUG)<sub>n</sub>·inhibitor complex indicating that the inhibition is not a straight forward competitive process. A simple ligand, highly selective for CUG<sup>exp</sup> was used to design a new dimeric ligand that binds to (CUG)<sub>n</sub> almost 50-fold more tightly and is more effective in destabilizing MBNL1·(CUG)<sub>4</sub>. The single-molecule method and the analysis framework might be extended to the study of other biomolecular interactions.

### 4.2 Introduction

Myotonic dystrophy type 1 (DM1) is a triplet-repeat disease originating from a progressive expansion of the CTG repeat (CTG<sup>exp</sup>) in the 3'-untranslated region of the *DMPK* gene.<sup>1</sup> The CTG<sup>exp</sup> produces a toxic RNA transcript (CUG<sup>exp</sup>) that does not exit the nucleus but associates with proteins. One of these proteins, MBNL1, is an important regulator of alternative splicing.<sup>2</sup> Sequestration of MBNL1 in nuclear foci leads to multiple misspliced pre-mRNAs, incorrect protein levels, and ultimately the signs and

---

<sup>2</sup> This Chapter is a manuscript under review at Nucleic Acids Research. It is a collaborative project with Maria Spies lab at University of Iowa.

symptoms of the disease.<sup>3</sup> In a mouse model of DM1 a morpholino antisense oligonucleotide (ASO)<sup>1</sup>, a 2'-O-(2-methoxyethyl) ASO<sup>4</sup> and a D-amino acid hexapeptide (ABP1)<sup>5</sup> reversed this process, rescued the missplicing and reversed the phenotype in mice thereby validating CUG<sup>exp</sup> as a therapeutic target. Because no drugs are currently available to treat DM1, there is intense interest in finding small molecules that may function in a manner similar to the morpholino antisense oligonucleotides, but avoid the limitations inherent in the antisense therapeutic approach.<sup>6</sup> Pentamidine,<sup>7</sup> benzo[g]quinolone heterocycle derivatives,<sup>8</sup> a Hoechst derivative (H1),<sup>9</sup> and modularly assembled Hoechst 33258<sup>10</sup> are examples of bioactive CUG<sup>exp</sup> binders that are able to restore MBNL1 function in DM1 cell and animal models.

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

The lack of a suitable method for studying the MBNL1·CUG<sup>exp</sup> interaction and its inhibition by small molecules means that the overall complexation process remains largely unexplored. The same holds true for many other systems where it is desirable to disrupt off-pathway protein·RNA interactions without affecting the biologically important function of the protein. To aid our drug discovery program and provide a potentially general approach to studying RNA·protein inhibition, we developed a single-molecule method to analyze the MBNL1·CUG<sup>exp</sup>·inhibitor system. The major advantage of our model-independent approach is the ability to detect and measure the individual binding events in real time under equilibrium conditions.<sup>14</sup> Analyzing the distributions of bound

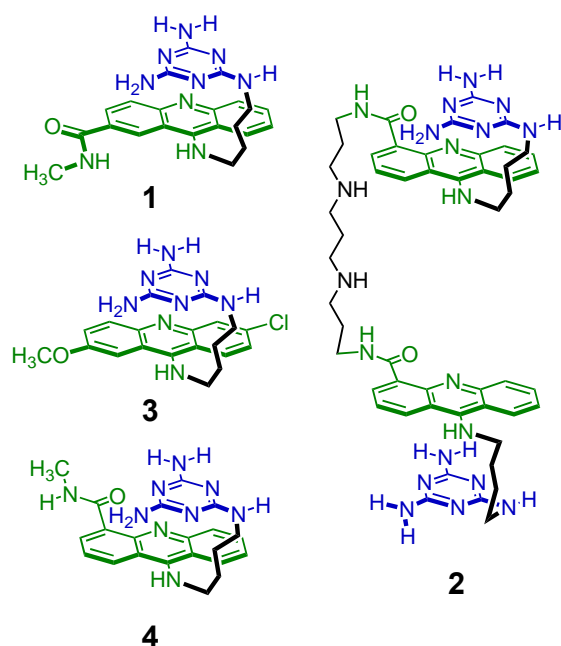
and free states gives the most reliable on and off rates and a full picture of the inhibition process.<sup>15, 16</sup>

Herein, we describe the use of total internal reflection fluorescence (TIRF) microscopy to study, at the single-molecule level, the inhibition of MBNL1 binding to CUG repeat by known small molecule inhibitor **1** (Figure 4.1). The synthesis of dimeric inhibitor ligand **2**, based on structure **1**, is described and evaluated using the TIRF method. Ligand **2** was shown to be bioactive in DM1 cell models (data will be published elsewhere). Both compounds are found to bind RNA at the same time as MBNL1 indicating that the inhibition of MBNL1·CUG<sup>exp</sup> interaction does not follow a simple competitive mechanism. Rather than competitive steric blocking of the MBNL1 complexation site on CUG repeat, it appears the ligands are accelerating the dissociation of the protein.

### 4.3 Results

**Design and Synthesis of Small Molecules.** Our previous studies validated **3** (Figure 4.1) as a selective CUG binder, which was able to inhibit the MBNL1·CUG interaction with an apparent  $IC_{50} = 52 \pm 20 \mu M$ .<sup>17</sup> Because the ultimate RNA target is CUG<sup>exp</sup>, a logical way to increase the selectivity and potency of this lead compound is to utilize the multivalent effect.<sup>18</sup> Disney and coworkers successfully applied this approach to the Hoechst 33258 ligand, which was known to bind CUG with modest selectivity.<sup>19</sup> Because the rationally designed ligand **3** already exhibited a high affinity to (CUG)<sub>4</sub> and showed excellent selectivity, we initially pursued a dimeric ligand.

The covalent linkage of ligand **3** to itself or other compounds required a functional derivative. The three possible sites for covalent modification of **3** are the acridine ring, the triaminotriazine recognition unit or the linking chain. The most expeditious synthetic approach utilized an acridine ring containing a carboxylic acid group to interconnect two ligands. It was found that the chloro- and methoxy-groups in **3** could be replaced with a 2- or 4-carboxamido group (see **1** and **4** in Figure 4.1, respectively) without altering its affinity for CUG or its inhibition of the MBNL1·CUG complex. Limited by the absence of structural data on the ligand·CUG<sup>exp</sup> complex, we designed and synthesized a small library of dimeric ligands that could potentially bind to consecutive CUG sites. This study, which will be published elsewhere, led us to **2** as the optimized dimeric ligand used in the current study (Figure 4.1).



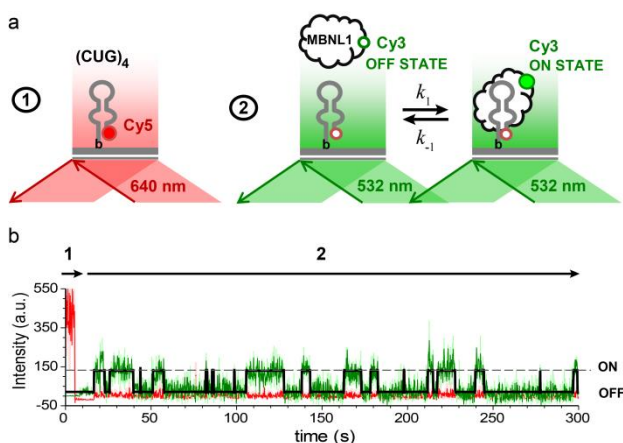
**Figure 4.1.** Compounds 1-4 used in this study.

Ligand **2** was designed to target two consecutive UU mismatches, with the *N,N'*-bis(3-aminopropyl)-1,3-propanediamine linker spanning two central GC base-pairs. Although the binding mode was not firmly established for **3**, it was designed to act as a “stacked intercalator” with the acridine and triaminotriazine rings  $\pi$ -stacked while the intercalator sits between the GC base pair and the U-triaminotriazine-U base triplet.<sup>17</sup> A recent combined experimental and computational study provides support for that binding model in the major groove of the CUG RNA.<sup>20</sup> By positioning the linking chain and triaminotriazine recognition units on opposite sides of the acridine unit, the bivalent complex likely requires a threading mechanism for binding.<sup>21</sup> This design element was intentional with the goal of increasing the binding affinity through a higher residence time.<sup>22, 23</sup>

**Establishing a TIRFM-based Assay to Monitor the MBNL1·(CUG)<sub>n</sub> Interaction.**<sup>3</sup> To investigate binding of MBNL1 to (CUG)<sub>n</sub> and the effect of **1** and our bioactive dimeric ligand, **2**, on MBNL1·(CUG)<sub>n</sub> interaction, we developed a TIRFM-based single-molecule binding assay.<sup>28, 29</sup> A stem-loop containing two CUG pairs separated by a tetraloop, (CUG)<sub>4</sub>, was used as an RNA substrate to analyze the MBNL1 binding. This RNA construct has been previously established as the shortest CUG repeat to which MBNL1

<sup>3</sup> The single molecule study has been performed in Maria Spies lab at University of Illinois (2011-12) and University of Iowa (2012-13).

binds with a similar affinity as to the long CUG repeats.<sup>30</sup> The predicted stem-loop structure of (CUG)<sub>4</sub> has been confirmed by melting studies ( $T_m = 60.6 \pm 0.6$ ).<sup>30, 31</sup> Therefore, (CUG)<sub>4</sub> is the shortest validated CUG repeat model. Moreover, it contains only one MBNL1 binding site per RNA molecule, which simplifies analysis of the association/dissociation events. Validation of the single site binding is described in the next section below.

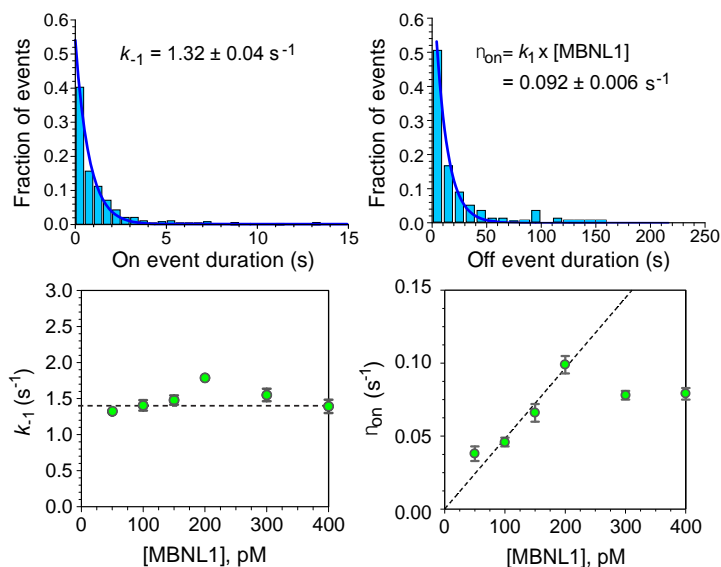


**Figure 4.2.** Single-molecule analysis of MBNL1 binding to (CUG)<sub>4</sub>. a. (CUG)<sub>4</sub> was immobilized on a PEG-coated surface of the slide *via* biotin-neutravidin interaction. The experiment was carried out in two steps. (1) Positions of (CUG)<sub>4</sub> molecules on the slide were triangulated by exciting Cy5 label of (CUG)<sub>4</sub> with a red laser. (2) MBNL1·(CUG)<sub>4</sub> interaction was monitored using green TIR illumination. Upon binding to (CUG)<sub>4</sub>, Cy3 label on MBNL1 is excited as it is sequestered within the evanescent field. b. Representative Cy3 (green) and Cy5 (red) fluorescence intensity trajectories from an individual (CUG)<sub>4</sub> molecule. Arrows 1 and 2 depict periods of red and green excitation, respectively. ON and OFF events are indicated on the right.

We followed the binding of Cy3 labeled MBNL1 to, and its dissociation from, individual (CUG)<sub>4</sub> RNA molecules. Biotinylated and Cy5-labeled (CUG)<sub>4</sub> was immobilized on a PEG-coated surface of the TIRFM flow chamber *via* biotin-neutravidin interaction (Figure 4.2a, left). First, locations of the surface-tethered (CUG)<sub>4</sub> molecules were identified by illuminating the slide with red laser (641 nm), which selectively excited the Cy5 fluorophore present on (CUG)<sub>4</sub>. Then, the laser was switched to green (532 nm) to monitor binding of Cy3 labeled MBNL1 to (CUG)<sub>4</sub> (Figure 4.2a, right). We followed MBNL1·(CUG)<sub>4</sub> interaction in real time, with a time resolution of 100 ms. In this experimental scheme, the surface-tethered (CUG)<sub>4</sub> and Cy3 labeled MBNL1 are at equilibrium during the length of the measurement. TIR-generated evanescent field excites only the Cy3 labeled MBNL1 molecules that reside near the surface longer than 2 frames of the camera, which only occurs when MBNL1 is bound to the surface-

tethered (CUG)<sub>4</sub>. This mode of excitation eliminates the background fluorescence of Cy3-MBNL1 in solution and thereby allows monitoring MBNL1·(CUG)<sub>4</sub> interactions over a broad range of MBNL1 concentrations.

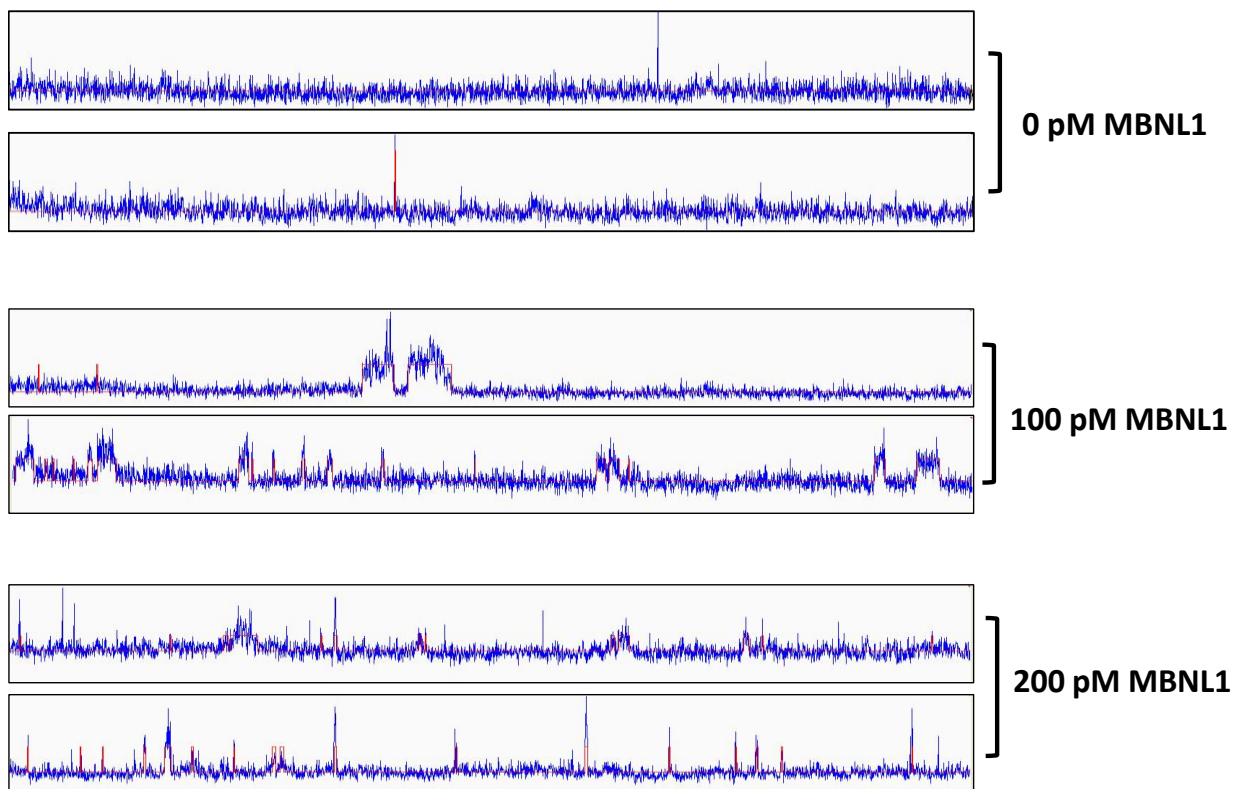
Each event of MBNL1 binding to the surface-tethered (CUG)<sub>4</sub> was observed as the appearance of a Cy3 signal at the location where (CUG)<sub>4</sub> resided. Conversely, dissociation of MBNL1 from (CUG)<sub>4</sub> resulted in the disappearance of the Cy3 signal. Fluorescence trajectories (Figure 4.2b) showed multiple events of two-state (ON and OFF) association and dissociation. The presence of only two states indicates that each observed event corresponded to binding of a single MBNL1 to a single (CUG)<sub>4</sub>. To confirm that potential non-specific interaction of MBNL1 with the surface does not interfere with our analysis, we substituted Cy5-labeled (CUG)<sub>4</sub> with Cy5-labeled streptavidin. Only a few fluorescence trajectories showing both Cy5 and Cy3 signals were observed in the presence of 300 pM Cy3-MBNL1 (about 5% of what was observed in the presence of Cy5-labeled (CUG)<sub>4</sub>) confirming that the vast majority of binding events described above resulted from specific MBNL1·(CUG)<sub>4</sub> interaction. Analysis of the Cy3-MBNL1·(CUG)<sub>4</sub> interaction in the presence of unlabeled MBNL1 confirmed that MBNL1 binding properties were not affected by the Cy3 conjugation (Figure 4.11).



**Figure 4.3.** a. Distributions of all ON (left) and OFF (right) event durations detected in the 50 trajectories of individual (CUG)<sub>4</sub> molecules are fit to single exponentials to yield  $k_{-1}$  ([MBNL1] = 50 pM) and  $v_{on}$  ([MBNL1] = 200 pM), respectively. b. Effect of MBNL1 concentration on the  $k_{-1}$  and  $v_{on}$  of the observed events is shown.  $k_{-1}$  (dissociation rate constant) was independent of [MBNL1], suggesting the absence of functional cooperation between monomers. Initially,  $v_{on}$  (association rate) increased with increasing [MBNL1]. It was saturated at the MBNL1 concentrations above 200 pM.



The Cy3 (green laser excitation) regions of the trajectories were fit to a two-state model using QuB software,<sup>32</sup> yielding dwell times in the OFF (dissociated) and ON (bound) states (Figure 4.4). We determined kinetic parameters of MBNL1 binding to the surface-tethered (CUG)<sub>4</sub> by globally analyzing distributions of ON and OFF dwell times in the presence of different concentrations of MBNL1 protein. All protein concentrations (50, 100, 150, 200, 300 and 400 pM) were used to determine the dissociation rate constant, whereas the association rate constant was determined using only 50 – 200 pM range of concentrations, which gave a linear dependence of the observed association rate on MBNL1 concentration.

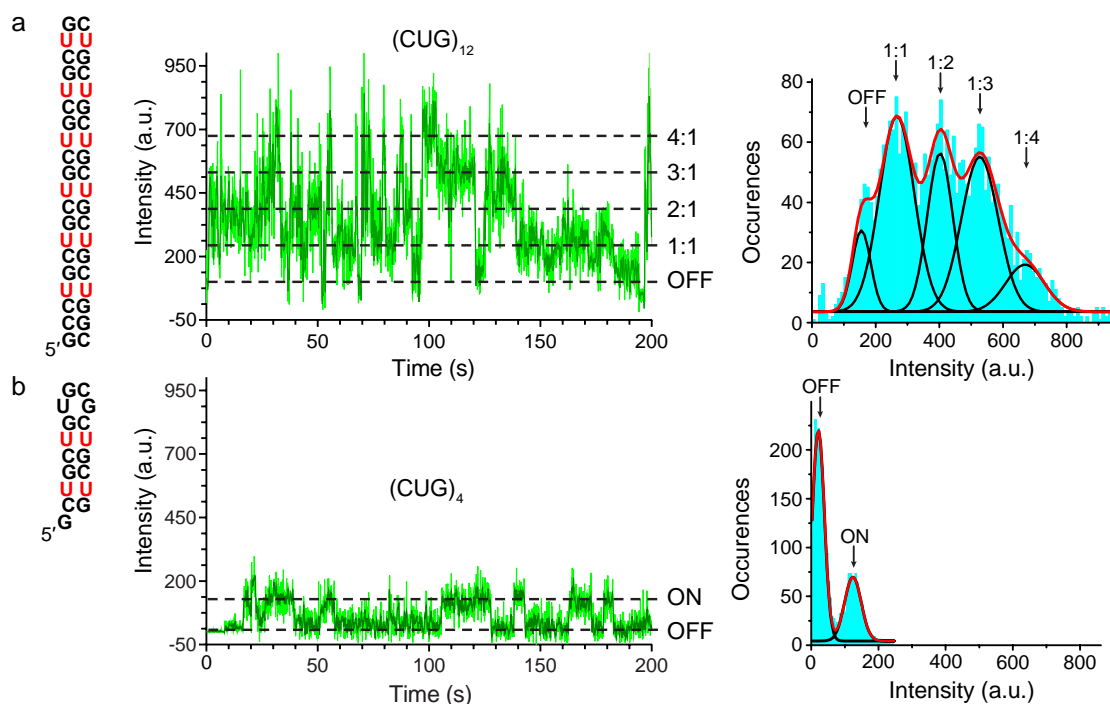


**Figure 4.4.** Representative trajectories with two-state fitting. Single-molecule trajectories were fit using QuB to reveal the binding and dissociation of MBNL1 to (CUG)<sub>4</sub>. The data was collected in the standard buffer in the presence of 0 pM, 100 pM and 200 pM MBNL and were globally fit by HMM using QuB software. For representative trajectories fit idealized states (red line) are shown superimposed on the actual Cy3 intensity trajectories (light blue). The trajectories in which (CUG)<sub>4</sub> molecules showed either abnormally high Cy5 intensity or more than two-step bleaching were omitted from analysis.

Figure 4.3a and 4.3b show individually fit ON and OFF time distributions obtained in the presence of 50 pM and 200 pM MBNL1, respectively. Global analyses of the dwell distributions yielded the dissociation rate constant for MBNL1·(CUG)<sub>4</sub> complex,  $k_{-1} =$

$1.44 \pm 0.3 \text{ s}^{-1}$ , and association rate constant,  $k_1 = (4.6 \pm 0.2) \times 10^8 \text{ s}^{-1} \text{ M}^{-1}$ , from which we calculated the equilibrium dissociation constant,  $K_{D1} = k_{-1}/k_1 = 3.1 \pm 0.1 \text{ nM}$ . The system was set up by Dr. Masayoshi Honda, a postdoctoral researcher in Maria Spies lab.

**Stoichiometry of the MBNL1·(CUG)<sub>n</sub> Interaction.** Biotinylated (CUG)<sub>12</sub> can simultaneously accommodate multiple Cy3 labeled MBNL1 molecules. Its Cy3 trajectories contained 5 states characterized by different Cy3 intensities (Figure 4.5a). An intensity histogram showed that there are three major Cy3 intensity populations consistent with the binding of one (around 196 a.u.), two (around 332 a.u.), and three (around 458 a.u.) MBNL1 proteins to the (CUG)<sub>12</sub>. A small fourth peak (around 632 a.u.) may represent more than one Cy3 label present on a fraction of MBNL1 molecules (Cy3 labeling efficiency of MBNL1 was 110%). Stoichiometry determination assay using SPR confirmed a stoichiometric ratio of MBNL1:(CUG)<sub>12</sub> close to 3:1 (Figure 4.12).



**Figure 4.5.** Each (CUG)<sub>4</sub> can accommodate a single MBNL1 protein, whereas three MBNL1 proteins can be simultaneously bound to (CUG)<sub>12</sub>. **a.** Schematic representation of (CUG)<sub>12</sub> (left), **b.** Schematic representation of (CUG)<sub>4</sub>. Representative fluorescence intensity trajectories for (MBNL1)<sub>3</sub>·(CUG)<sub>12</sub> (a) and MBNL1·(CUG)<sub>4</sub> (b) are shown in the middle (light green line shows the raw data collected at 100 ms time resolution, dark green line shows intensity averaged for 9 data points). The fluorescence intensity histogram compiled from the trajectory is shown on the right (aqua) and is fit to 5 Gaussian peaks (black lines), corresponding to stepwise increase in (CUG)<sub>12</sub> occupancy by MBNL1 (a) whereas only a single 1:1 binding state is observed for MBNL1·(CUG)<sub>4</sub> (b).

### **Revealing the Mechanism of the MBNL1·(CUG)<sub>n</sub> Interaction Inhibition by **1** and **2**.**

We then carried out MBNL1·(CUG)<sub>4</sub> binding studies in the presence of increasing concentrations of **1** and **2**. Before starting the recording, the Cy3 labeled MBNL1 and the ligands were incubated for at least 5 minutes in the TIRFM reaction chamber to ensure that all components of the system are at equilibrium. The difference in the dwell time distributions for the MBNL1·(CUG)<sub>4</sub> complex in the absence and presence of **1** and **2**, and therefore the difference in the apparent  $k_{off}$  of the MBNL1·(CUG)<sub>4</sub> complex, suggested that both **1** and **2** do not act as simple competitive inhibitors, which should only affect the on-rate (Figures 4.6-4.9).

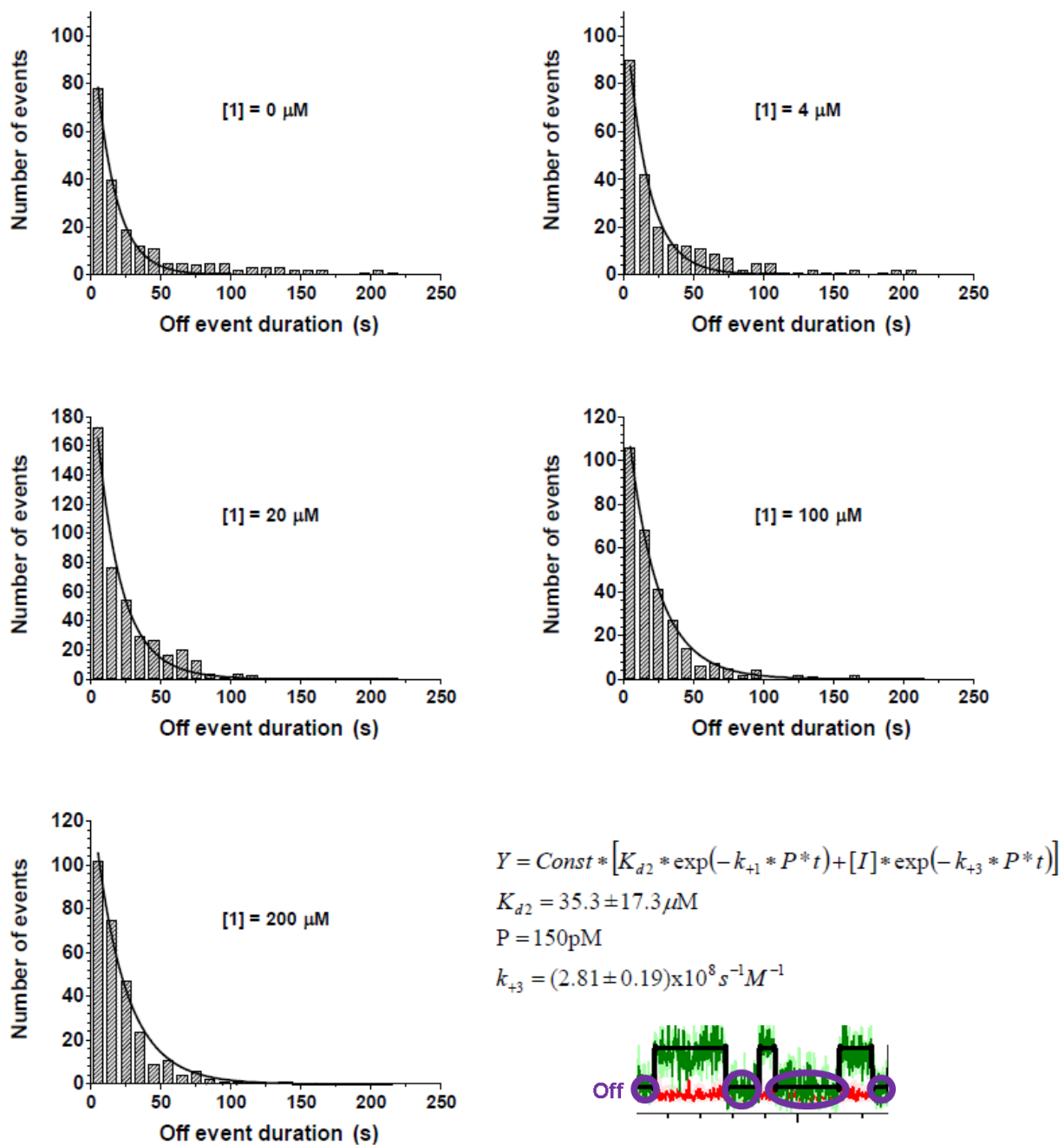


Figure 4.6. Representative global fitting of off-event dwell time for 1 at different concentrations.

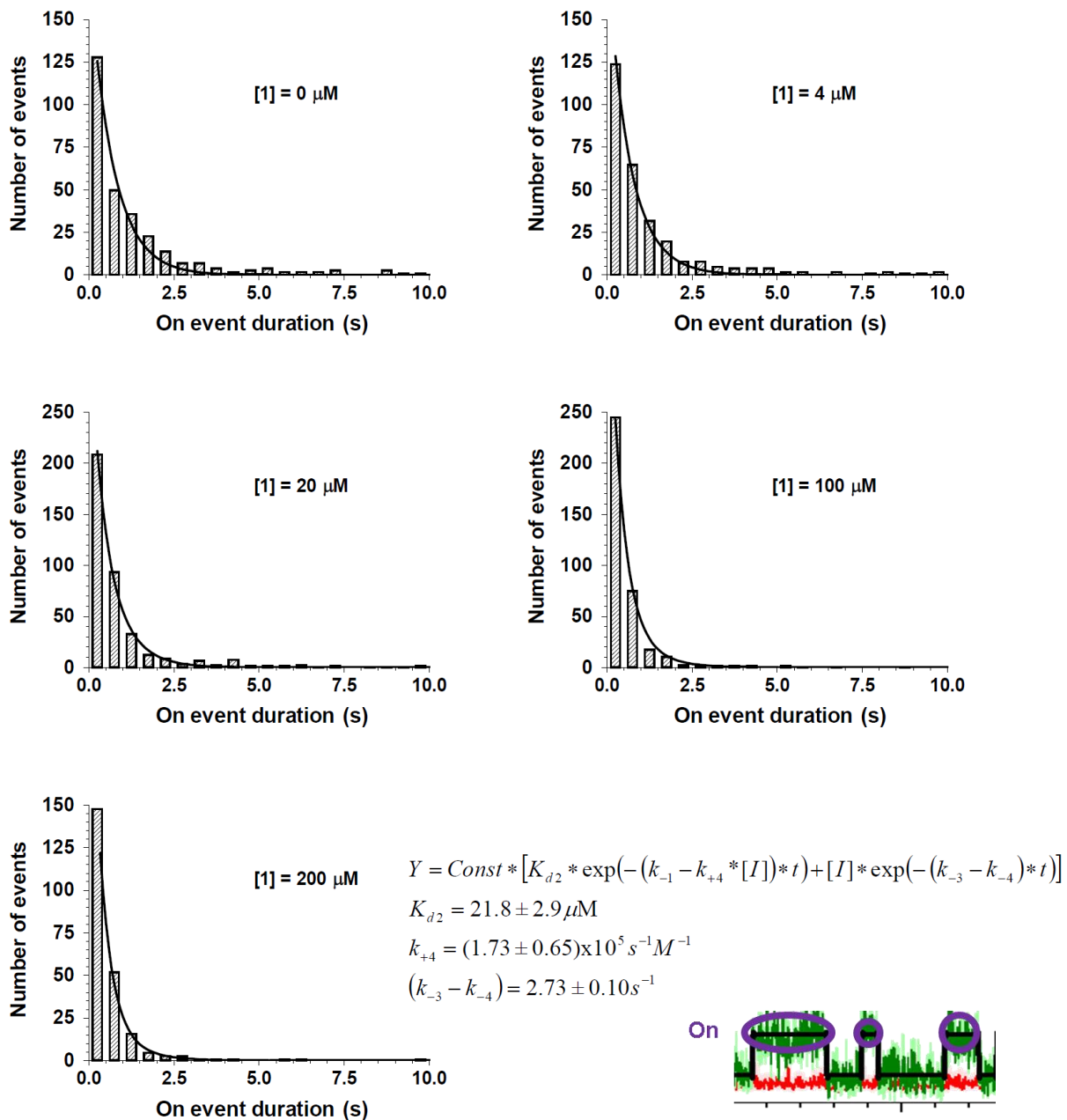
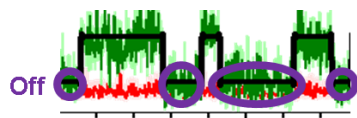


Figure 4.7. Representative global fitting of on-event dwell time for 1 at different concentrations.



$$Y = Const * [K_{d2} * \exp(-k_{+1} * P * t) + [I] * \exp(-k_{+3} * P * t)]$$

$$K_{d2} = 0.54 \pm 0.39 \mu\text{M}$$

$$P = 150 \text{ pM}$$

$$k_{+3} = (2.71 \pm 0.16) \times 10^8 \text{ s}^{-1} \text{ M}^{-1}$$

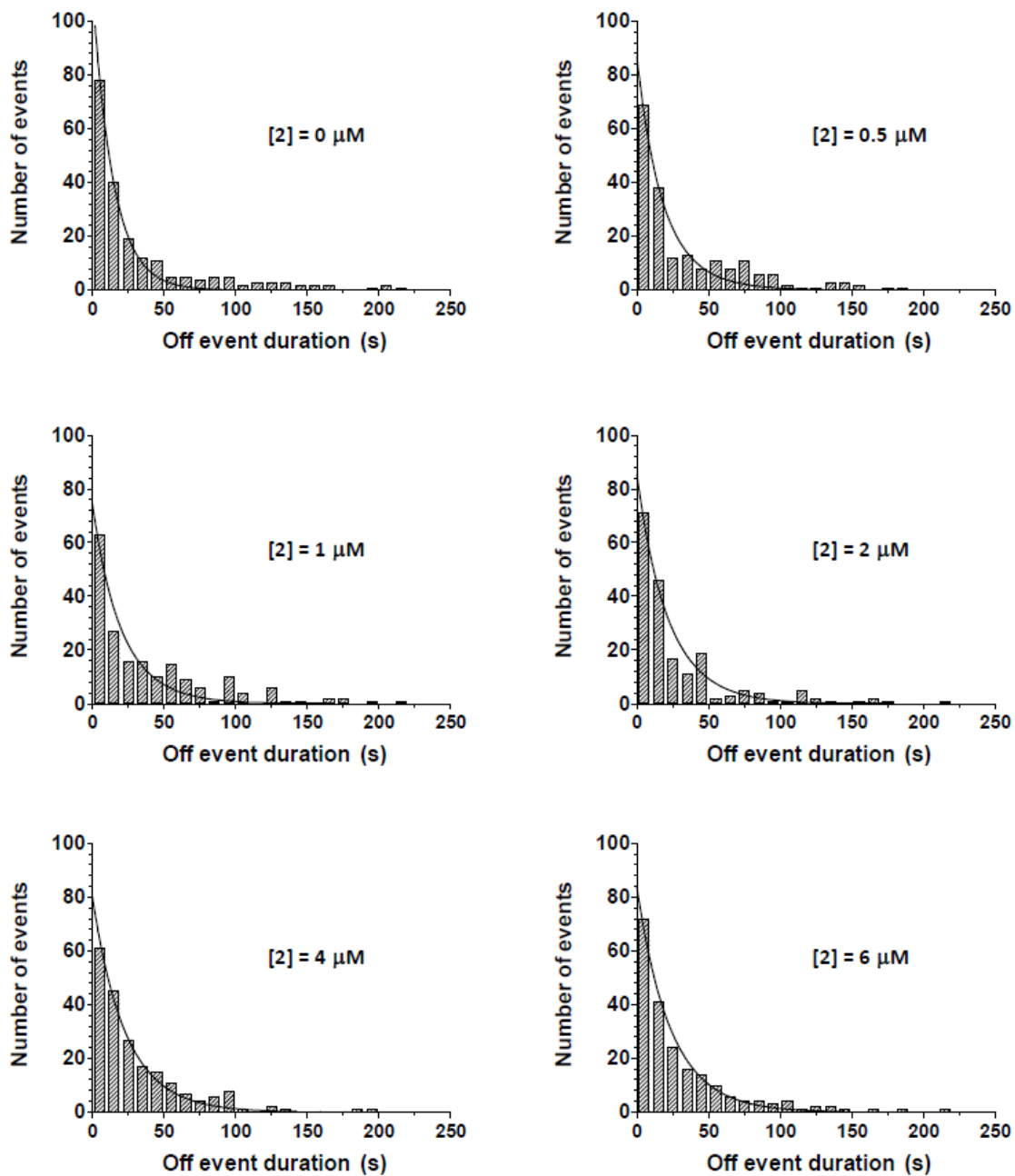


Figure 4.8. Representative global fitting of off-event dwell time for 2 at different concentrations.

$$Y = Const * [K_{d2} * \exp(-(k_{-1} - k_{+4} * [I]) * t) + [I] * \exp(-(k_{-3} - k_{-4}) * t)]$$

$$K_{d2} = 0.45 \pm 0.02 \mu\text{M}$$

$$k_{+4} = (1.93 \pm 0.03) \times 10^5 \text{ s}^{-1} \text{ M}^{-1}$$

$$(k_{-3} - k_{-4}) = 4.85 \pm 0.26 \text{ s}^{-1}$$

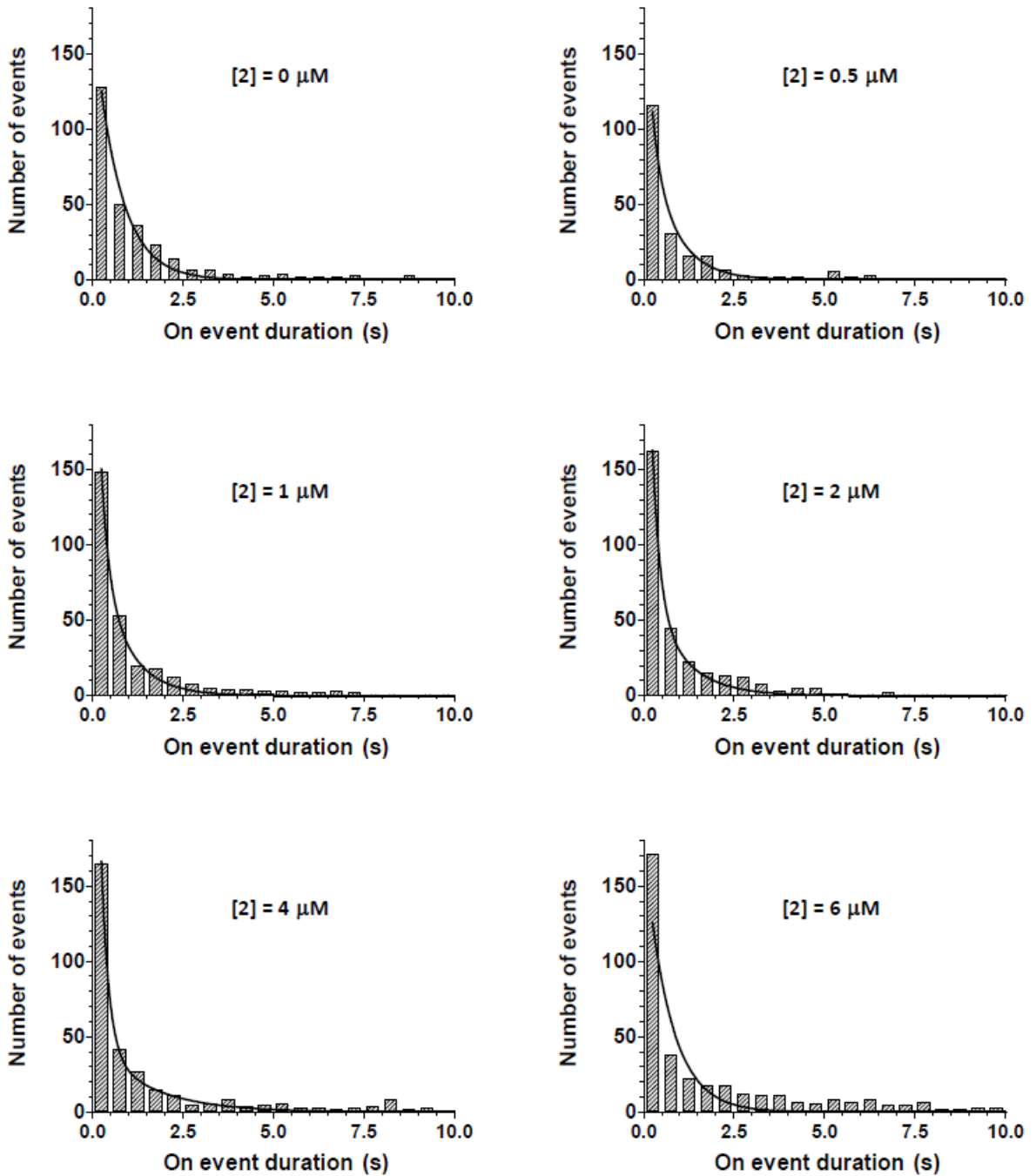
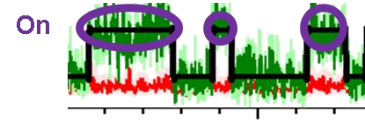


Figure 4.9. Representative global fitting of on-event dwell time for 2 at different concentrations.

Based on the proposed modes of (CUG)<sub>4</sub> binding by the ligand and MBNL1, it is possible that both can coexist on the same (CUG)<sub>4</sub> molecule: whereas the ligand is expected to interact with the U-U mismatch,<sup>17</sup> MBNL1 specifically binds to consecutive GC nucleotides.<sup>33</sup> The observed inhibition likely stems from different affinity of the MBNL1 protein for naked and ligand-bound (CUG)<sub>4</sub>. Most of this effect originates from increased off-rate of (CUG)<sub>4</sub>·MBNL1·ligand complex compared to (CUG)<sub>4</sub>·MBNL1. This situation can be described by a closed scheme of linked equilibria (Table 1a).

Assumptions derived from our experimental design and conditions allowed us to globally fit distributions of the time intervals between binding events, obtained at a range of ligand concentrations, to the equation 1.

$$K_{D2} = \frac{k_{-2}}{k_{+2}} = \frac{[RNA^{free}][I]}{[RNA \cdot I]},$$

is the equilibrium dissociation constant for ligand (1 or 2)

binding to (CUG)<sub>4</sub>. Free and ligand-bound RNA exist in equilibrium defined by  $K_{D2}$  and ligand concentration. All single-molecule TIRFM studies were carried out under conditions where the concentration of the inhibitor was much greater than was the concentration of surface-tethered RNA. Hence, it can be safely assumed that the total concentration of ligand equals the concentration of free ligand and is designated as [I] below. The second assumption is that only a small fraction of the protein is in the RNA·P or RNA·I·P complex because its concentration in the microscope flow cell is significantly lower than the expected  $K_{D1}$  or  $K_{D3}$ . The binding kinetics can be then deduced from the ON time distributions analyzed using the open scheme shown in Table 4.1b where the forward flux through the two manifolds will be determined by this ratio:

$$\frac{[RNA^{free}]}{[RNA \cdot I]} = \frac{K_D^1}{[I]} = \frac{flux(1)}{flux(2)}$$

Equation 1 assumes a simplified open scheme (see Table 4.1b) and contains a double exponential

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

where the decay constants correspond to kinetic association constants for MBNL1 binding to (CUG)<sub>4</sub>,  $k_1$ , and (CUG)<sub>4</sub>·I complex,  $k_3$ , respectively. The two exponentials are weighted by the abundance of each manifold under given conditions, which are defined by the equilibrium between free and ligand-bound (CUG)<sub>4</sub>. The experiments were carried



out in the presence of 0, 4, 20, 100 and 200  $\mu\text{M}$  of ligand **1** or in the presence of 0, 0.5, 1, 2, 4, 6 and 8  $\mu\text{M}$  of ligand **2**.

The closed scheme in Table 4.1a was used to analyze the ON time distributions. These distributions were globally fit to equation (2).

$$Y = \text{Const} \times [K_{D2} \times \exp(-(k_{-1} - k_4 \times [I]) \times t) + [I] \times \exp(-(k_{-3} - k_{-4}) \times t)] \quad (2)$$

$\text{Const} \times K_{D2}$  is the weight of manifold (1) whereas  $\text{Const} \times [I]$  is the weight of manifold (2) adjusted by the number of observed events. These analyses yielded  $K_{D2}$ ,  $k_4$  and  $(k_{-3} - k_{-4})$ . Parameters  $k_{-3}$  and  $k_{-4}$  are linked as  $(k_{-3} - k_{-4})$  and therefore cannot be determined individually by fitting but can be calculated from the linked equilibria as  $k_{-3} = \frac{(k_{-3} - k_{-4}) \times k_3 \times K_{D1}}{k_3 \times K_{D1} - k_4 \times K_{D2}}$ . The results are summarized in Table 4.1.<sup>4</sup>

Indicative of its function as a bivalent inhibitor, the binding affinity of **2** for  $(\text{CUG})_4$  was almost 50-fold higher than that of **1**. Because neither **1** nor **2** behave as typical competitive inhibitors, but instead allow formation of the  $(\text{CUG})_4 \cdot \text{MBNL1} \cdot \text{ligand}$  complex, the traditionally defined  $K_i$  is replaced by  $K_{D2}$ , which represents the affinity of the ligand for  $(\text{CUG})_4$ . We can, however, define an apparent  $\text{IC}_{50}$  as the concentration of ligand at which 50% of MBNL1 is free. It is notable, however, that this apparent  $\text{IC}_{50}$  depends on the concentration of the MBNL1 and  $(\text{CUG})_4$  (Figure 4.10) making each inhibitor to be effective in a relatively narrow concentration range.

Because of the kinetic scheme, which includes two pathways, we cannot use the traditional definition of  $K_i$ . We can, however, define the apparent  $\text{IC}_{50}$  as a concentration of the ligand at which 50% of MBNL1 is free.

---

<sup>4</sup> The single molecule data was collected by Dr. Masayoshi Honda, a postdoctoral researcher in Maria Spies lab. The analysis platform was developed by Professor Maria Spies (University of Iowa).

**Table 4.1.** Parameters from global fitting of the dwell-time distributions to the closed (a) and open (b) schemes of linked equilibria describing state interconversion in the (CUG)<sub>4</sub>·MBNL1·ligand system.

**a**

**b**

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

$k_i$  is assoc. rate const. and  $k_{-i}$  is dissoc. rate const.  $K_{Di} = k_{-i}/k_i$  is equilibrium dissoc. const. for the respective step in the Scheme. I = Inhibitor (1 or 2).

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

\* Errors were calculated as: 
$$\Delta(K_D) = K_D \times \sqrt{\left(\frac{\Delta k_{off}}{k_{off}}\right)^2 + \left(\frac{\Delta k_{on}}{k_{on}}\right)^2}$$

Apparent  $IC_{50} = I^{50}$ ; MBNL1 protein and  $(CUG)_4$  are shown as P and RNA, respectively.

$$\text{Total protein concentration } \hat{P}^{total} = \hat{P}^{FREE} + [RNA \times P] + [RNA \times I \times P] = P;$$

$$\text{Total RNA } \hat{RNA}^{total} = \hat{RNA}^{FREE} + [RNA \times I] + [RNA \times P] + [RNA \times I \times P] = R.$$

We need to express  $IC_{50}$  in terms of known concentrations and equilibrium dissociation constants:

$$\left\{ \begin{array}{l} (1) [P^{FREE}] = [RNA \cdot P] + [RNA \cdot I \cdot P] = \frac{P}{2} \\ (2) K_{D1} \times [RNA \cdot P] = [RNA^{FREE}] \times \frac{P}{2} \\ (3) K_{D3} \times [RNA \cdot I \cdot P] = [RNA \cdot I] \times \frac{P}{2} \\ (4) K_{D2} \times [RNA \cdot I] = [RNA^{FREE}] \times I^{50} \\ (5) R = [RNA^{FREE}] + [RNA \cdot I] + [RNA \cdot P] + [RNA \cdot I \cdot P] = [RNA^{FREE}] + [RNA \cdot I] + \frac{P}{2} \end{array} \right.$$

$$(4) \rightarrow [RNA \cdot I] = \frac{[RNA^{FREE}] \times I^{50}}{K_{D2}}$$

$$(5) \& (4) \rightarrow [RNA^{FREE}] + [RNA^{FREE}] \times \frac{I^{50}}{K_{D2}} = R - \frac{P}{2}$$

$$[RNA^{FREE}] = \frac{K_{D2} \times \left( R - \frac{P}{2} \right)}{K_{D2} + I^{50}} \quad (6)$$

$$(3) \& (4) \rightarrow [RNA \cdot I] = \frac{K_{D3} \times [RNA \cdot I \cdot P]}{\frac{P}{2}} = \frac{[RNA^{FREE}] \times I^{50}}{K_{D2}}$$

$$(1) \rightarrow [RNA \cdot I \cdot P] = \frac{P}{2} - [RNA \cdot P] = \frac{P}{2} - \frac{[RNA^{FREE}] \times \frac{P}{2}}{K_{D1}} = \frac{P}{2} \times \frac{K_{D1} - [RNA^{FREE}]}{K_{D1}}$$

$$[RNA \cdot I \cdot P] = \frac{\frac{P}{2} \times [RNA^{FREE}] \times I^{50}}{K_{D3} \times K_{D2}} = \frac{P}{2} \times \frac{K_{D1} - [RNA^{FREE}]}{K_{D1}}$$

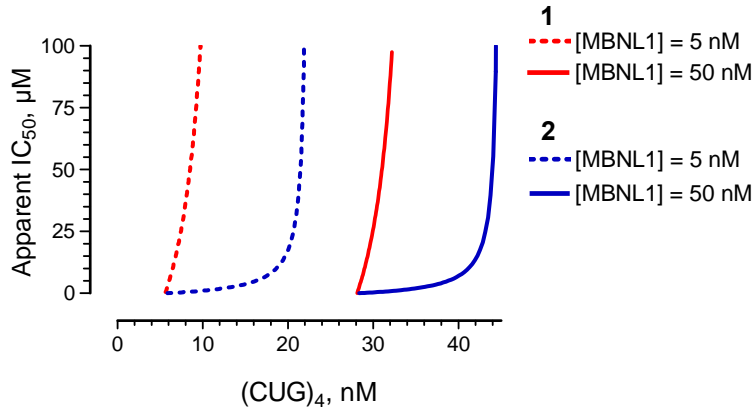
$$K_{D1} \times K_{D2} \times K_{D3} - K_{D2} \times K_{D3} \times [RNA^{FREE}] = K_{D1} \times [RNA^{FREE}] \times I^{50}$$

$$[RNA^{FREE}] \times (K_{D1} \times I^{50} + K_{D2} \times K_{D3}) = K_{D1} \times K_{D2} \times K_{D3}$$

$$[RNA^{FREE}] = \frac{K_{D1} \times K_{D2} \times K_{D3}}{(K_{D1} \times I^{50} + K_{D2} \times K_{D3})} \quad (7)$$

$$\begin{aligned}
(6) \& (7) \rightarrow \frac{K_{D1} \times K_{D2} \times K_{D3}}{K_{D1} \times I^{50} + K_{D2} \times K_{D3}} &= \frac{K_{D2} \times \left(R - \frac{P}{2}\right)}{K_{D2} + I^{50}} \\
(K_{D2} + I^{50}) \times K_{D1} \times K_{D3} &= (K_{D1} \times I^{50} + K_{D2} \times K_{D3}) \times \left(R - \frac{P}{2}\right) \\
I^{50} \times K_{D1} \times K_{D3} + K_{D1} \times K_{D2} \times K_{D3} &= I^{50} \times K_{D1} \times \left(R - \frac{P}{2}\right) + K_{D3} \times K_{D2} \times \left(R - \frac{P}{2}\right) \\
I^{50} \times \left(K_{D1} \times K_{D3} - K_{D1} \times \left(R - \frac{P}{2}\right)\right) &= K_{D3} \times K_{D2} \times \left(R - \frac{P}{2}\right) - K_{D1} \times K_{D2} \times K_{D3} \\
I^{50} &= \frac{K_{D2} \times K_{D3} \times \left(R - \frac{P}{2}\right) - K_{D1} \times K_{D2} \times K_{D3}}{\left(K_{D1} \times K_{D3} - K_{D1} \times \left(R - \frac{P}{2}\right)\right)} \quad (8)
\end{aligned}$$

Equation (8) relates the apparent  $IC_{50}$  to the absolute concentrations of protein and RNA and to equilibrium dissociation constants. It is obvious that the amount of ligand required to observe half of the protein in the free form will greatly depend on actual concentrations of ligand and RNA. Thus, it will be potent under relatively narrow set of conditions.

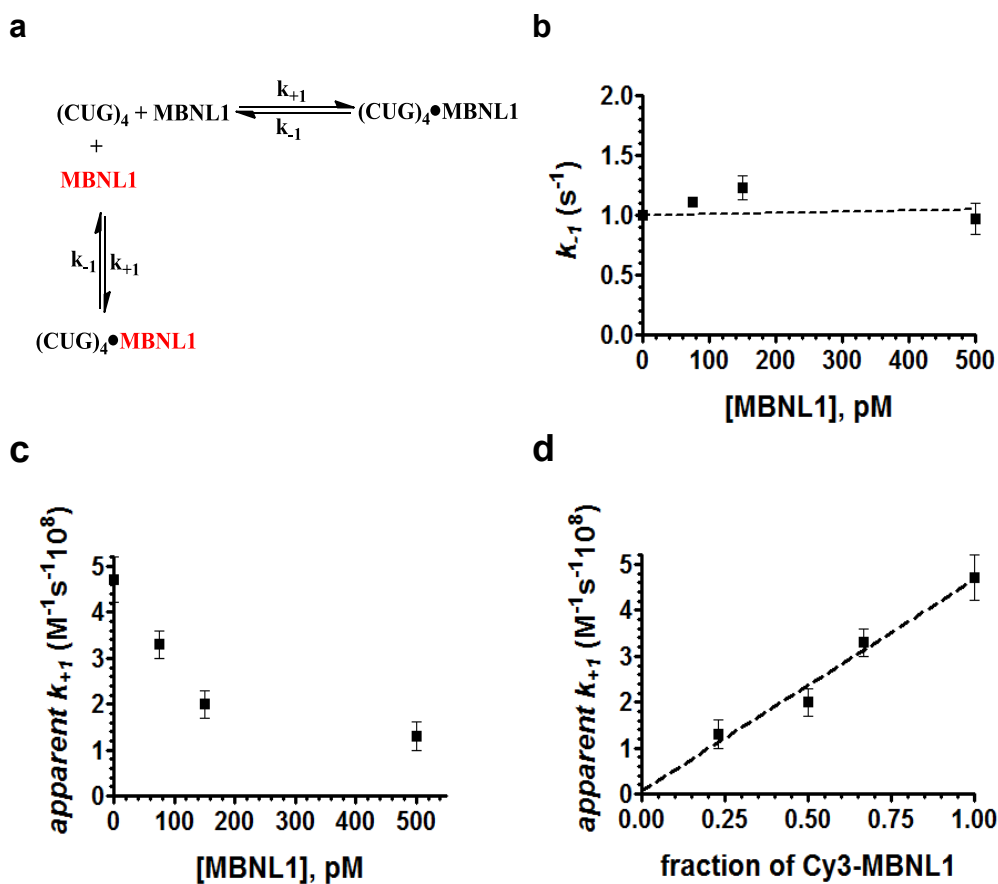


**Figure 4.10.** Apparent  $IC_{50}$  values for 1 and 2 at  $[MBNL1] = 5$  and  $50$  nM. The concentration of ligand required to observe half of MBNL1 in the free form,  $IC_{50}$ , depends on actual concentrations of MBNL1 and  $(CUG)_4$ .

At micromolar concentrations, many small molecules self-associate into colloidal aggregates that non-specifically inhibit protein activity. Aggregate-based inhibition is

sensitive to detergent,<sup>34</sup> therefore we confirmed that the inhibition by **2** was not affected by the presence of Triton X-100, eliminating the possibility of ligand aggregation.

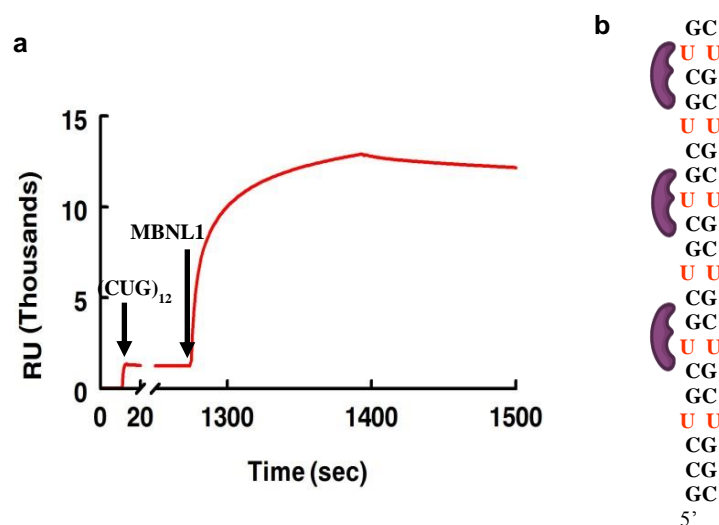
To ensure that the observed inhibition mode does indeed stem from the  $(\text{CUG})_4 \cdot \text{MBNL1}$  complex formation and not an experimental artifact, we analyzed the  $(\text{CUG})_4 \cdot \text{MBNL1}$  binding in the presence of the unlabeled MBNL1 protein, which should act a competitive inhibitor of Cy3-MBNL1 binding to the immobilized  $(\text{CUG})_4$ . Addition of unlabeled MBNL1 had no effect on  $k_{-1}$  for the  $(\text{CUG})_4 \cdot \text{MBNL1}$  complex; in contrast  $k_1$  decreased proportionally to the fraction of Cy3-MBNL1 (Figure 4.11). Not only did this control experiment confirmed that we can clearly distinguish the different modes of inhibition, but also that Cy3-labeling does not interfere with  $(\text{CUG})_4 \cdot \text{MBNL1}$  interaction.



**Figure 4.11.** Unlabeled MBNL1 acts as a competitive inhibitor. (a) Schematic representation of competitive inhibition. (b and c)  $k_{-1}$  is unaffected by addition of unlabeled MBNL1 whereas apparent  $k_{+1}$  is decreased. (d)  $k_{+1}$  is linearly dependent on  $[\text{Cy3-MBNL1}] / ([\text{unlabeled MBNL1}] + [\text{Cy3-MBNL1}])$  ratio.

**Binding Affinity and MBNL1 Inhibition Potency of Ligand 2 are Nearly 100-fold That of 1 in Bulk Experiments.** To provide a more conventional confirmation of single-molecule results, we used SPR to study the inhibition of MBNL1 binding to (CUG)<sub>4</sub> and (CUG)<sub>12</sub>. To distinguish binding of ligands from MBNL1 to the immobilized RNA constructs, the injection of MBNL1 was preceded by ligand injection. The experiments were done in the presence of a large excess of competitor tRNA (1 μM yeast tRNA) to confirm the specificity of ligands for CUG repeat. The maximum RU upon MBNL1 injection at various concentrations of each inhibitor was recorded. These values were normalized and plotted versus increasing concentration of each ligand to yield a binding isotherm. Apparent IC<sub>50</sub> values against (CUG)<sub>4</sub> and (CUG)<sub>12</sub> for **1** were 174 ± 12 μM and 293 ± 19 μM, respectively, whereas for **2**, the values were 1.3 ± 0.2 μM, and 1.1 ± 0.1 μM, respectively (Figure 4.13). Biotinylated RNA constructs were immobilized on a streptavidin coated surface. The immobilization density for most of the experiments was kept under 100 RU. For the stoichiometry study of MBNL1 binding to (CUG)<sub>12</sub>, 1100 RU of (CUG)<sub>12</sub> was immobilized. Stoichiometry of MBNL1 binding to CUG12 was obtained by measuring the steady state binding response from the injection of a highly concentrated MBNL1 solution (12 μM) for 120 s. Although the RU is still not fully reached the maximum steady state, we could approximately derive the stoichiometry ratio for MBNL1 binding to (CUG)<sub>12</sub> using this equation:  $R_L = \left(\frac{R_{max}}{S_m}\right)\left(\frac{MW_L}{MW_A}\right)$

R<sub>L</sub> is the density of immobilized (CUG)<sub>12</sub>, ligand, on the chip surface; R<sub>max</sub> is the maximum binding response for MBNL1 binding; S<sub>m</sub> is the binding stoichiometry; MW<sub>L</sub> is the molecular weight of (CUG)<sub>12</sub>, 13293 Da, and MW<sub>A</sub> is the molecular weight of GST-MBNL1 analyte, 60489 Da. Analysis of the steady-state SPR response unit (RU) yielded a binding stoichiometry of almost 3:1 for MBNL1 binding to (CUG)<sub>12</sub>, which implies that one (CUG)<sub>12</sub> can bind to three MBNL1 proteins.



**Figure 4.12.** Representative sensogram for stoichiometric study of MBNL1·(CUG)<sub>12</sub> interaction. (a) The arrows show the starting point of (CUG)<sub>12</sub> immobilization and MBNL1 injection. The sensogram suggests stoichiometric ratio of 3:1 for MBNL1:(CUG)<sub>12</sub> binding as shown schematically in (b).

RU change in the sensogram was observed upon addition of inhibitors **1** and **2** as well as MBNL1, indicating binding to the biotinylated CUG repeats. Therefore, for the inhibition studies at different concentrations of small molecules, simultaneous injection of both MBNL1 and inhibitors made the analysis impossible. This limitation results from the fact that the RU signal would have to be deconvoluted to determine whether the binding is from MBNL1 or the small molecule. To circumvent this problem, we set up an experiment in which injection of the small molecule preceded the injection of MBNL1. This way, by pre-incubation of the surface with small molecule, we were able to reach the plateau binding level of the small molecule. Then, while keeping the small molecule concentration the same, the second injection gave rise only to the MBNL1 binding. With the assumption that ligand does not dissociate significantly under these conditions, the proportion of MBNL1 binding that is inhibited at any concentration of small molecule could be determined. The highest RU was measured at the point that the injection was stopped. This is end point of the association phase and start of the dissociation phase. This RU is an indication of MBNL1 binding level, therefore the higher the inhibitor concentration, the lower this would be point.

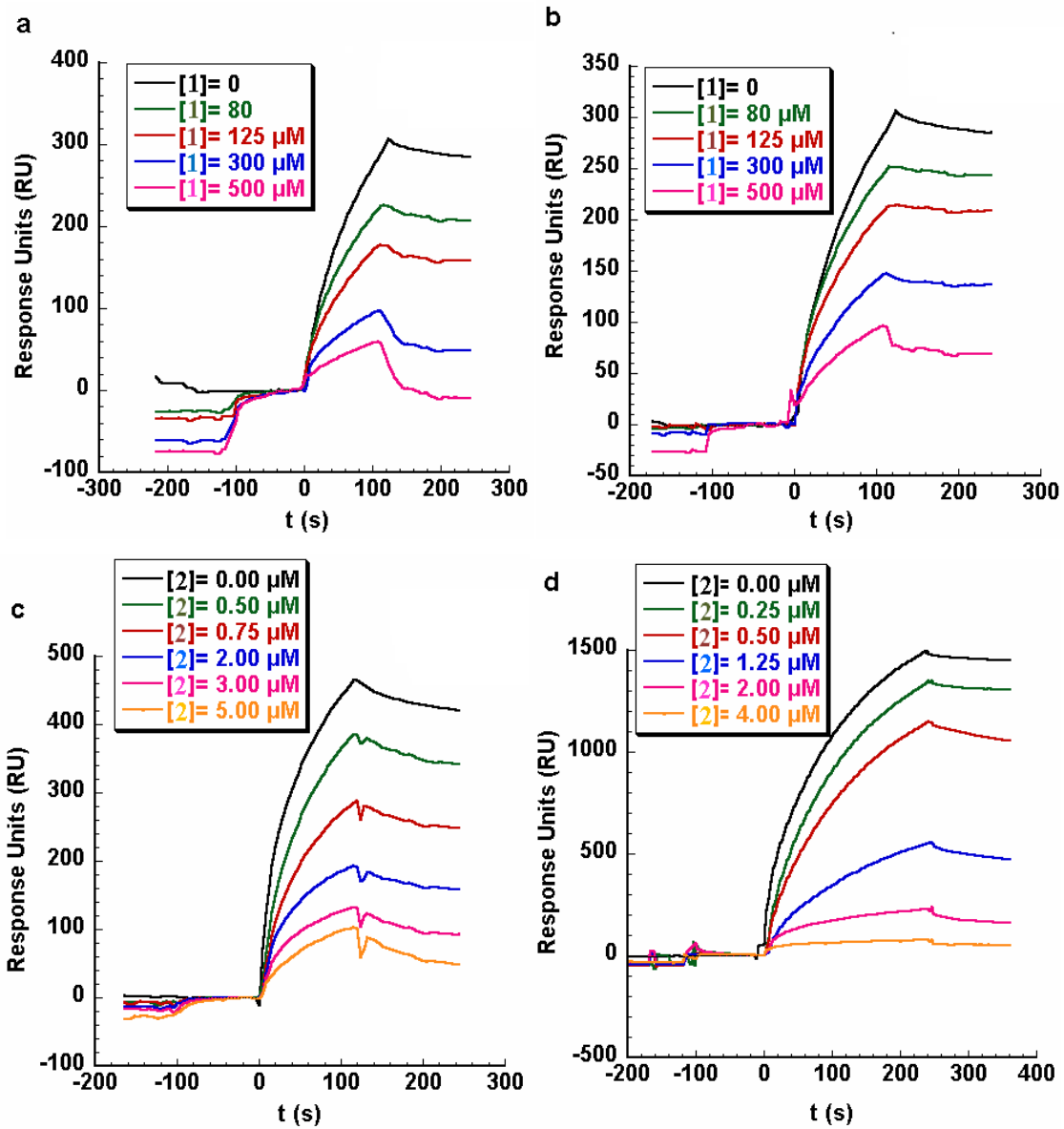
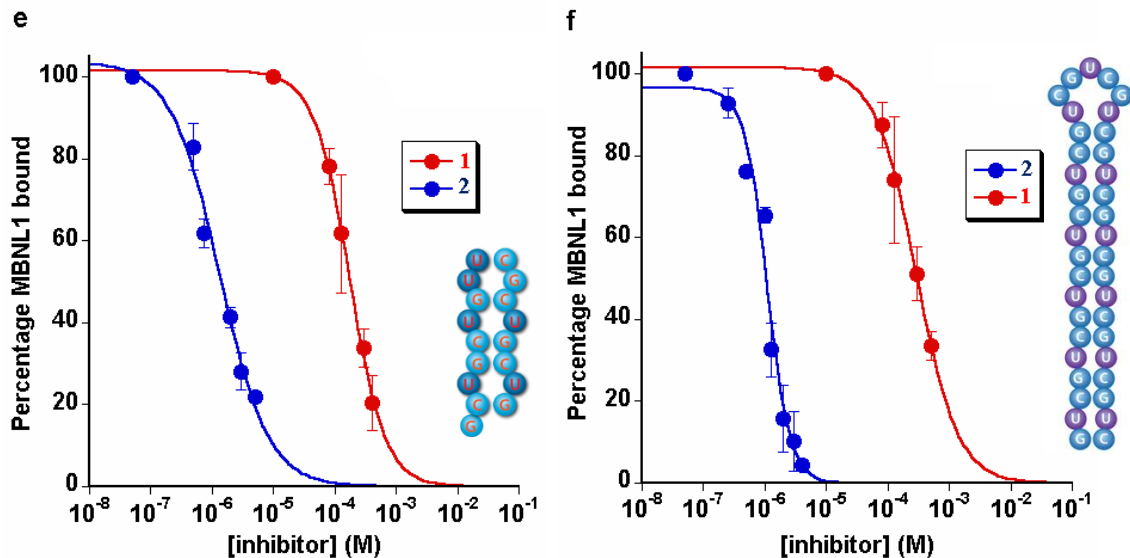


Fig 4.13 (cont. on next page)

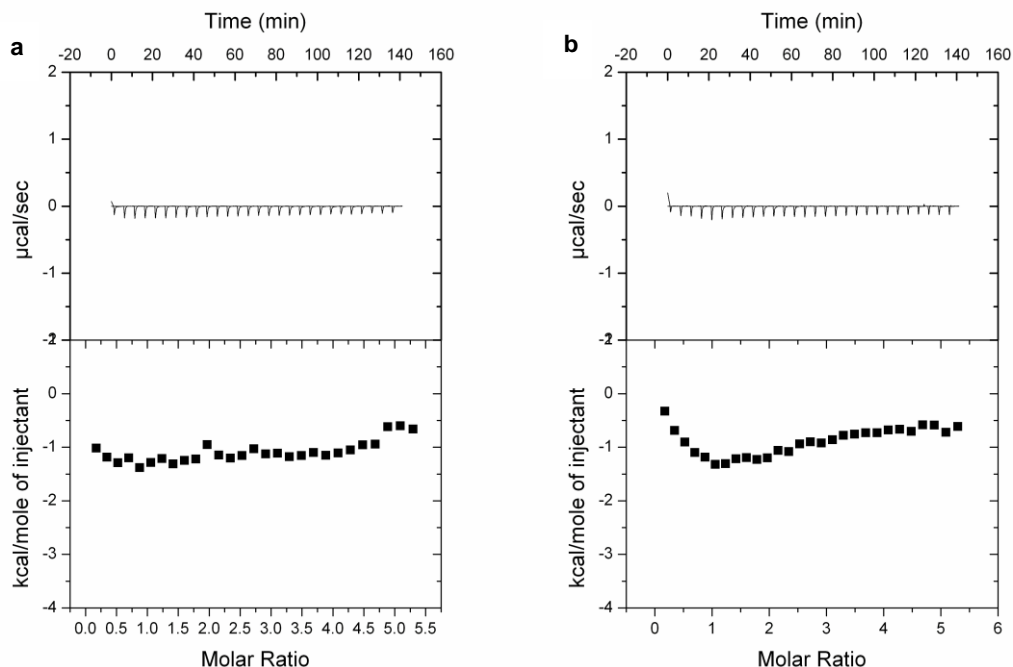




**Figure 4.13.** Representative Sensograms from SPR studies and fitting data to dose-response curves. Biotinylated (CUG)<sub>4</sub> is the immobilized ligand in (a) and (c). Biotinylated (CUG)<sub>12</sub> is the immobilized ligand in (b) and (d). Varying concentration of **1** and **2** are injected from  $t = -120$  s to either  $t = 120$  s (a, b and c) or  $t = 240$  s (d). GST-MBNL1,  $0.65 \mu\text{M}$ , is injected from  $t = 0$  s to either  $t = 120$  s (a, b and c) or  $t = 240$  s (d). Baseline for the curves was set to RU = 0 at  $t = 0$ . (e) Inhibition of MBNL1 binding to (CUG)<sub>4</sub> in the presence of varying concentrations of **1** and **2**. (f) Inhibition of MBNL1 binding to (CUG)<sub>12</sub> in the presence of varying concentrations of **1** and **2**. Error bars represent mean  $\pm$  s.d. of three replicates.

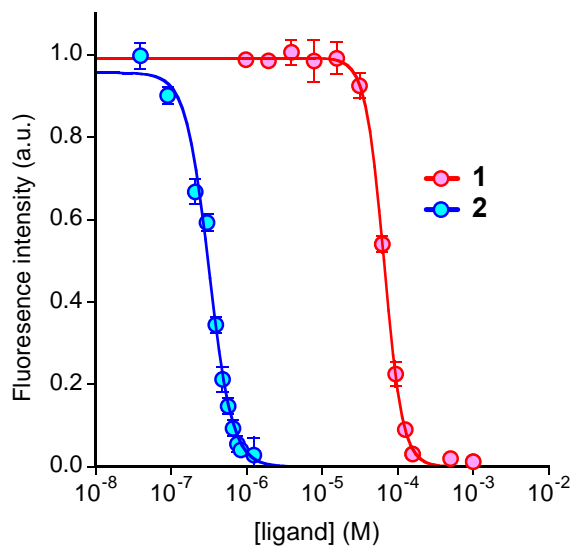
For each ligand the fact that apparent  $\text{IC}_{50}$  values for inhibition of MBNL1 binding to either (CUG)<sub>4</sub> and (CUG)<sub>12</sub> are close, supports the previous findings that both these constructs are acceptable models for CUG repeat.<sup>30</sup> Normalized  $\text{IC}_{50}$  values can be calculated by multiplying the actual  $\text{IC}_{50}$  by the number of binding modules it has. The ratio between normalized  $\text{IC}_{50}$  values is the bivalent effect, which is almost 120-fold for **2**. This improvement in inhibition confirms the inhibition result obtained by single molecule TIRF microscopy, however, by this method, we were unable to measure how the binding kinetic change in the presence of inhibitors or reveal the inhibition mode of the ligands. This is due to the fact that this technique doesn't allow us to differentiate binding of inhibitor from MBNL1. Therefore at any concentration of inhibitor, whether in association or dissociation phase, we are detecting association and dissociation of both MBNL1 and small molecules to the immobilized RNA construct. This deficiency is overcome in the single molecular TIRFM experiments by only MBNL1 with Cy3 which enables us to detect binding and unbinding of an individual MBNL1 protein.

Due to non-competitive inhibition mechanism,  $IC_{50}$  values should depend on the concentration of immobilized  $(CUG)_n$  on the SPR sensor chip, which is not measurable. Nonetheless, obtaining smaller  $IC_{50}$  value for **2**, compared to **1**, is in agreement with the single-molecule results. To verify that **1** and **2** don't bind to MBNL1 protein we carried out Isothermal Titration Calorimetry experiments. ITC titration curves show no measureable interaction between the ligands and MBNL1 protein (Figure 4.14).



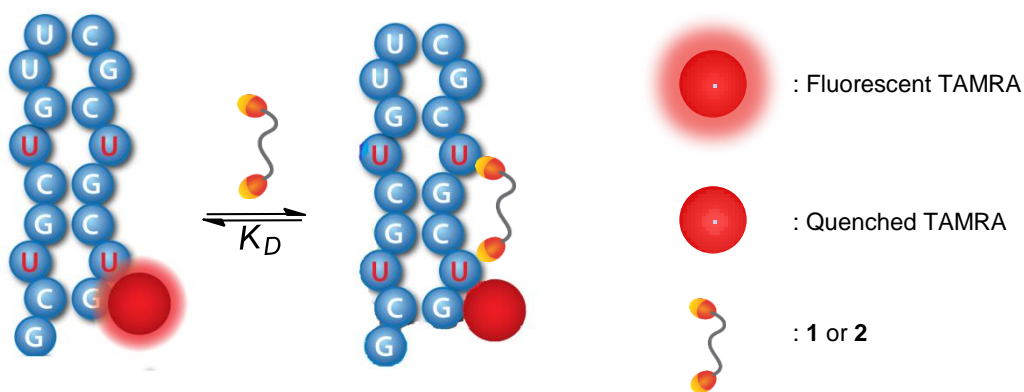
**Figure 4.14.** ITC binding isotherms and titration curves of **1** (a) and **2** (b) to MBNL1 solution. No measureable binding is observed.

To measure the binding constants of **1** and **2** to CUG repeat by another bulk approach, a steady state fluorescence titration method with TAMRA- $(CUG)_6$  was utilized. It is known that guanosine quenches the TAMRA fluorescence through photoinduced electron transfer.<sup>35, 36</sup> Therefore, it is possible that binding of a ligand to the UU mismatch close to 5'-TAMRA makes a structural change that can lead to quenching of the 5'-TAMRA by the 3'-G in 5'-TAMRA- $(CUG)_6$ -3'. Ligands **1** and **2** were titrated into TAMRA- $(CUG)_6$  solution.



**Figure 4.15.** Fluorescence titrations of TAMRA-(CUG)<sub>6</sub> with **1** and **2**. Comparison of normalized fluorescence intensity change of TAMRA-(CUG)<sub>6</sub> in the presence of increasing concentrations of **1** and **2**. TAMRA was excited at 560 nm and its emission was recorded at 590 nm. Error bars represent mean  $\pm$  s.d. of three replicates.

Upon increasing the ligand concentration, TAMRA fluorescence intensity gradually decreased as a result of fluorophore quenching by the bound ligand. A plot of normalized fluorescence intensity versus increased concentrations of each ligand yielded a binding isotherm with  $K_D$  of  $66 \pm 2 \mu\text{M}$  and  $318 \pm 15 \text{ nM}$  for **1** and **2**, respectively (Figure 4.15). These  $K_D$  values are similar to what we derived from the single-molecule study (Table 1).



**Figure 4.16.** Schematic representation of steady-state fluorescence titrations of TAMRA-(CUG)<sub>6</sub> with **1** and **2**.

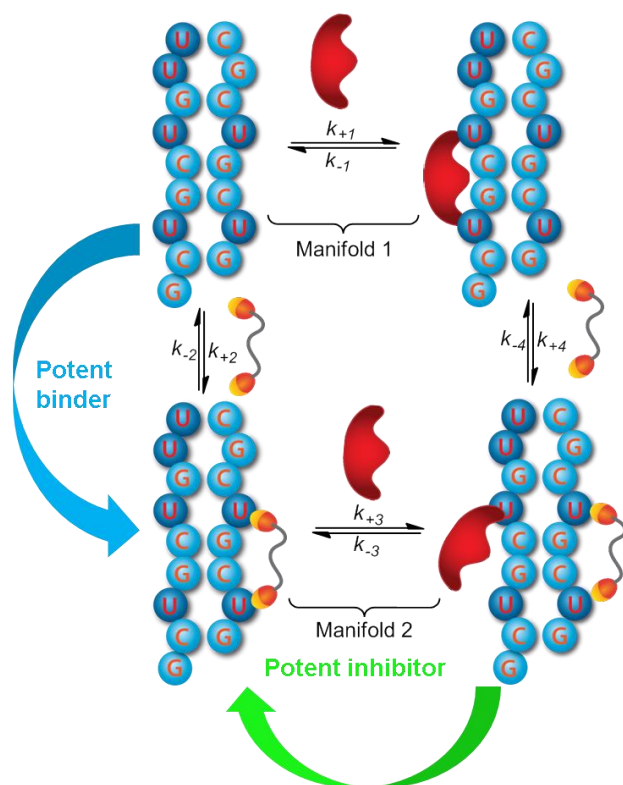
## 4.4 Discussion

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

We developed and employed a TIRFM-based single molecule analysis to study the interaction of MBNL1, a key regulatory protein in alternative splicing process, and (CUG)<sub>4</sub>, a validated model of CUG<sup>exp</sup> as well as the inhibition of this interaction by the small molecule RNA binding ligands. Our analysis revealed the affinity of MBNL1 for (CUG)<sub>4</sub> to be  $3.1 \pm 0.1$  nM, whereas using EMSA we,<sup>17</sup> and others<sup>30</sup> reported  $K_D$  values of  $26 \pm 4$  nM and  $170 \pm 20$  nM, respectively. This was not surprising because the single-molecule measurements were carried out under true equilibrium conditions, whereas EMSA can underestimate affinity because the short-lived complexes may dissociate while being resolved. Additionally, only productive (native) interactions between (CUG)<sub>4</sub> and MBNL1 molecules were detected by the single-molecule measurements thereby canceling out any possible errors resulting from heterogeneity of MBNL1 or (CUG)<sub>4</sub>.

Single-molecule analyses of the MBNL1·(CUG)<sub>4</sub> interaction in the presence of two inhibitors allowed us to determine the affinities of the two tested ligands for (CUG)<sub>4</sub> as well as how they affect MBNL1·(CUG)<sub>4</sub> complex formation. Notably, the equilibrium dissociation constants for the ligand·(CUG)<sub>n</sub> binding determined in bulk and at the single-molecule level were very similar. Converging of the binding affinity of the 1·

(CUG)<sub>n</sub> and 2·(CUG)<sub>n</sub> complexes (n = 4 or 6) measured by the fluorescence and single-molecule methods to the same value validated the single-molecule method. Our single-molecule analysis was model-independent and suggested an inhibition mode that has not been considered previously: instead of competing for the same binding site the ligands and MBNL1 can form a ternary complex on (CUG)<sub>4</sub>.



**Figure 4.17.** Schematic representation of proposed non-competitive MBNL1 · CUG<sup>exp</sup> inhibition by **2**.

In the more physiologically relevant buffer used in this study, ligand **1** was a relatively weak inhibitor, which bound to (CUG)<sub>4</sub> with a low affinity (i.e.,  $K_D = 22 \pm 3 \mu\text{M}$ ). Moreover, it allowed MBNL1 to bind to the 1·(CUG)<sub>4</sub> complex with a very high affinity ( $K_D = 10.2 \pm 0.8 \text{ nM}$ ). This interesting finding prompted us to design more effective small molecules. We pursued a bivalent ligand with the idea that it would: (1) bind to CUG<sup>exp</sup> with a higher affinity and ideally weaken the MBNL1 interaction with ligand·CUG<sup>exp</sup>, (2) exhibit a longer occupancy time and increase the RNA ds stabilization, if MBNL1 indeed binds preferentially to the ss form of RNA as has been suggested,<sup>41, 42</sup> and (3) bind a longer segment of RNA to achieve more effective coverage and steric blocking of CUG<sup>exp</sup>.

To reach the drug development goals outlined above, we designed a library of dimeric ligands and found **2** to be the most potent inhibitor of CUG repeat·MBNL1 interaction.

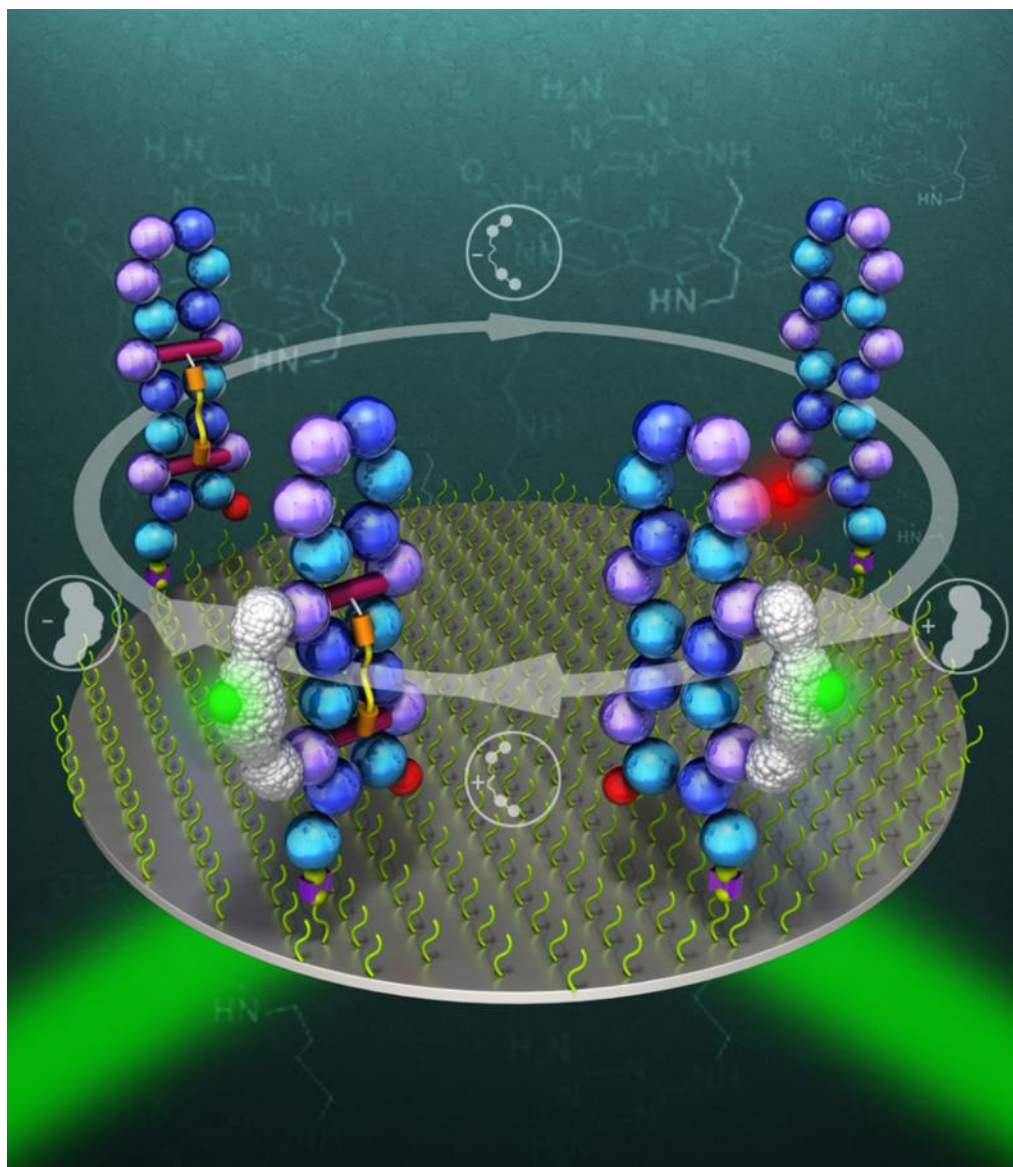
The affinity of **2** for (CUG)<sub>4</sub> was measured by the single-molecule method to be  $K_D = 450 \pm 20$  nM, a value that is almost 50-fold lower than that measured for **1**. Despite this large increase in affinity, MBNL1 can still bind to the **2**·(CUG)<sub>4</sub> complex. This unexpected observation of a non-competitive inhibition mode for both **1** and **2** is the key finding of this study. Its implication for drug design is obvious. The **2**·(CUG)<sub>4</sub> complex reduces the MBNL1 affinity only 6-fold relative to naked (CUG)<sub>4</sub> and 2-fold relative to that measured for the **1**·(CUG)<sub>4</sub> complex. Thus, tight binding is an insufficient criterion for effective inhibition because it may shift the overall equilibrium into manifold **2** of Table 4.1a (i.e., low  $K_{D2}$ ) without necessarily affecting MBNL1 binding. Highly effective inhibition will result from agents that exhibit both tight and selective binding to CUG<sup>exp</sup> (low  $K_{D2}$ ) and high  $K_{D3}$  values.

Another significant result is that the apparent IC<sub>50</sub> values for the ligands depend markedly on both (CUG)<sub>4</sub> and MBNL1 concentrations (Figure 4.10). Thus, the CUG<sup>exp</sup> length and cellular copy number and the MBNL1 concentration will determine the drug effectiveness. Because the CUG repeat continues to expand over the patient's life, an ideal therapeutic agent should be able to inhibit binding of MBNL1 to CUG<sup>exp</sup> over a broad range of nuclear CUG<sup>exp</sup> copy number and repeat lengths. By single molecule and bulk methods (Figures 4.10, 4.13 and Table 4.1) we found that, given the same concentration of MBNL1, **2** effectively inhibits (CUG)<sub>n</sub>·MBNL1 (n= 4 or 12) interaction to a greater extent than **1** and over a broader range of (CUG)<sub>n</sub> concentrations, a trend that can be further improved in the future ligands.

## 4.5 Conclusion

This study revealed that inhibition of (CUG)<sub>4</sub>·MBNL1 interaction by small molecules does not occur by a simple competitive mechanism. These findings aided in the design of a more potent (CUG)<sub>4</sub>·MBNL1 inhibitor and are guiding our current optimization efforts. More generally, the increasingly frequent discovery of key roles for RNA outside its established role in protein synthesis suggests new opportunities for RNA-targeted therapeutics. Indeed, RNA is now a validated, yet underutilized drug target, with less off-pathway binding. As the development of inhibitors for protein·RNA interactions becomes more widespread, we believe the inhibition model revealed in this study will need to be considered. It is likely that the dynamic and versatile structure of RNA allows it to form ternary ligand·protein·RNA complexes, although with lower stability.<sup>43</sup> Finally, the single-

molecule methodology described herein may prove to be a powerful method to unravel the inhibition mechanism for any biomacromolecular interaction.



**Figure 4.18.** Schematic representation of four states in equilibrium in the proposed non-competitive MBNL1 · CUG<sup>exp</sup> inhibition by ligands.

## 4.6 Materials and Methods

**Instrumentation and Chemicals.** All reagents were purchased from commercial suppliers and were used without further purification. Anhydrous solvents were obtained from an anhydrous solvent dispensing system. For all reactions employing anhydrous solvents, glassware was oven-dried, cooled under vacuum, and then purged with dry nitrogen; all reactions were conducted under dry nitrogen. Purified compounds were further dried under high vacuum (0.01–0.05 Torr) or lyophilized using a Labconco lyophilizer. Yields refer to purified and spectroscopically pure compounds. NMR spectra were recorded at 23 °C on either Varian Unity 500 or Varian Unity Inova 500NB, operating at 500 MHz and 125 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  acquisitions, respectively. NMR spectra were processed using MestReNova software. Chemical shifts are reported in ppm and referenced to the corresponding residual nuclei in the following deuterated solvents:  $\text{CDCl}_3$  (7.26 ppm  $^1\text{H}$ , 77.16 ppm  $^{13}\text{C}$ ); DMSO (2.50 ppm  $^1\text{H}$ , 39.52 ppm  $^{13}\text{C}$ );  $\text{D}_2\text{O}$  (4.79 ppm  $^1\text{H}$ );  $\text{CD}_3\text{OD}$  (3.31 ppm  $^1\text{H}$ , 50.41 ppm  $^{13}\text{C}$ ). Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), sext (sextet), dd (doublet of doublets), ddd (doublet of doublet of doublets), td (triplet of doublets), dt (doublet of triplets), m (multiplet), b (broad). Integration is provided and coupling constants,  $J$ , are reported in Hertz (Hz). ESI mass spectra were recorded using the Quattro or ZMD mass spectrometer. High resolution mass spectra (HRMS) were obtained at the University of Illinois mass spectrometry facility. All compounds described herein gave NMR and mass spectral data in accord with their structures. Ligands **1** and **2** gave a HRMS within 5 ppm of calculated values.

**MBNL1N Plasmid and RNA.** The expression vector pGEX-6p-1/MBNL1N was obtained from Maurice S. Swanson (University of Florida, College of Medicine, Gainesville, FL, USA)<sup>44</sup>. MBNL1N is comprised of the four zinc finger motifs of MBNL1. It contains a 6xHis tag at the C-terminus and the GST tag at the N-terminus. MBNL1N binds RNA with similar affinity as the full-length MBNL1. It is referred to as MBNL1 throughout this paper for the sake of simplicity. All the oligonucleotides were purchased from Integrated DNA Technology and were HPLC purified. The sequences and modifications for RNA constructs used in each study are shown in the respective material section.

**MBNL1N Protein Expression and Purification.** Using BL21-CodonPlus(DE3)-RP competent cells (Stratagene), the expression of MBNL1N protein was induced with 1mM



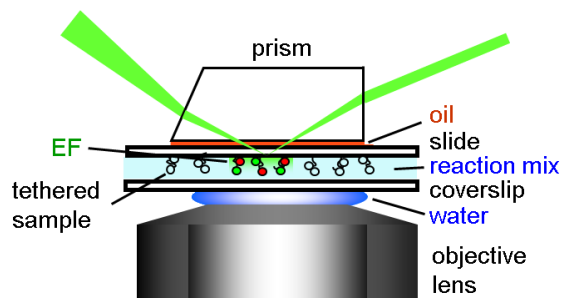


TAMRA-(CUG)<sub>6</sub> at various ligand concentrations. The assays were performed using a Cary Eclipse Fluorescence Spectrophotometer (Varian). TAMRA-(CUG)<sub>6</sub> was excited at 560 nm and its emission was recorded at 590 nm. Stoichiometric titrations were carried out at 20 °C in PBS, 1X buffer.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

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

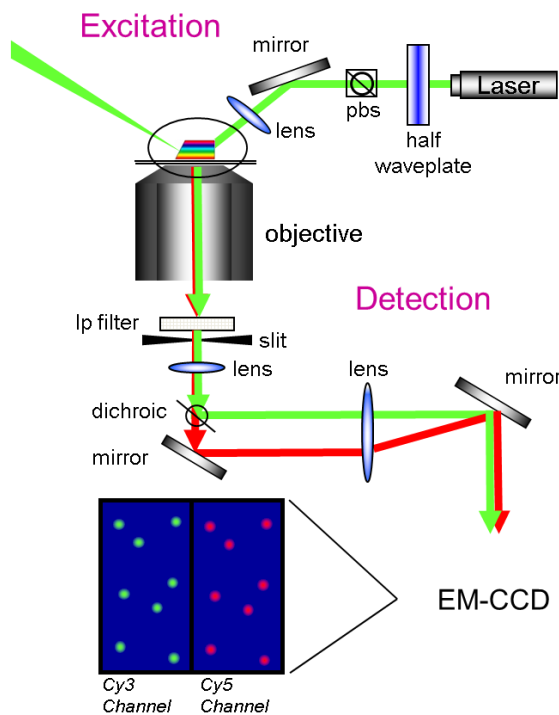
$K_D$  is the dissociation binding constant,  $[L]$  is the ligand concentration,  $n$  is the Hill coefficient and  $F_0$  and  $F_{max}$  are the fluorescence intensity of free and fully bound RNA, respectively. In the control experiment, ligands **1** and **2** had no effect on the fluorescence intensity of the free TAMRA fluorophore. The sequences and modifications for (CUG)<sub>6</sub> RNA construct used in this study: 5'- TAMRA-CUGCUGCUGCUGCUGCUG – 3'

**Reaction Conditions for the Single Molecule Assay.** Biotinylated (CUG)<sub>4</sub> or (CUG)<sub>12</sub> were immobilized on a quartz surface, which was coated with polyethylene glycol in order to eliminate nonspecific surface adsorption of proteins. The immobilization was mediated by biotin-neutravidin interaction between biotinylated Cy5-labeled (CUG)<sub>4</sub>, neutravidin and biotinylated polymer (PEG, MW = 5,000). Standard PBS buffer contained the oxygen scavenging system consisting of 1 mg/mL glucose oxidase, D-glucose (0.4%, w/v), 2-mercaptoethanol (1%, v/v) and 0.04 mg/mL catalase. Immobilization of 50 pM of each oligonucleotide allowed for detection of 100–600 individual molecules per slide. Cy3 labeled MBNL1 was then added and incubated for 5 min at 25°C in the standard PBS, 1X buffer in the absence or presence of the indicated concentrations of **1** and **2**. Sample excitation and data recording were initiated after all components of the MBNL1·(CUG)<sub>4</sub>·inhibitor system were allowed to equilibrate in the TIRFM sample chamber. The presence of ligands had no effect on the fluorescence intensity of the Cy3 labeled MBNL1.



**Figure 4.19.** Schematic representation of TIRF set-up used in the MBNL1·CUG<sup>exp</sup> inhibition study by ligands (courtesy of Dr. Masayoshi Honda).

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



**Figure 4.20.** Schematic representation of TIRF microscopy set-up used to localize Cy5-labeled (CUG)<sub>4</sub> and Cy3-labeled MNBL1 (courtesy of Dr. Masayoshi Honda).

**Surface Plasmon Resonance (SPR) Analysis.** All SPR experiments were conducted on a streptavidin coated sensor chip using a Biacore 3000 instrument. Streptavidin coated research grade sensor chips were preconditioned with three consecutive 1-minute injections of 1 M NaCl/ 50 mM NaOH before the immobilization was started. 3'-biotin labeled RNA (CUG<sub>4</sub> or CUG<sub>12</sub>) was captured on flow cell 2 (Response Unit, RU, between 100-1100). Flow cell 1 was used as a reference. Inhibition analysis was carried out in PBS, 1X buffer, pH = 7.4, containing 0.05% Tween-20 and 0.2 mg/mL (7.4  $\mu$ M or 580  $\mu$ M nucleotides) bulk yeast t-RNA to confirm the specificity of inhibition. Various concentrations of **1** and **2** were flowed at a rate of 20  $\mu$ L/min for 240 s over the immobilized RNA. After the initial 120 s, a solution of GST-MBNL1 protein, 600 nM, in the same buffer was also flowed over the surface for the rest of time, 120 s. The reference-subtracted sensograms were recorded.

After the dissociation phase, the surface was regenerated, with a pulse of 0.5% SDS and/or 100 mM NaOH, for a few times followed by buffer wash to re-establish baseline. RU upon the injection of PBS buffer was subtracted from sensograms. For inhibition studies, the resulting sensograms were baselined at  $t = 120$  s to offset the binding of small molecule to the immobilized RNA surface. The peak RU ( $t = 120$  s) was recorded.

Two or three separate SPR experiments on different sensor chips with different levels of RNA immobilization were performed to verify that the values are not affected by surface RNA density. The sequences and modifications for RNA constructs used in this study:

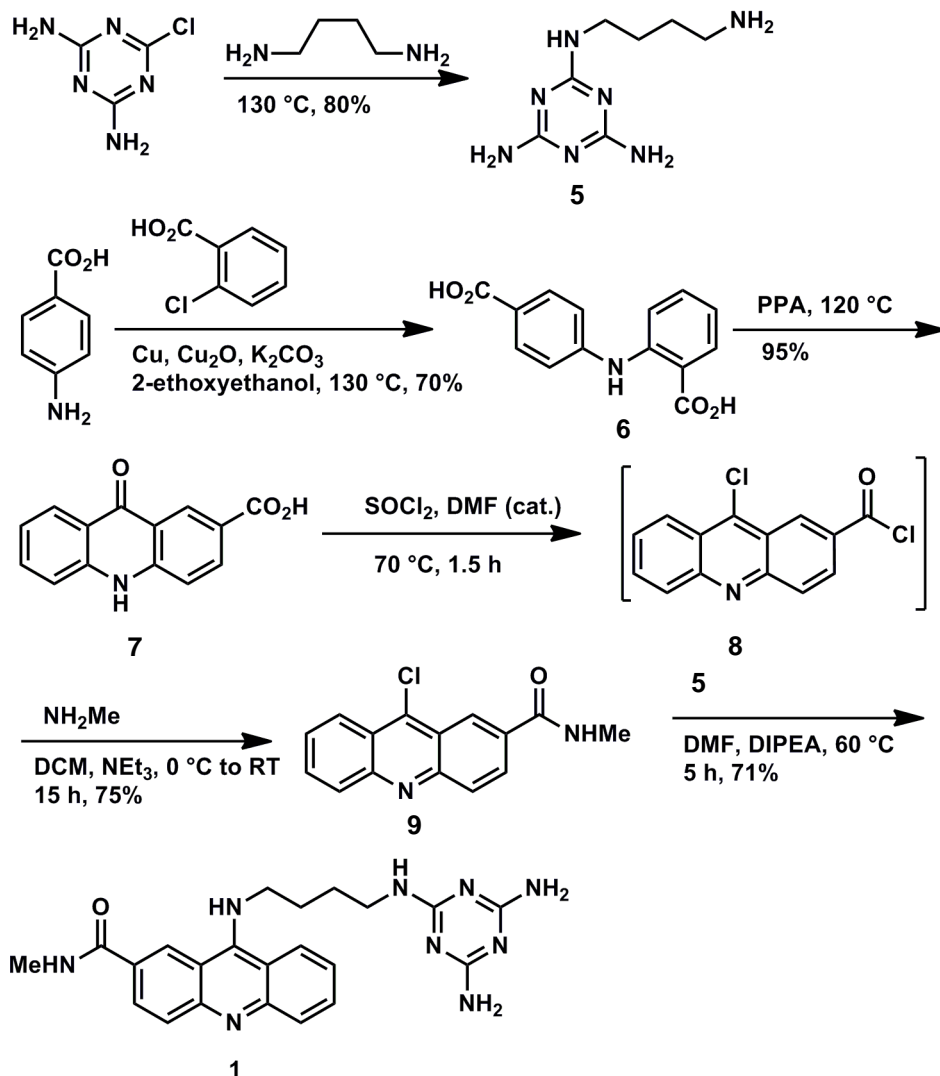
(CUG)<sub>4</sub> construct: 5'- GCUGCUGUUCGCUGCUG-TEG-Biotin - 3'

(CUG)<sub>12</sub> construct:

5'- GCCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGCUGGC-TEG-Biotin – 3'

**Isothermal Titration Calorimetry.** ITC measurements were performed at 25 °C on a MicroCal VP-ITC (MicroCal). Experiment consisted of titrating 10 µL of a 5 µM ligand from a 280 µL syringe (rotating at 300 rpm), for a total of 28 injections, into a sample cell containing 1.42 mL of a 0.2 µM MBNL1N solution. The duration of the injection was set to 24 s, and the delay between injections was 300 s. The initial delay before the first injection was 60 s. To derive the heat associated with each injection, the area under each isotherm (microcalories per second versus seconds) was determined by integration by the graphing program Origin 5.0 (MicroCal). The stock solution of **1** and **2** were 10mM in DMSO and water, respectively. The buffer solution for ITC experiments was 1X PBS, pH 7.0. 5% DMSO was added to the buffer for experiments with **1** to balance the residual DMSO in the ligand solution.

## 4.7 Experimental Synthetic Procedures



**Figure 4.21.** Synthesis of 9-((4-((4,6-Diamino-1,3,5-triazin-2-yl)amino)butyl)amino)-*N*-methylacridine-2-carboxamide (1).

***N*2-(4-Aminobutyl)-1,3,5-triazine-2,4,6-triamine (5).** Title compound was prepared as described previously,<sup>17</sup> in 80% yield. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 6.40 (t, *J* = 5.6 Hz, 1H), 6.05 (s, 2H), 5.88 (s, 2H), 3.14 (q, *J* = 6.3 Hz, 2H), 2.54 – 2.52 (m, 2H), 1.44 (p, *J* = 7.1 Hz, 2H), 1.37 – 1.31 (m, 2H).

**2-((4-Carboxyphenyl)amino)benzoic acid (6).** Title compound was prepared as described previously,<sup>24</sup> with minor changes in the work-up procedure, in 70% yield. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 9.87 (bs, 1H), 7.94 (d, *J* = 8.0 Hz, 1H), 7.87 (d, *J* = 8.7 Hz,

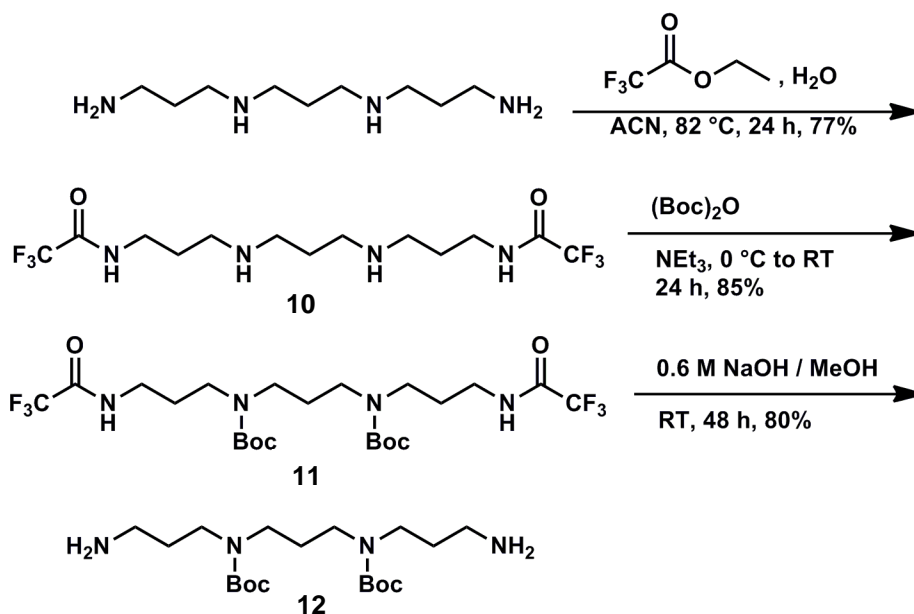
2H), 7.50 – 7.46 (m, 2H), 7.27 (d, J = 8.7 Hz, 2H), 6.96 – 6.91 (m, 1H).; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 258.1; found 258.1.

**9-Oxo-9,10-dihydroacridine-2-carboxylic acid (7).** Title compound was prepared as described previously,<sup>25</sup> with minor changes in the work-up procedure, in 95% yield. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 12.90 (bs, 1H), 12.06 (s, 1H), 8.83 (d, J = 2.0 Hz, 1H), 8.22 (ddd, J = 17.7, 8.4, 1.6 Hz, 2H), 7.8 – 7.76 (m, 1H), 7.6 – 7.57 (m, 2H), 7.36 – 7.27 (m, 1H); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 177.00, 167.41, 143.13, 140.90, 134.12, 133.88, 133.62, 128.38, 126.10, 121.76, 120.85, 119.69, 117.69, 117.38; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 240.1; found 240.1.

**9-Chloro-N-methylacridine-2-carboxamide (9).** A round-bottom flask, equipped with a stir bar, was charged with **7** (375 mg, 1.6 mmol, 1 equiv.) and freshly distilled thionyl chloride (2 mL, 27.5 mmol, 17.2 equiv.). A catalytic amount of DMF was added and the solution was heated gently under reflux at 69 °C, stirring until homogeneous and then for 1 h. The excess thionyl chloride was distilled off and the last traces were removed azeotropically via coevaporation with DCM (3 x 50 mL). The mixture was left under vacuum (minimally) for 1 hr to afford the crude intermediate, **8**, as a yellow powder. <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ 9.48 (d, J = 1.5 Hz, 1H), 9.25 (d, J = 9.2 Hz, 1H), 9.16 (d, J = 8.8 Hz, 1H), 8.72 – 8.62 (m, 2H), 8.37 – 8.30 (m, 1H), 8.06 – 7.99 (m, 1H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 276.0; found 276.0. The crude intermediate, without further purification, was dissolved in anhydrous DCM. Anhydrous triethylamine was added to the solution until the pH = 11 and it was cooled to 0 °C. A 2M solution of methylamine in methanol (0.9 mL, 1.8 mmol, 1.1 equiv.) was added and the solution was stirred at 0 °C for 2 h and then slowly warmed to room temperature overnight. The solvent was removed by rotary evaporation and the crude mixture was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH, 98:2 to 95:5) to yield **9** as a yellow solid (325 mg, 1.2 mmol, 75%). <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ 8.79 (d, J = 1.6 Hz, 1H), 8.40 (d, J = 8.6 Hz, 1H), 8.24 – 8.19 (m, 2H), 8.14 (dd, J = 9.0, 1.8 Hz, 1H), 7.86 – 7.83 (m, 1H), 7.69 – 7.62 (m, 1H), 6.56 (bs, 1H), 3.13 (d, J = 4.8 Hz, 3H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 271.1; found 271.1.

**9-((4-((4,6-Diamino-1,3,5-triazin-2-yl)amino)butyl)amino)-N-methylacridine-2-carbox-amide (1).** A round-bottom flask, equipped with a stir bar, was charged with **9** (290 mg, 1.07 mmol, 1 equiv.) and **5** (233 mg, 1.18 mmol, 1.1 equiv.). DIPEA (305 mg, 2.36 mmol, 2.2 equiv.) and anhydrous DMF (20 mL) were added. The solution was heated at 70 °C for 5 h. The solvent was removed by rotary evaporation and the product

was purified via flash chromatography (basic alumina; DCM:methanol:NH<sub>4</sub>OH, 95:4.9:0.1 to 85:14:1) to yield a yellow solid (377 mg, 0.76 mmol, 71%). <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>) δ 8.91 (s, 1H), 8.33 (d, *J* = 8.6 Hz, 1H), 8.02 (dd, *J* = 9.1, 1.9 Hz, 1H), 7.85 (t, *J* = 7.6 Hz, 2H), 7.75 – 7.68 (m, 1H), 7.42 – 7.35 (m, 1H), 4.03 (t, *J* = 7.2 Hz, 2H), 3.37 – 3.33 (m, 2H), 3.00 (s, 3H), 1.93 (p, *J* = 7.4 Hz, 2H), 1.70 (p, *J* = 7.2 Hz, 2H); *m/z* HRMS (ESI) calculated for [M+H]<sup>+</sup>: 432.2260; found 432.2267.



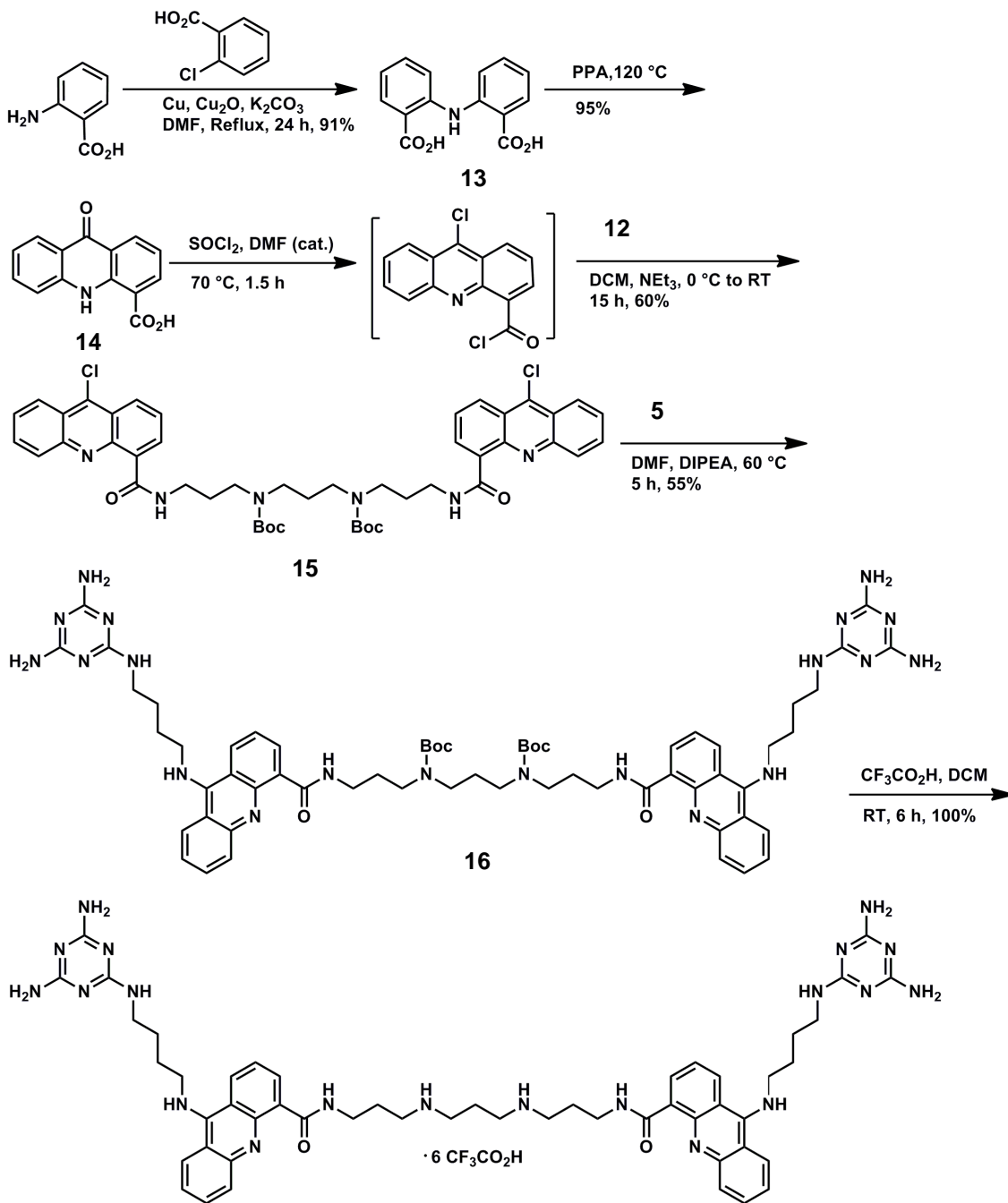
**Figure 4.22.** Synthesis of Di-*tert*-Butyl propane-1,3-diyldis((3-aminopropyl)carbamate) (**12**).

***N,N'*-((Propane-1,3-diyldis(azanediyl))bis(propane-3,1-diyldis(2,2,2-trifluoroacetamido))** (**10**). The title compound was prepared from *N,N'*-(propane-1,3-diyldis(propane-1,3-diamine) as described previously,<sup>4</sup> with minor changes, in 77% yield. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 9.61 (s, 2H), 8.75 (bs, 2H), 3.26 (q, *J* = 6.5 Hz, 4H), 3.02 – 2.95 (m, 4H), 2.94 – 2.88 (m, 4H), 1.91 (p, *J* = 7.8 Hz, 2H), 1.82 (p, *J* = 7.0 Hz, 4H); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 156.40, 117.08, 44.63, 44.02, 36.61, 25.22, 22.56; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 381.2; found 381.2.

**Di-*tert*-Butyl propane-1,3-diyldis((3-(2,2,2-trifluoroacetamido)propyl)carbamate)** (**11**). The title compound was prepared as described previously,<sup>4</sup> with minor changes. The product was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH, 98:2 to 95:5) to afford **11** in 85% yield. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 9.42 (bs, 2H), 3.18 – 3.07 (m, 12H), 1.71 – 1.60 (m, 6H), 1.37 (s, 18H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 581.3; found 581.3.



**Di-*tert*-Butyl propane-1,3-diylbis((3-aminopropyl)carbamate (12).** The title compound was prepared as described previously,<sup>26</sup> with minor changes. The product was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH:NH<sub>4</sub>OH, 80:19:1 to 67:30:2) to afford S8 in 80% yield. <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ 3.28 (s, 4H), 3.16 (s, 4H), 2.69 (t, J = 6.4 Hz, 4H), 1.75 (s, 2H), 1.64 (p, J = 6.8 Hz, 4H), 1.45 (s, 18H); <sup>13</sup>C NMR (126 MHz, chloroform-*d*) δ 155.29, 79.06, 44.44, 43.68, 38.96, 32.34, 31.73, 28.17; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 389.3; found 389.3.



**2,2'-Iminodibenzoic (13).** A round-bottom flask, equipped with a stir bar, was charged with 2-bromobenzoic acid (3.00 g, 14.9 mmol, 1 equiv.), anthranilic acid (2.25 g, 16.4 mmol, 1.1 equiv.), K<sub>2</sub>CO<sub>3</sub> (6.00 g, 44 mmol, 2.95 equiv.), Cu (0.20 g, 2.9 mmol, 0.2 equiv.) and Cu<sub>2</sub>O (0.22 g, 1.45 mmol, 0.1 equiv.). To this mixture, anhydrous DMF (50 mL) was added and it was stirred at reflux at 130 °C for 16 h. The work-up procedure was carried out as reported previously,<sup>27</sup> to yield 2,2'-iminodibenzoic acid in 91% yield as a light green solid. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 13.06 (bs, 1H), 10.85 (bs, 1H), 7.91 (s, 2H), 7.46 (s, 4H), 6.95 (s, 2H); <sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>) δ 168.41, 143.58, 133.38, 131.81, 119.99, 117.56, 113.56; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 258.1; found 258.1.

**9-Oxo-9,10-dihydroacridine-4-carboxylic acid (14).** The title compound was prepared as described previously,<sup>25</sup> with minor changes in the work-up procedure, in 95% yield. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 11.96 (s, 1H), 8.53 (dd, *J* = 8.0, 1.5 Hz, 1H), 8.45 (dd, *J* = 7.5, 1.7 Hz, 1H), 8.24 (d, *J* = 8.1 Hz, 1H), 7.82 – 7.73 (m, 2H), 7.39 – 7.30 (m, 2H); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 176.53, 169.14, 141.20, 139.92, 136.90, 134.11, 132.41, 125.89, 122.32, 121.63, 120.60, 120.24, 118.63, 115.01; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 240.1; found 240.1.

**Di-*tert*-Butylpropane-1,3-diylbis((3-(9-chloroacridine-4-carboxamido)propyl) carbamate) (15).** A round-bottom flask equipped with a stir bar was charged with **14** (600 mg, 2.5 mmol, 1 equiv.) and freshly distilled thionyl chloride (3 mL, 41 mmol, 16.4 equiv.). A catalytic amount of DMF was added and heated gently under reflux at 69 °C, stirring until homogeneous and then for 1 h. The excess thionyl chloride was distilled off and the last traces removed azeotropically via coevaporation with DCM (3 x 50 mL). It was left under vacuum for (minimally) 1 h to afford the crude intermediate as a yellow powder. The crude intermediate was dissolved in anhydrous DCM. Anhydrous triethylamine was added to the solution until the pH was 11 and it was cooled to 0 °C. **12** (437 mg, 1.125 mmol, 0.45 equiv.) was added and the solution was stirred at 0 °C for 2 h and slowly warmed to room temperature overnight. The solvent was removed by rotary evaporation and the crude mixture was purified via flash chromatography (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>:MeOH, 98:2 to 95:5) to yield a yellow solid (586 mg, 0.675 mmol, 60%). <sup>1</sup>H NMR (500 MHz, chloroform-*d*) δ 9.00 (s, 2H), 8.36 (d, *J* = 14.0 Hz, 4H), 8.27 (s, 2H), 8.20 (d, *J* = 16.2 Hz, 4H), 7.80 (s, 2H), 7.65 – 7.60 (m, 2H), 3.54 (s, 4H), 3.46 (d, *J* = 12.0 Hz, 4H), 3.23 (s, 4H), 1.86 (dd, *J* = 14.5, 7.5 Hz, 6H), 1.52 (s, 18H); <sup>13</sup>C NMR (125 MHz,

chloroform-*d*)  $\delta$  166.30, 156.65, 149.56, 149.23, 142.50, 132.76, 131.12, 130.12, 129.75, 128.54, 127.13, 124.59, 124.33, 123.26, 80.31, 44.94, 43.54, 36.14, 28.52, 27.87, 27.63; *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 867.3; found 867.3.

**Di-*tert*-Butylpropane-1,3-diylbis((3-(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamido)propyl)carbamate) (16).** A round-bottom flask, equipped with a stir bar, was charged with **15** (500 mg, 0.576 mmol, 1 equiv.) and *N*2-(4-Aminobutyl)-1,3,5-triazine-2,4,6-triamine (250 mg, 1.27 mmol, 2.2 equiv.). DIPEA (327 mg, 2.53 mmol, 4.4 equiv.) and anhydrous DMF (25 mL) were added. The solution was heated at 70 °C for 5 h. The solvent was removed by rotary evaporation and the product purified via flash chromatography (basic alumina; DCM:methanol:NH<sub>4</sub>OH, from 95:4.9:0.1 to 90:9.5:0.5) to yield a yellow solid (377 mg, 0.317 mmol, 55%). <sup>1</sup>H NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  8.60 (s, 2H), 8.37 (d, *J* = 8.5 Hz, 2H), 8.20 (d, *J* = 8.5 Hz, 2H), 7.96 – 7.76 (m, 2H), 7.63 (s, 2H), 7.37 – 7.27 (m, 4H), 3.82 (s, 4H), 3.49 (s, 4H), 3.35 (s, 4H), 3.27 (t, *J* = 6.8 Hz, 4H), 3.19 (s, 4H), 1.94 – 1.76 (m, 10H), 1.59 (d, *J* = 7.0 Hz, 4H), 1.34 (d, *J* = 35.9 Hz, 18H); *m/z* LRMS (ESI) calculated for [M+H]<sup>+</sup>: 1189.7; found 1189.7.

***N,N'*-((Propane-1,3-diylbis(azanediyl))bis(propane-3,1-diyl))bis(9-((4-((4,6-diamino-1,3,5-triazin-2-yl)amino)butyl)amino)acridine-4-carboxamide) (2).** A round-bottom flask, equipped with a stir bar, was charged with **16** (310 mg, 0.261 mmol, 1 equiv.). TFA (30 mL) and anhydrous DCM (70 mL) was added and the mixture stirred at room temperature for 2 h. The solvents were removed to afford a yellow solid (437 mg, 0.261 mmol, 100%). <sup>1</sup>H NMR (500 MHz, deuterium oxide)  $\delta$  8.34 (d, *J* = 8.3 Hz, 2H), 8.17 (d, *J* = 8.3 Hz, 2H), 8.11 (d, *J* = 7.4 Hz, 2H), 7.78 (t, *J* = 7.7 Hz, 2H), 7.52 (d, *J* = 8.3 Hz, 2H), 7.48 – 7.39 (m, 4H), 4.07 (t, *J* = 6.8 Hz, 4H), 3.58 (t, *J* = 6.7 Hz, 4H), 3.37 – 3.29 (m, 4H), 3.26 (t, *J* = 5.9 Hz, 8H), 2.31 (s, 2H), 2.14 – 2.08 (m, 4H), 1.98 – 1.91 (m, 4H), 1.71 – 1.65 (m, 4H); Elemental analysis, Calculated for C<sub>63</sub>H<sub>70</sub>F<sub>18</sub>N<sub>20</sub>O<sub>14</sub>: C, 45.22 %; H, 4.22 %; F, 20.44 %; N, 16.74 %; Found: C, 45.15 %; H, 4.20 %; F, 20.09 %; N, 16.46 %, *m/z* HRMS (ESI) calculated for [M+H]<sup>+</sup>: 989.5599; Found 989.5590.

## 4.8 Acknowledgements

I would like to thank Professor Maria Spies (University of Iowa) for the single molecule collaboration, Dr. Masayoshi Honda and for performing the single molecule experiments, Dr. Yuan Fu and Kali A. Miller for their assistance with MBNL1 expression, purification and the synthetic procedures.

## 4.9 References

- (1) Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., and Thornton, C. A. (2009) Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA, *Science* 325, 336-339.
- (2) O'Rourke, J. R., and Swanson, M. S. (2009) Mechanisms of RNA-mediated disease, *J. Biol. Chem.* 284, 7419-7423.
- (3) Echeverria, G. V., and Cooper, T. A. (2012) RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity, *Brain Res.* 1462, 100-111.
- (4) Wheeler, T. M., Leger, A. J., Pandey, S. K., MacLeod, A. R., Nakamori, M., Cheng, S. H., Wentworth, B. M., Bennett, C. F., and Thornton, C. A. (2012) Targeting nuclear RNA for in vivo correction of myotonic dystrophy, *Nature* 488, 111-115.
- (5) Garcia-Lopez, A., Llamusi, B., Orzaez, M., Perez-Paya, E., and Artero, R. D. (2011) In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models, *Proc. Natl. Acad. Sci. U. S. A.* 108, 11866-11871.
- (6) Foff, E. P., and Mahadevan, M. S. (2011) Therapeutics development in myotonic dystrophy type 1, *Muscle Nerve* 44, 160-169.
- (7) Warf, M. B., Nakamori, M., Matthys, C. M., Thornton, C. A., and Berglund, J. A. (2009) Pentamidine reverses the splicing defects associated with myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A.* 106, 18551-18556.
- (8) Ofori, L. O., Hoskins, J., Nakamori, M., Thornton, C. A., and Miller, B. L. (2012) From dynamic combinatorial 'hit' to lead: in vitro and in vivo activity of compounds targeting the pathogenic RNAs that cause myotonic dystrophy, *Nucleic Acids Res.* 40, 6380-6390.
- (9) Parkesh, R., Childs-Disney, J. L., Nakamori, M., Kumar, A., Wang, E., Wang, T., Hoskins, J., Tran, T., Housman, D., Thornton, C. A., and Disney, M. D. (2012) Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching, *J. Am. Chem. Soc.* 134, 4731-4742.
- (10) Pushechnikov, A., Lee, M. M., Childs-Disney, J. L., Sobczak, K., French, J. M., Thornton, C. A., and Disney, M. D. (2009) Rational design of ligands targeting

- triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3, *J. Am. Chem. Soc.* **131**, 9767-9779.
- (11) Holdgate, G. A., and Gill, A. L. (2011) Kinetic efficiency: the missing metric for enhancing compound quality?, *Drug Discov. Today* **16**, 910-913.
- (12) Blow, N. (2009) Proteins and proteomics: life on the surface, *Nat. Methods* **6**, 389-393.
- (13) Navratilova, I., and Myszka, D. G. (2006) Investigating Biomolecular Interactions and Binding Properties Using SPR Biosensors, *Surface Plasmon Resonance Based Sensors* **4**, 155-176.
- (14) Elenko, M. P., Szostak, J. W., and van Oijen, A. M. (2009) Single-molecule imaging of an in vitro-evolved RNA aptamer reveals homogeneous ligand binding kinetics, *J. Am. Chem. Soc.* **131**, 9866-9867.
- (15) Moffitt, J. R., Chemla, Y. R., and Bustamante, C. (2010) Methods in statistical kinetics, *Methods Enzymol.* **475**, 221-257.
- (16) Karunatilaka, K. S., Solem, A., Pyle, A. M., and Rueda, D. (2010) Single-molecule analysis of Mss116-mediated group II intron folding, *Nature* **467**, 935-939.
- (17) Arambula, J. F., Ramisetty, S. R., Baranger, A. M., and Zimmerman, S. C. (2009) A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16068-16073.
- (18) Mammen, M., Choi, S. K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors, *Angew. Chem. Int. Ed.* **37**, 2755-2794.
- (19) Childs-Disney, J. L., Hoskins, J., Rzuczek, S. G., Thornton, C. A., and Disney, M. D. (2012) Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive, *ACS Chem. Biol.* **7**, 856-862.
- (20) Wong, C. H., Richardson, S. L., Ho, Y. J., Lucas, A. M., Tuccinardi, T., Baranger, A. M., and Zimmerman, S. C. (2012) Investigating the Binding Mode of an Inhibitor of the MBNL1RNA Complex in Myotonic Dystrophy Type 1 (DM1) Leads to the Unexpected Discovery of a DNA-Selective Binder, *ChemBioChem*, DOI: 10.1002/cbic.201200602.
- (21) Zimmerman, S. C., Lamberson, C. R., Cory, M., and Fairley, T. A. (1989) Topologically Constrained Bifunctional Intercalators - DNA Intercalation by a Macrocyclic Bisacridine, *J. Am. Chem. Soc.* **111**, 6805-6809.

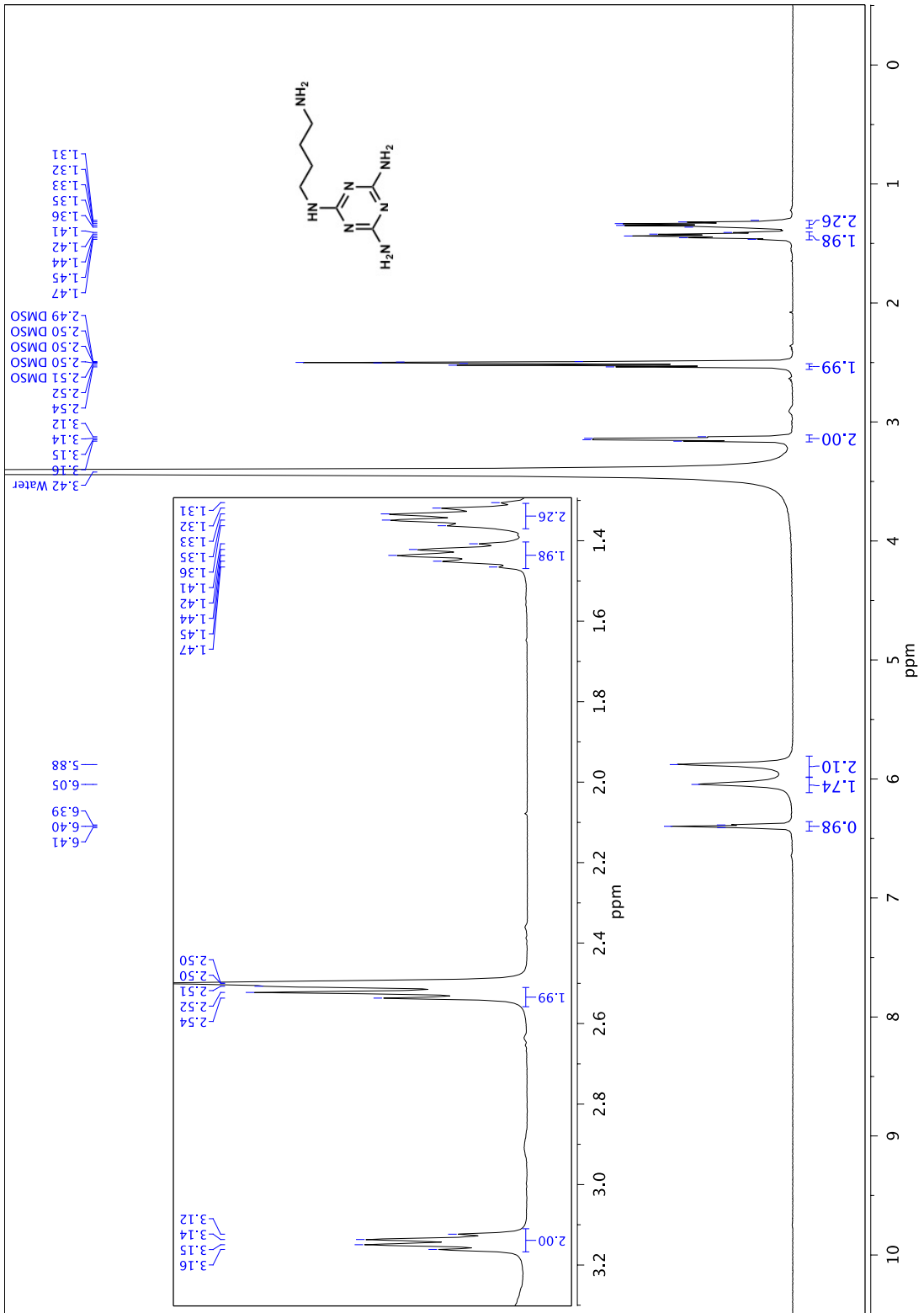
- (22) Fechter, E. J., Olenyuk, B., and Dervan, P. B. (2004) Design of a sequence-specific DNA bisintercalator, *Angew. Chem. Int. Ed. Engl.* **43**, 3591-3594.
- (23) Holman, G. G., Zewail-Foote, M., Smith, A. R., Johnson, K. A., and Iverson, B. L. (2011) A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant, *Nat. Chem.* **3**, 875-881.
- (24) Mei, X., August, A. T., and Wolf, C. (2006) Regioselective copper-catalyzed amination of chlorobenzoic acids: synthesis and solid-state structures of N-aryl anthranilic acid derivatives., *J. Org. Chem.* **71**, 142-149.
- (25) Kauffman, M. (1990) 4-(2=Carboxyphenyl)aminobenzenealkanoic, *J. Pharm. Sci.* **79**, 173-178.
- (26) Ilies, M. A., Seitz, W. a., Johnson, B. H., Ezell, E. L., Miller, A. L., Thompson, E. B., and Balaban, A. T. (2006) Lipophilic pyrylium salts in the synthesis of efficient pyridinium-based cationic lipids, gemini surfactants, and lipophilic oligomers for gene delivery., *J. Med. Chem.* **49**, 3872-3887.
- (27) Cuenca, F., Moore, M. J., Johnson, K., Guyen, B., De Cian, A., and Neidle, S. (2009) Design, synthesis and evaluation of 4,5-di-substituted acridone ligands with high G-quadruplex affinity and selectivity, together with low toxicity to normal cells, *Bioorg. Med. Chem. Lett.* **19**, 5109-5113.
- (28) Wildenberg, S. M. J. L. V. D., Prevo, B., and Peterman, E. J. G. (2011) Single Molecule Analysis, *Methods* **783**, 81-99.
- (29) Reck-Peterson, S. L., Derr, N. D., and Stuurman, N. (2010) Imaging single molecules using total internal reflection fluorescence microscopy (TIRFM), *Cold Spring Harb. Protoc.* **2010**, doi:10.1101/pdb.top1173.
- (30) Warf, M. B., and Berglund, J. A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T, *RNA* **13**, 2238-2251.
- (31) Cass, D., Hotchko, R., Barber, P., Jones, K., Gates, D. P., and Berglund, J. A. (2011) The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements, *BMC Mol. Biol.* **12**, 20.
- (32) Qin, F., Auerbach, A., and Sachs, F. (1996) Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events, *Biophys. J.* **70**, 264-280.

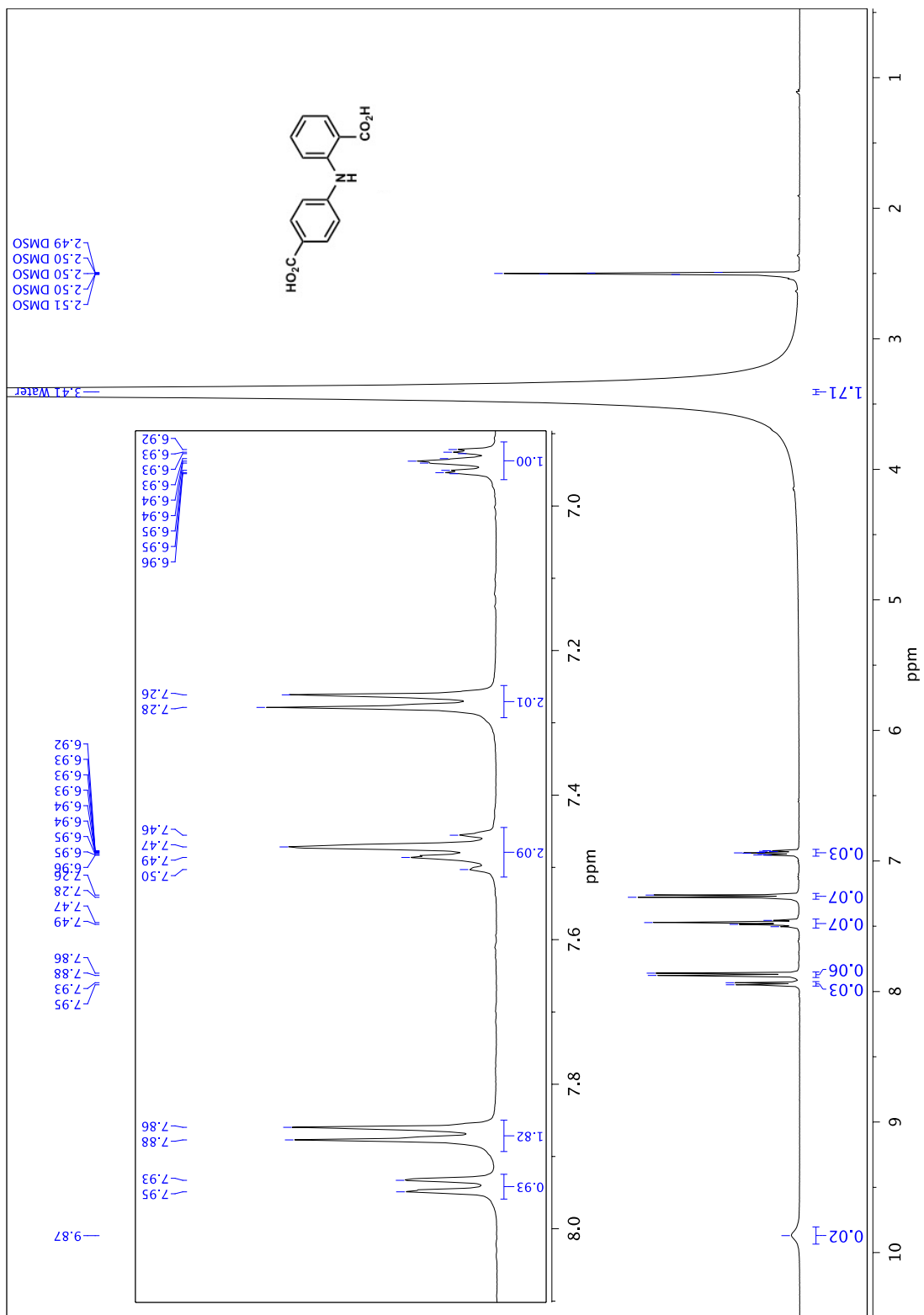
- (33) Teplova, M., and Patel, D. J. (2008) Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1, *Nat. Struct. Mol. Biol.* 15, 1343-1351.
- (34) Feng, B. Y., and Shoichet, B. K. (2006) A detergent-based assay for the detection of promiscuous inhibitors, *Nat. Protoc.* 1, 550-553.
- (35) Vamosi, G., Gohlke, C., and Clegg, R. M. (1996) Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: multiple conformers of single-stranded and double-stranded dye-DNA complexes, *Biophys. J.* 71, 972-994.
- (36) Qu, P., Chen, X. D., Zhou, X. X., Li, X., and Zhao, X. S. (2009) Fluorescence quenching of TMR by guanosine in oligonucleotides, *Sci. China, Ser. B* 52, 1653-1659.
- (37) Xu, Y., and Komiyama, M. (2012) Structure, function and targeting of human telomere RNA, *Methods* 57, 100-105.
- (38) Vicens, Q., and Westhof, E. (2003) RNA as a drug target: the case of aminoglycosides, *ChemBioChem* 4, 1018-1023.
- (39) Vicens, Q. (2009) RNA's coming of age as a drug target, *J. Incl. Phenom. Macro.* 65, 171-188.
- (40) Cooper, T. A., Wan, L., and Dreyfuss, G. (2009) RNA and disease, *Cell* 136, 777-793.
- (41) Fu, Y., Ramisetty, S. R., Hussain, N., and Baranger, A. M. (2012) MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation, *ChemBioChem* 13, 112-119.
- (42) Laurent, F. X., Sureau, A., Klein, A. F., Trouslard, F., Gasnier, E., Furling, D., and Marie, J. (2012) New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats, *Nucleic Acids Res.* 40, 3159-3171.
- (43) Stelzer, A. C., Frank, A. T., Kratz, J. D., Swanson, M. D., Gonzalez-Hernandez, M. J., Lee, J., Andricioaei, I., Markovitz, D. M., and Al-Hashimi, H. M. (2011) Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble, *Nat. Chem. Biol.* 7, 553-559.
- (44) Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., and Swanson, M. S. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs, *Nucleic Acids Res.* 35, 5474-5486.

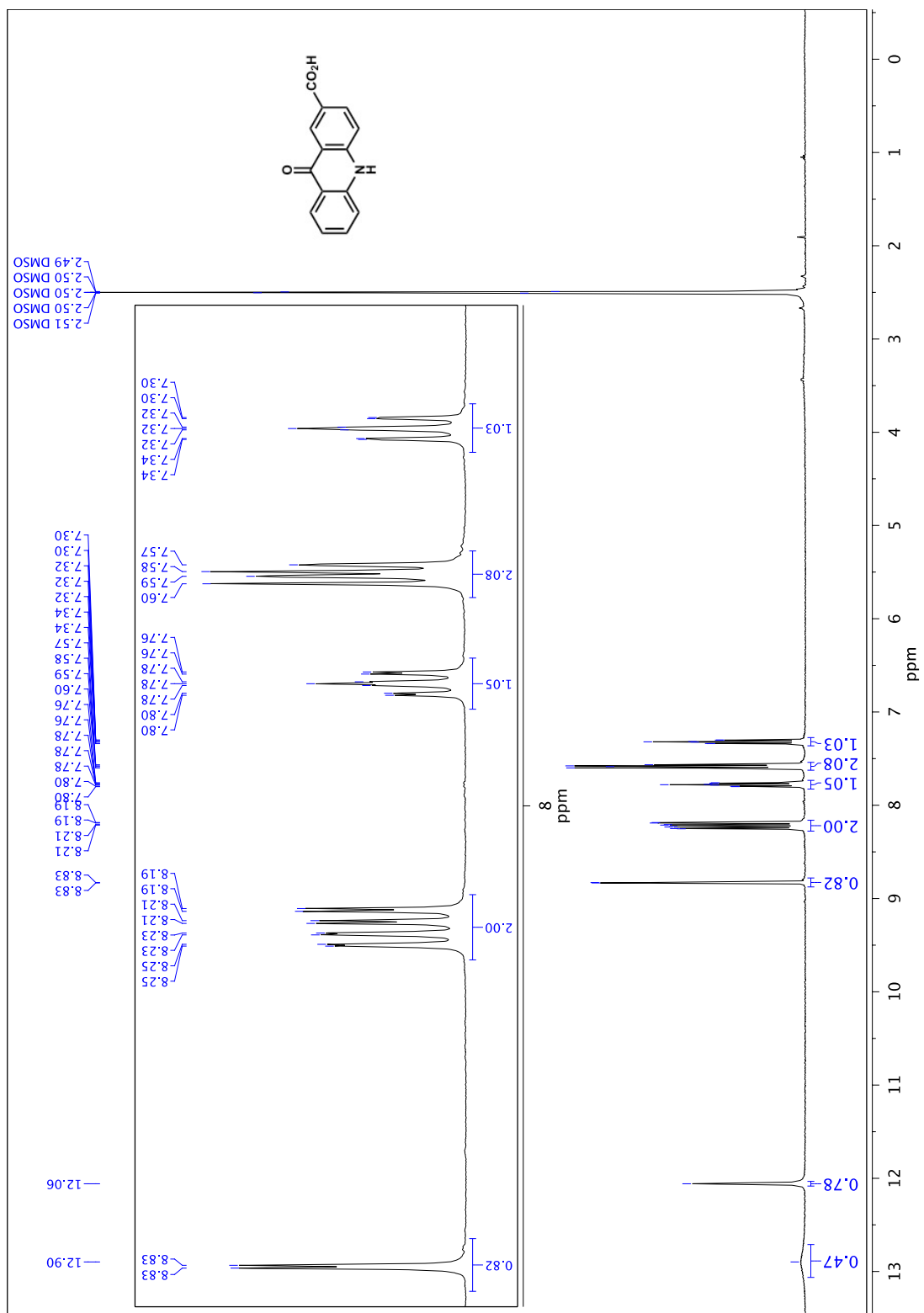
- (45) Galletto, R., Amitani, I., Baskin, R. J., and Kowalczykowski, S. C. (2006) Direct observation of individual RecA filaments assembling on single DNA molecules., *Nature* 443, 875-878.
- (46) Kim, Y. T., Tabor, S., Churchich, J. E., and Richardson, C. C. (1992) Interactions of gene 2.5 protein and DNA polymerase of bacteriophage T7., *J. Biol. Chem.* 267, 15032-15040.
- (47) Honda, M., Park, J., Pugh, R. A., Ha, T., and Spies, M. (2009) Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase, *Mol. Cell* 35, 694-703.

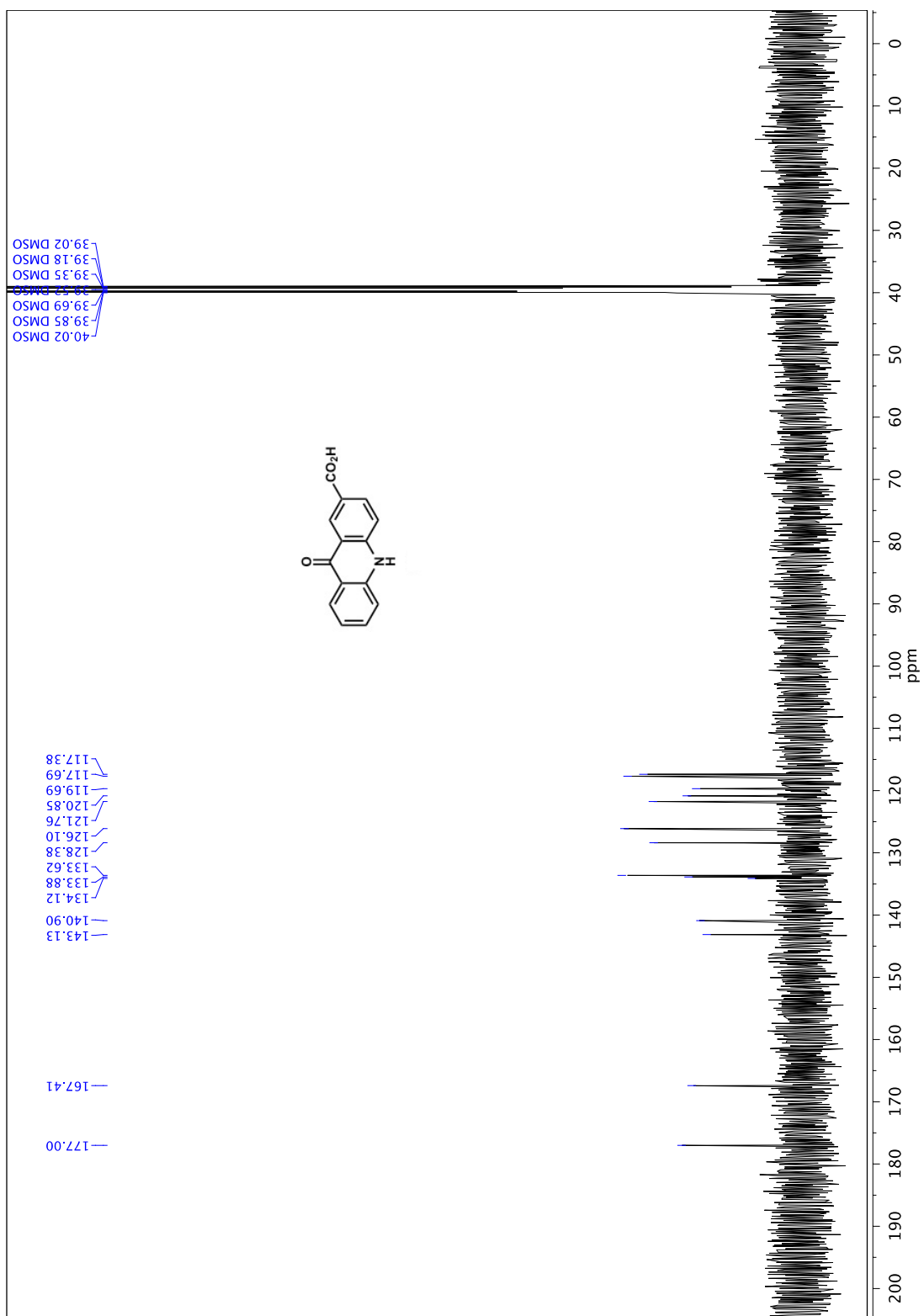
## 4.10 NMR Spectra

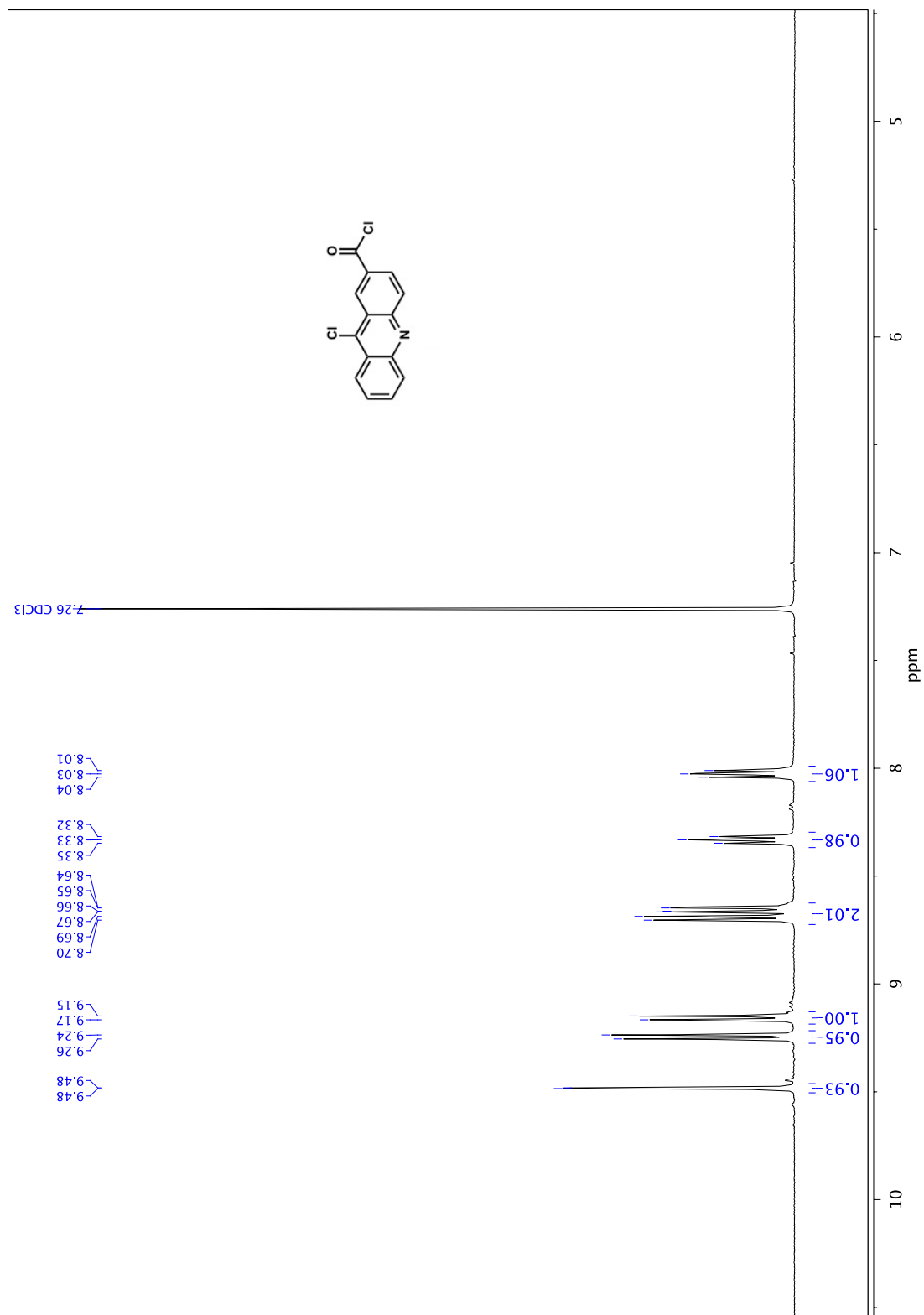


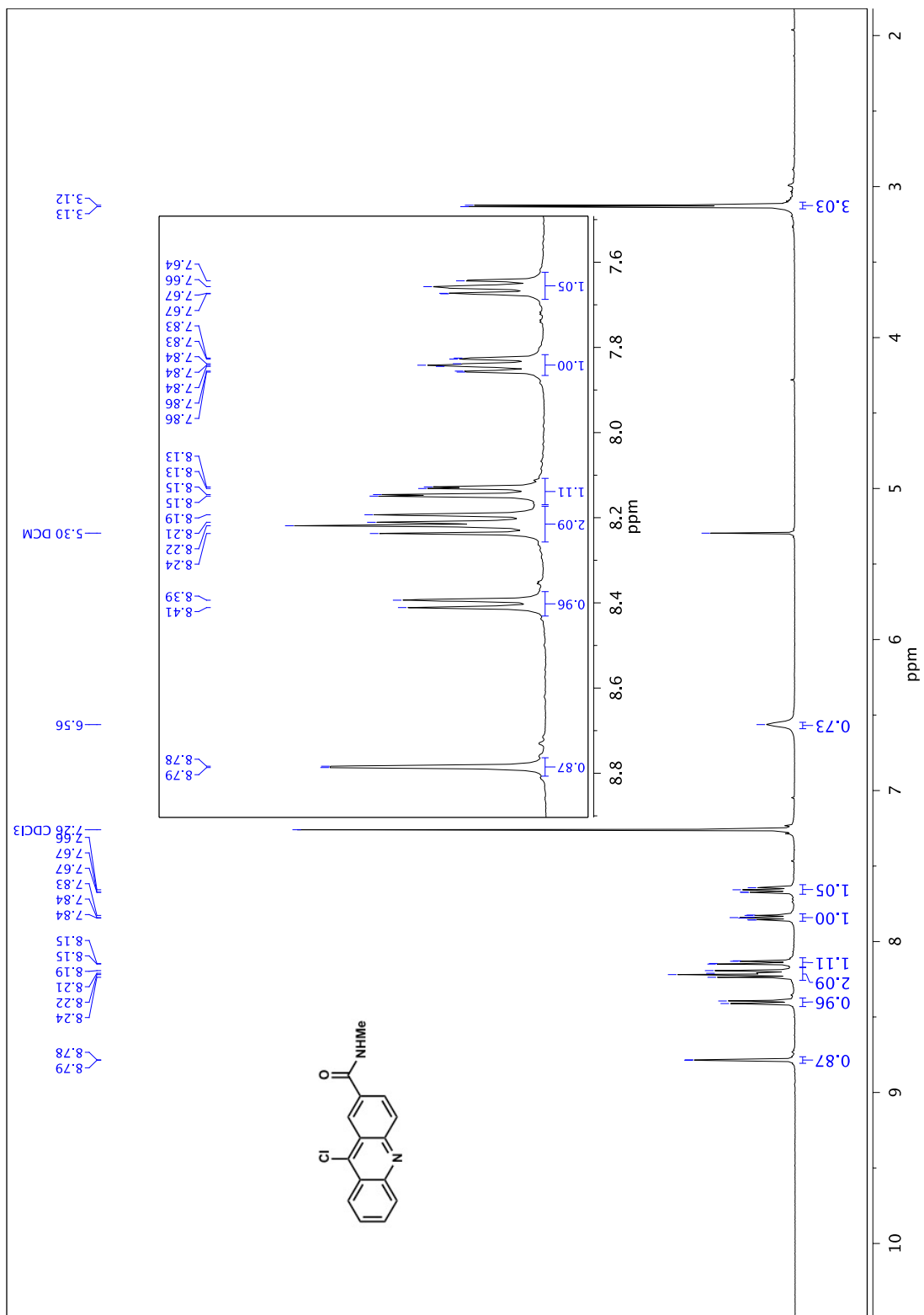


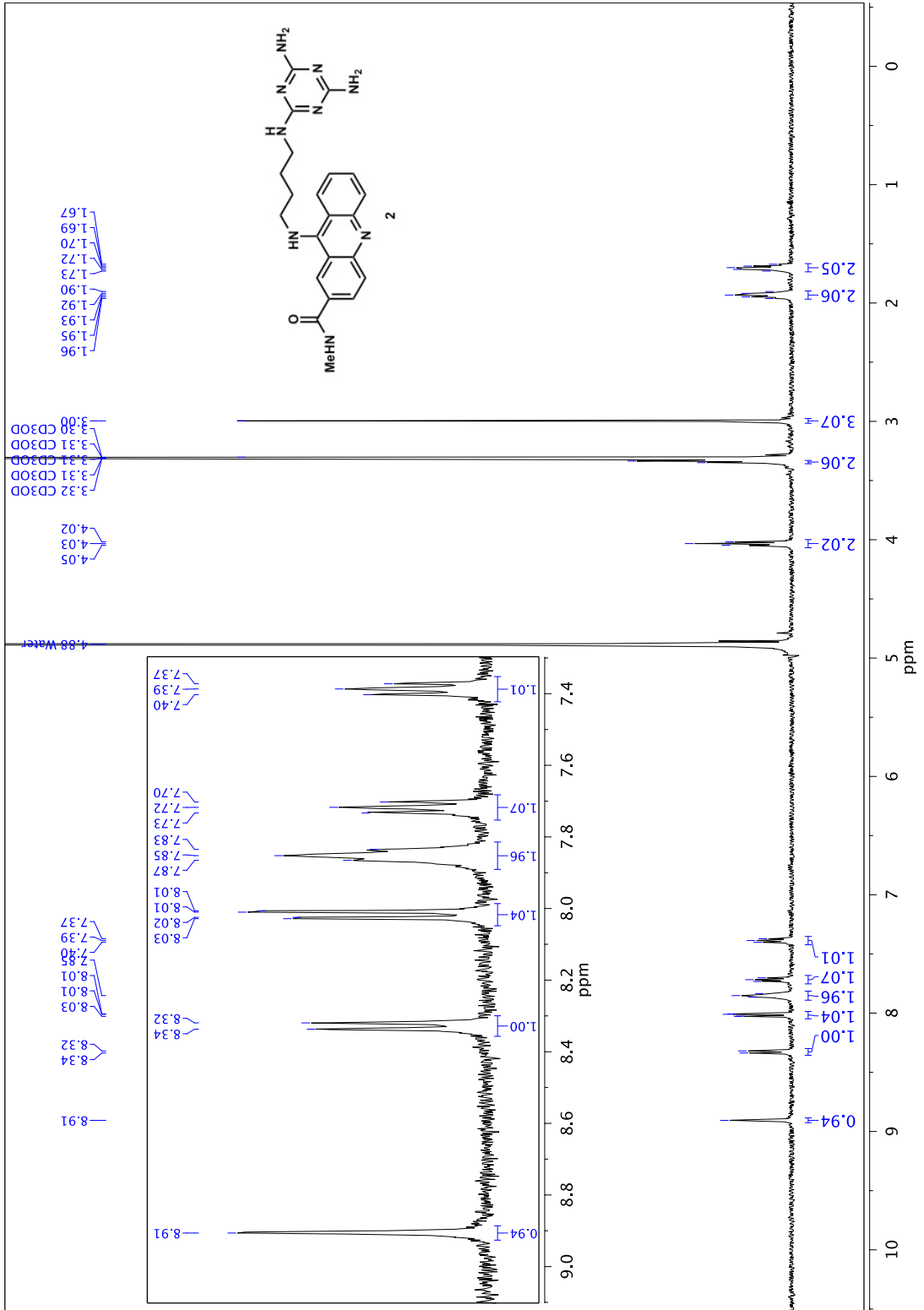


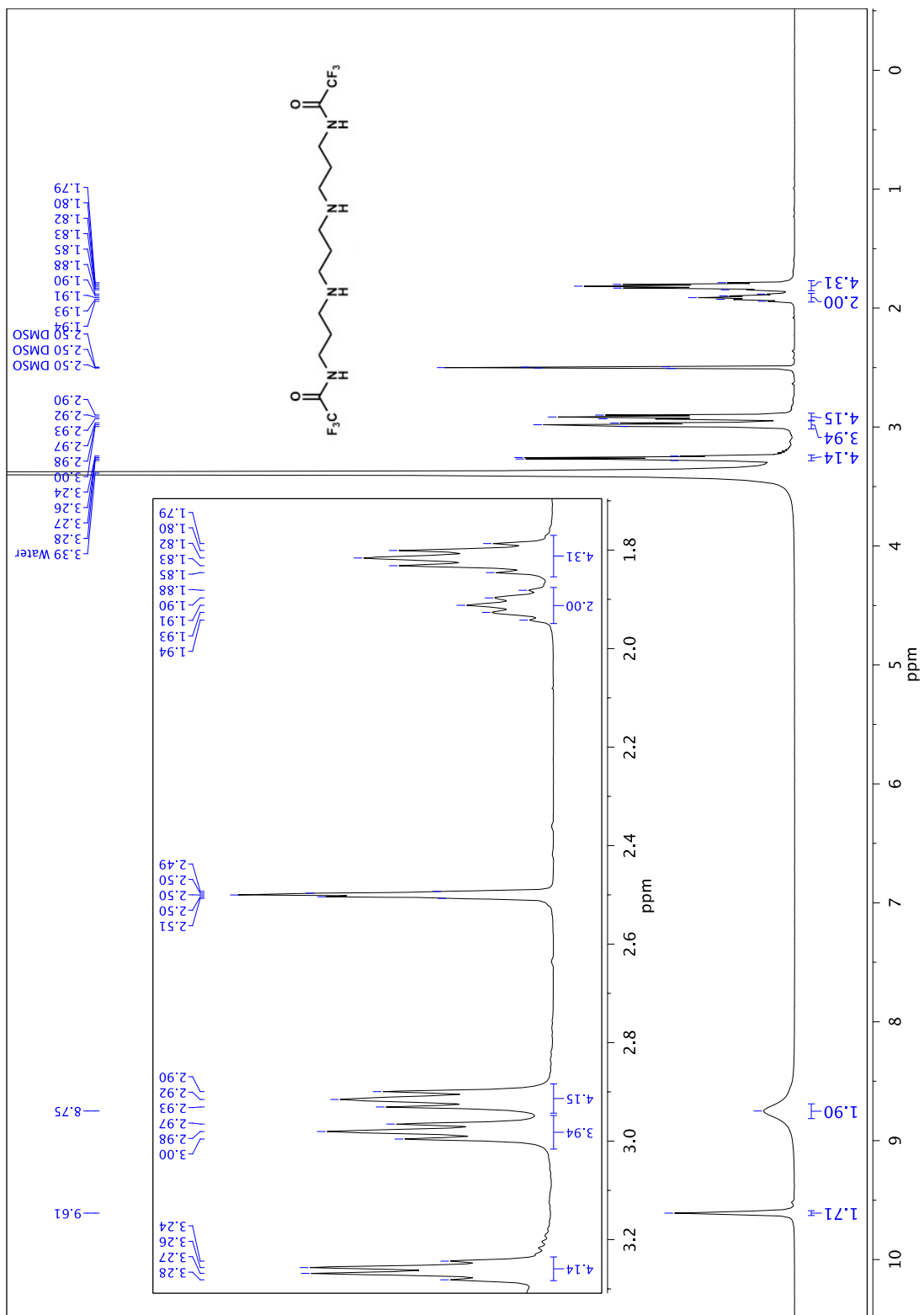




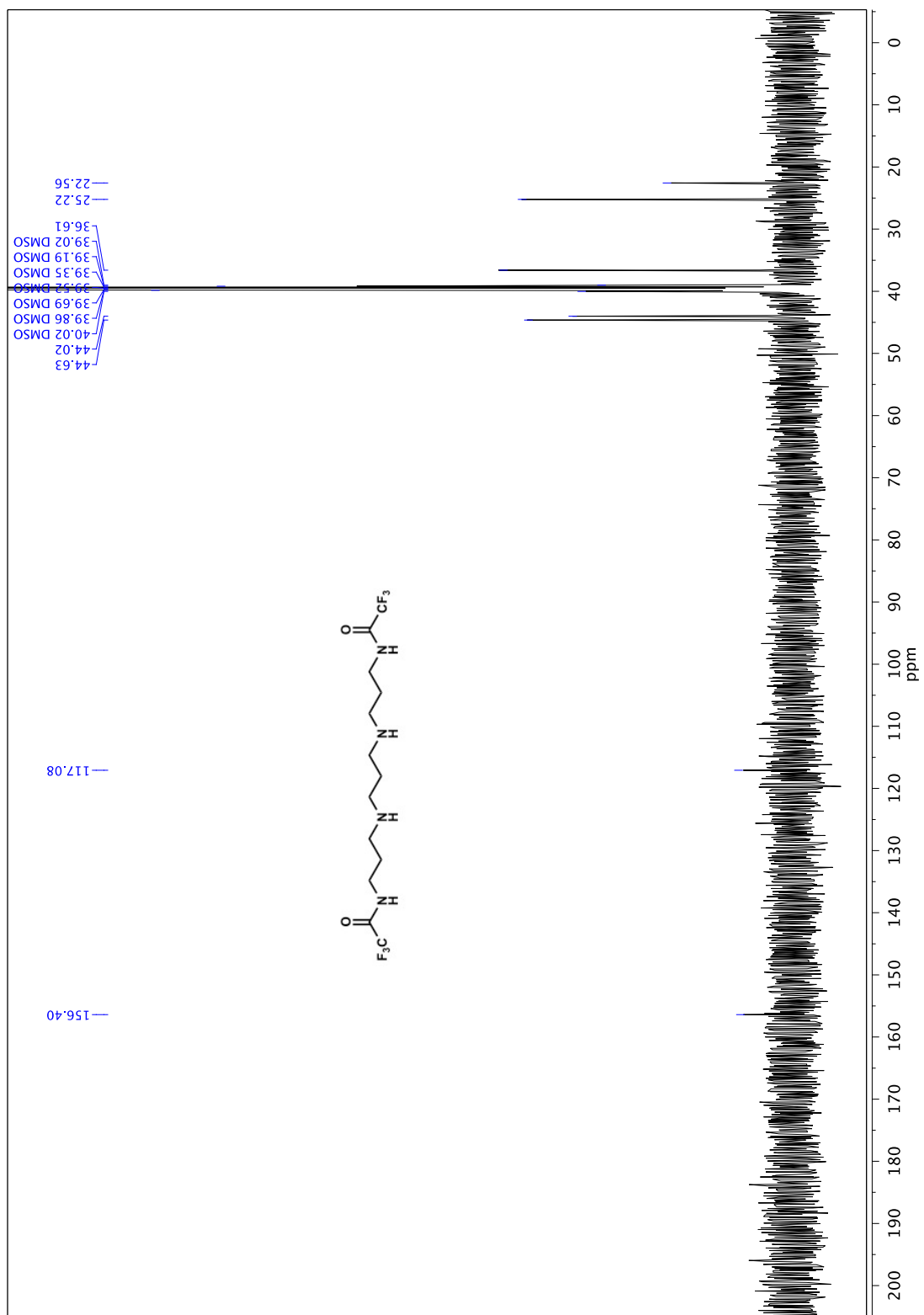


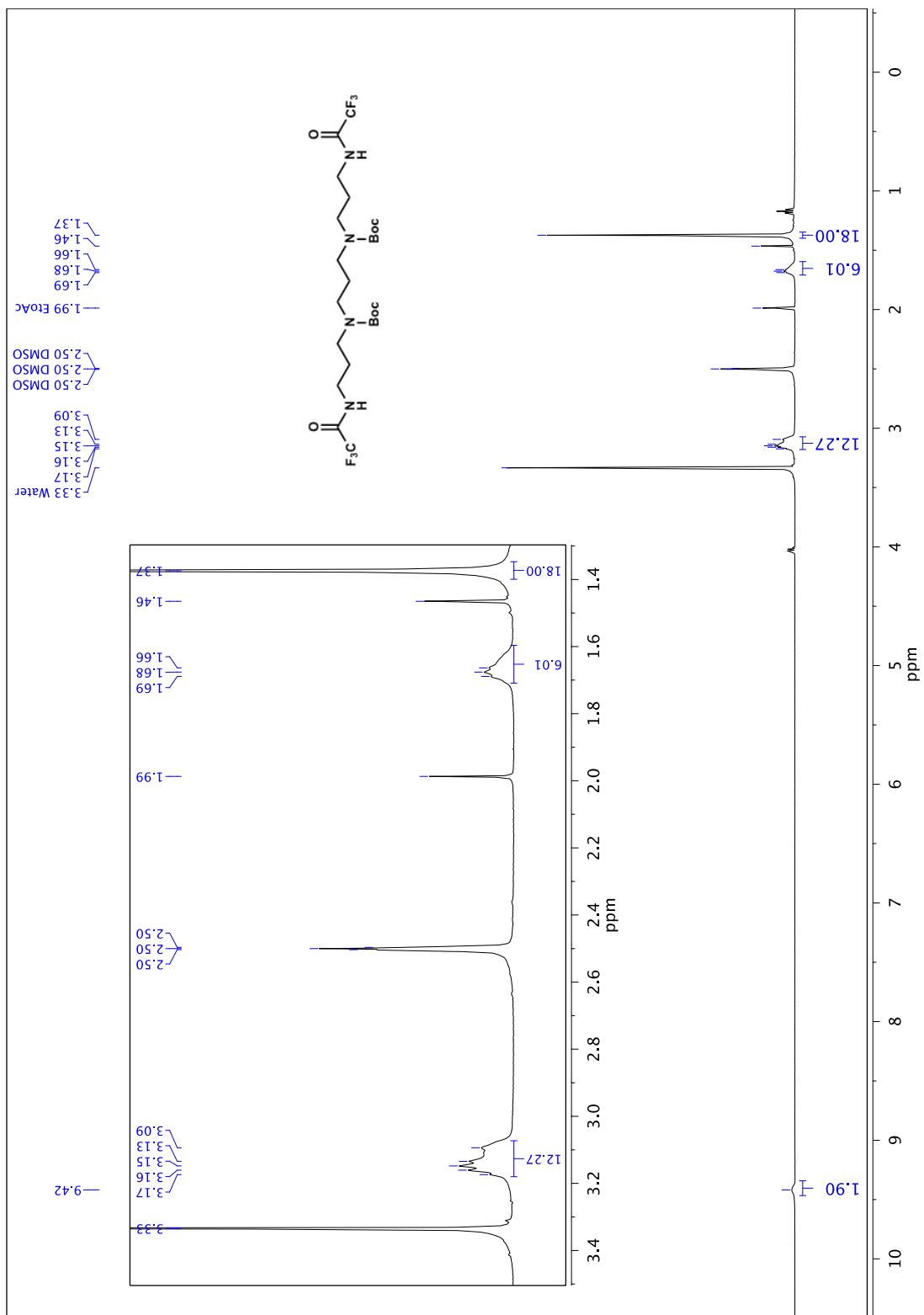




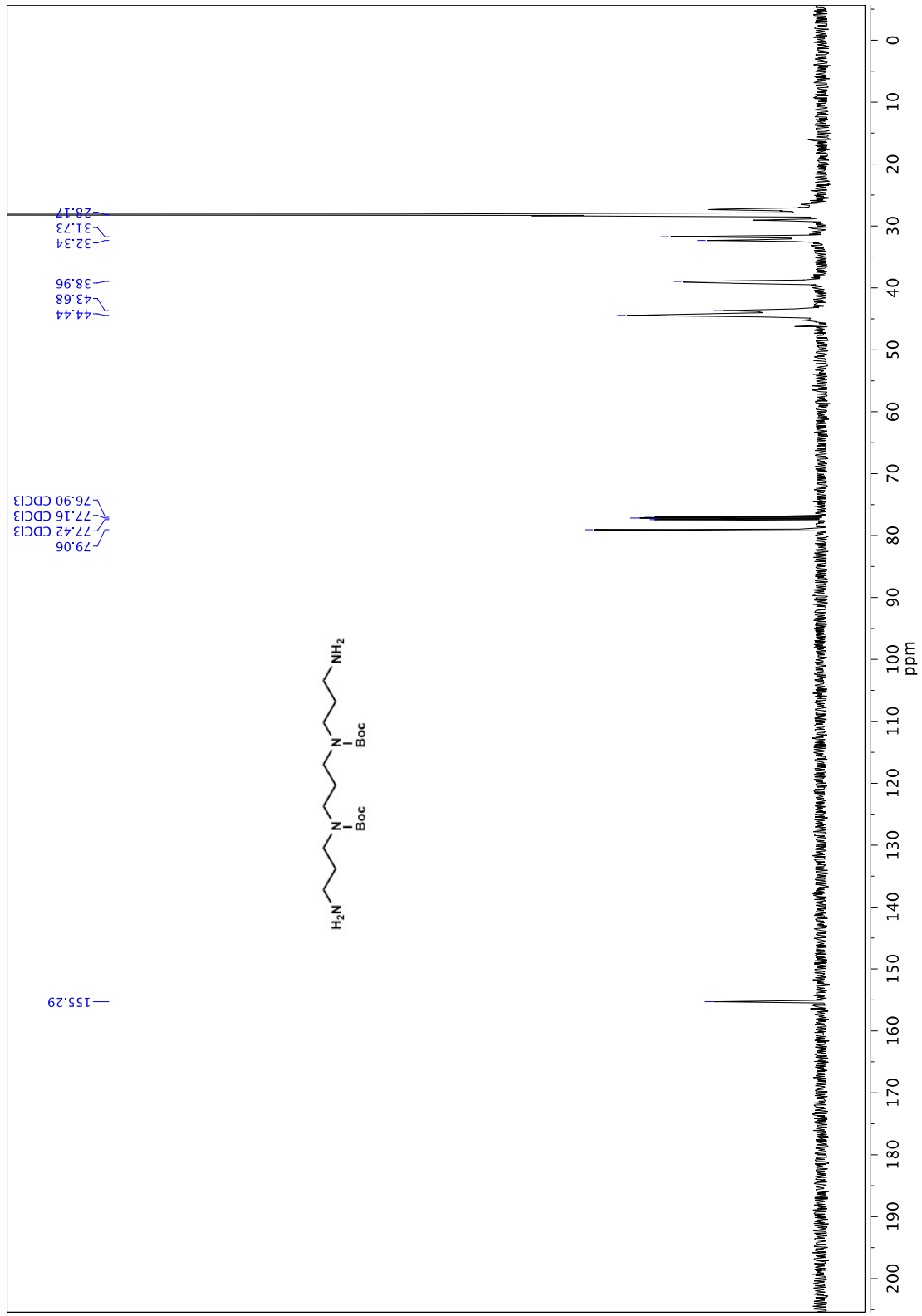


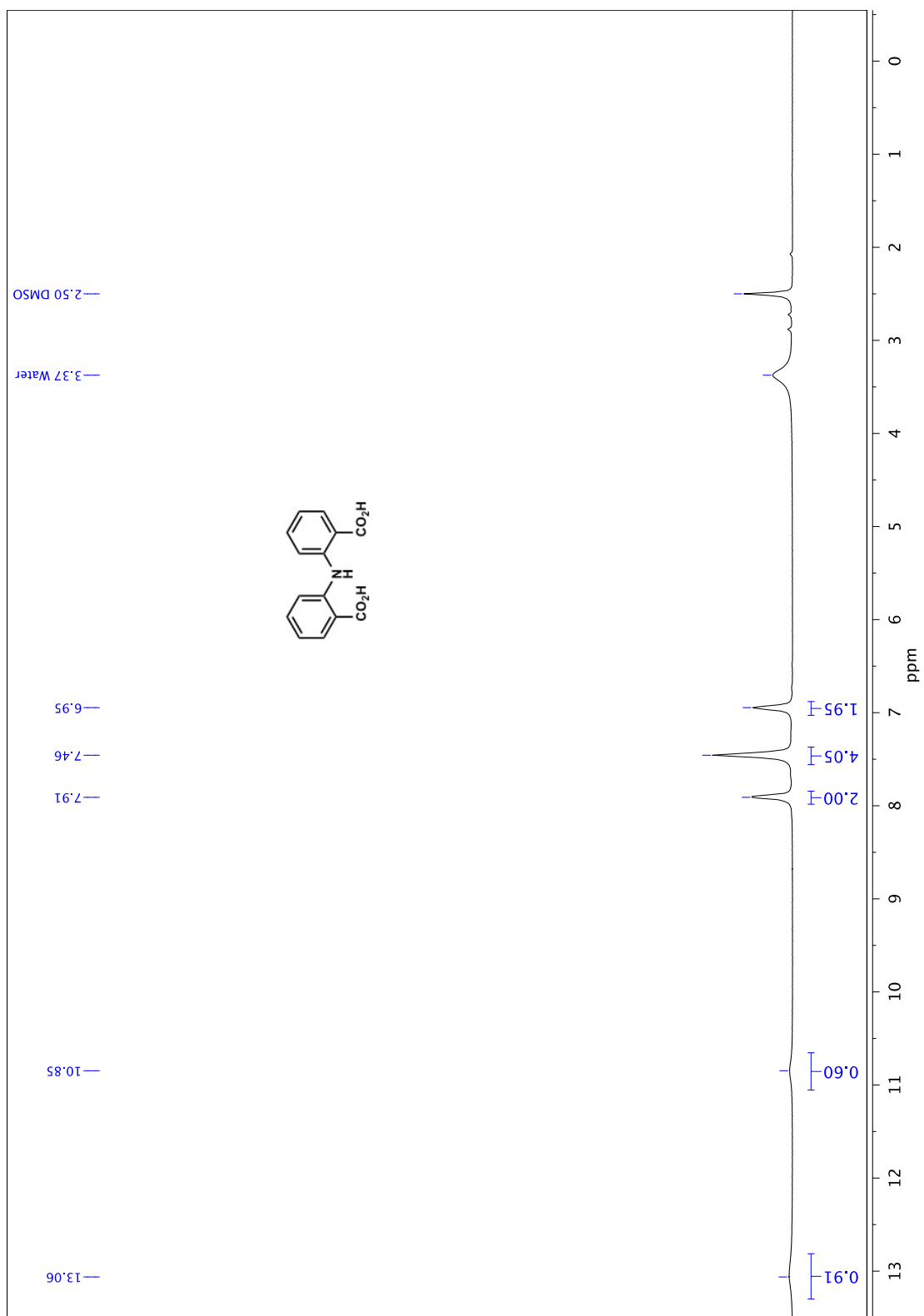


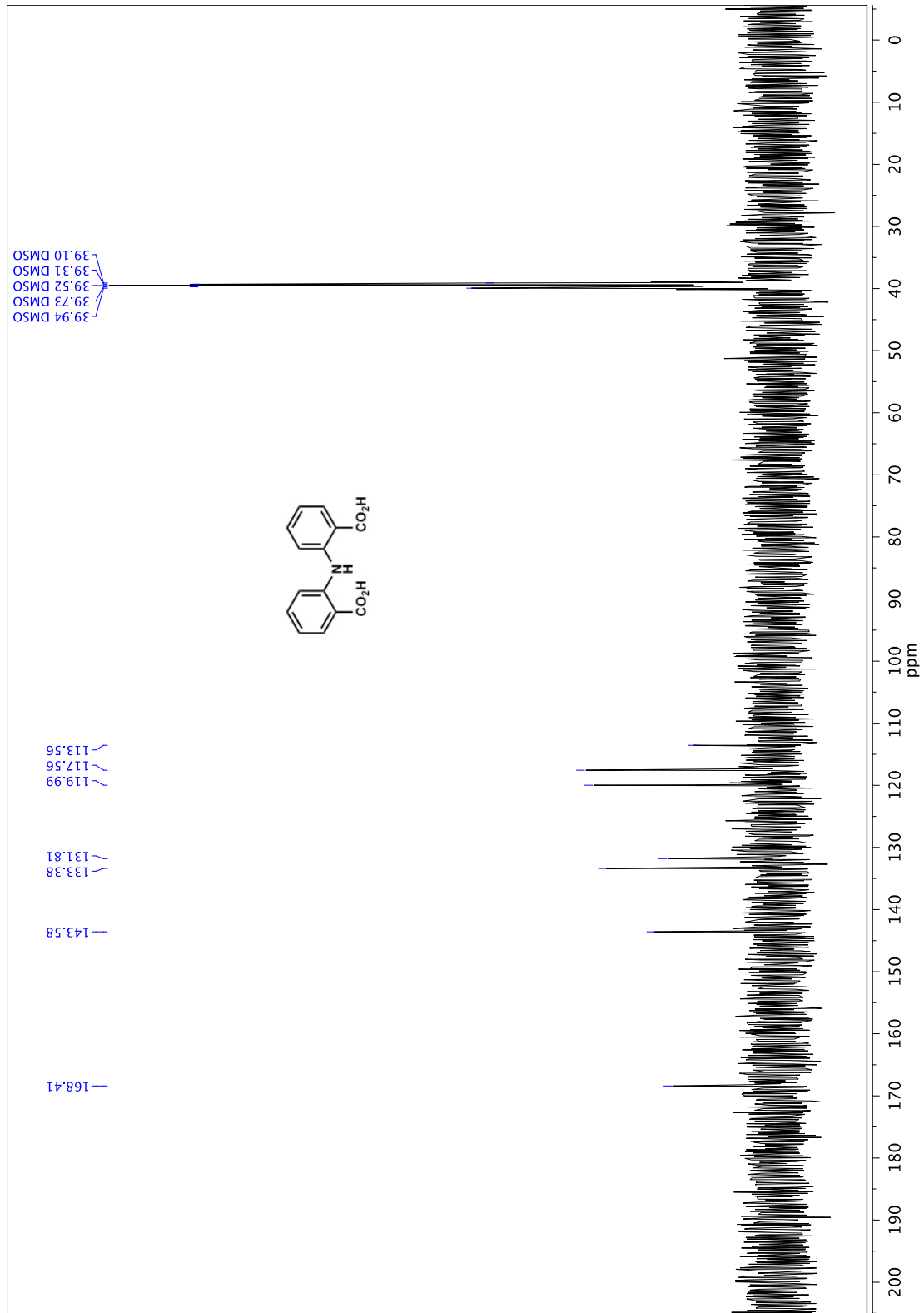


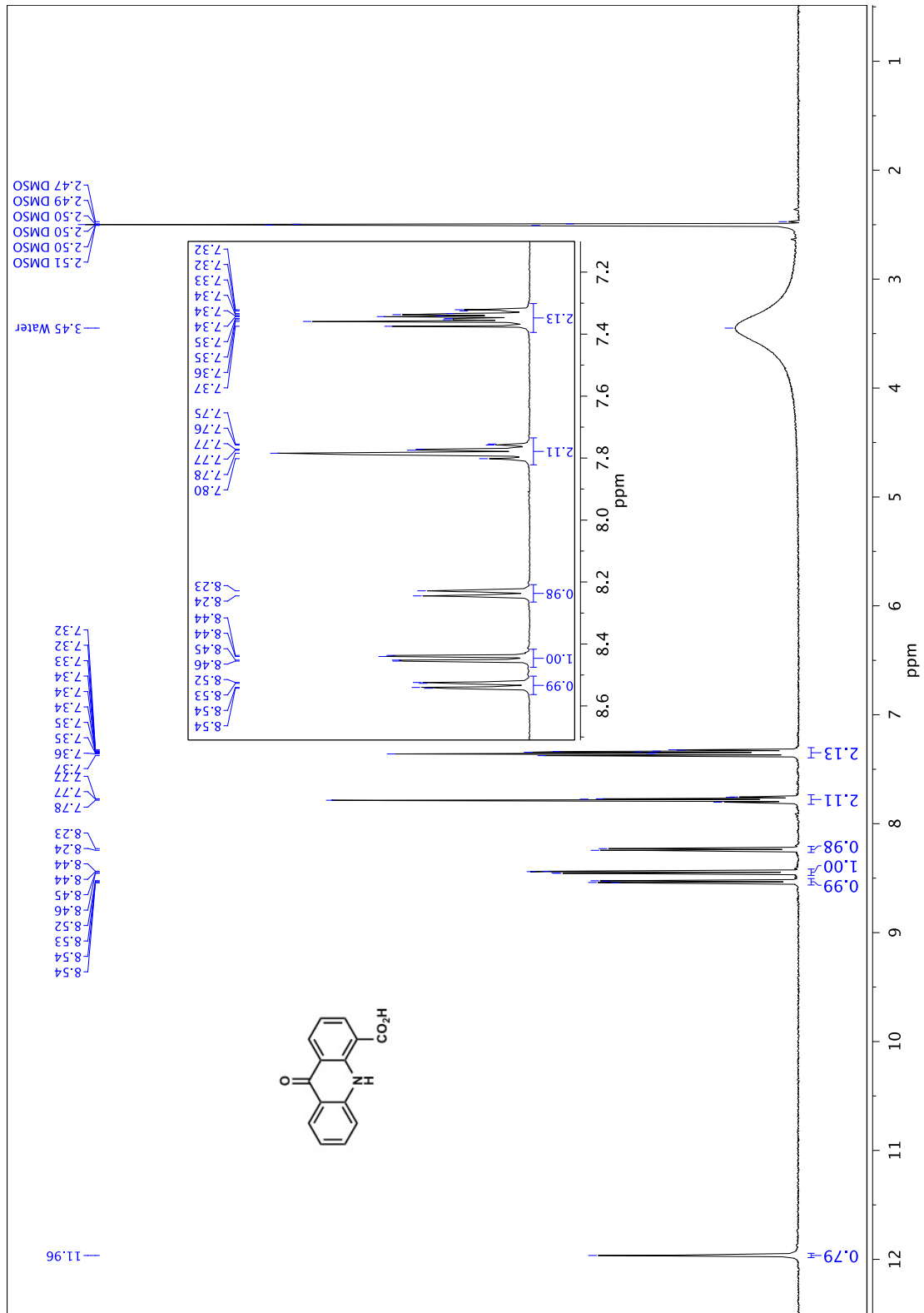


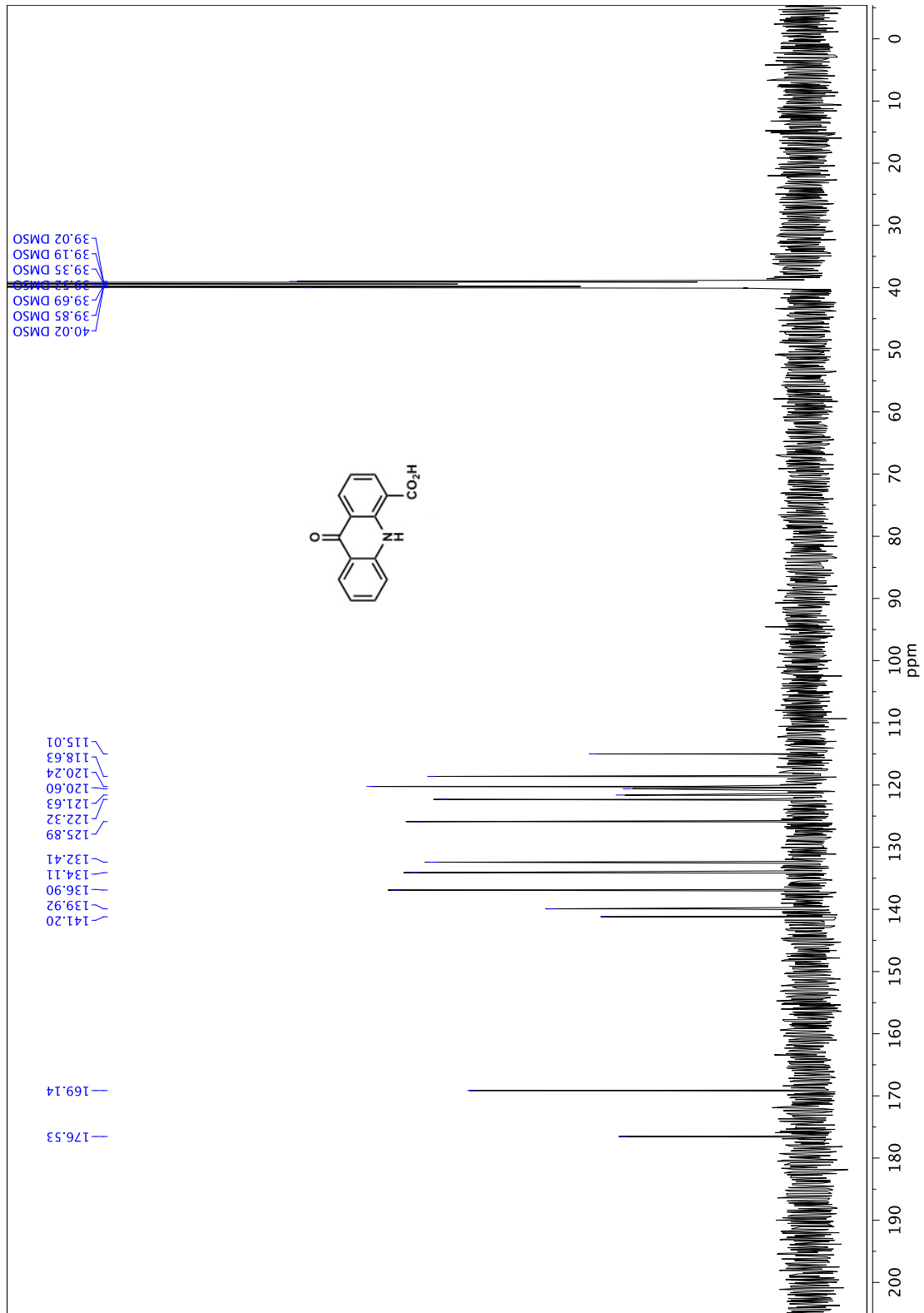




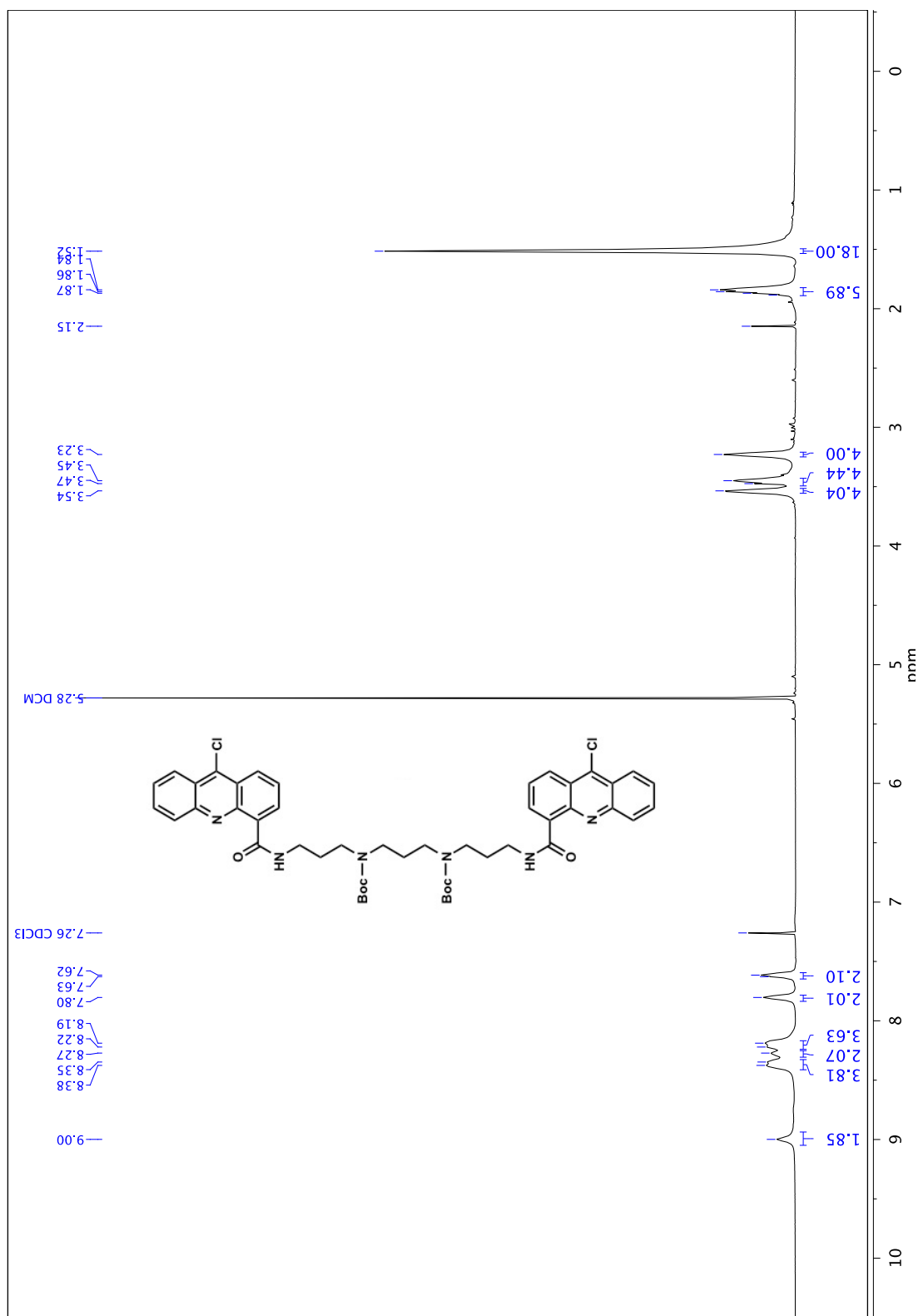


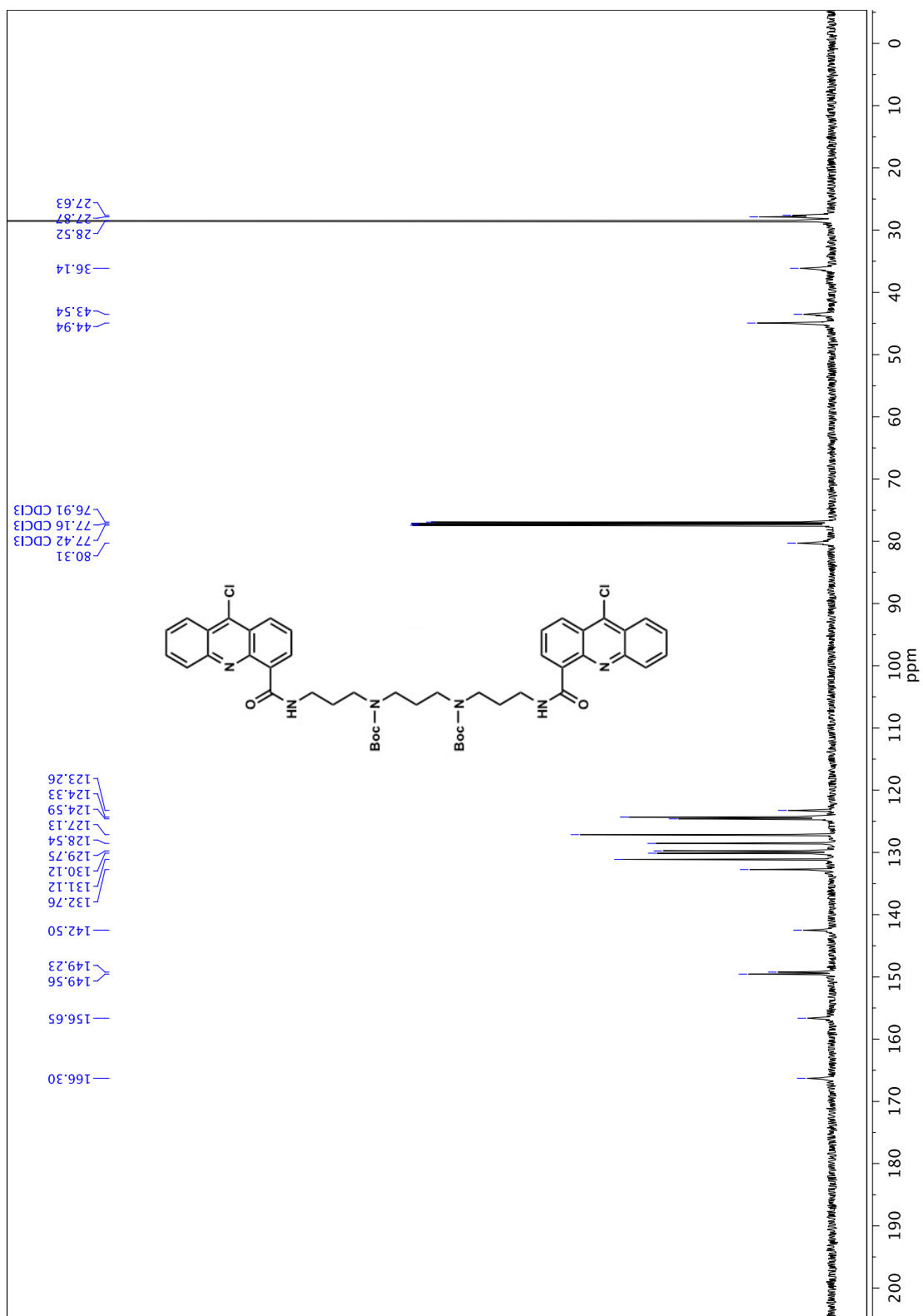




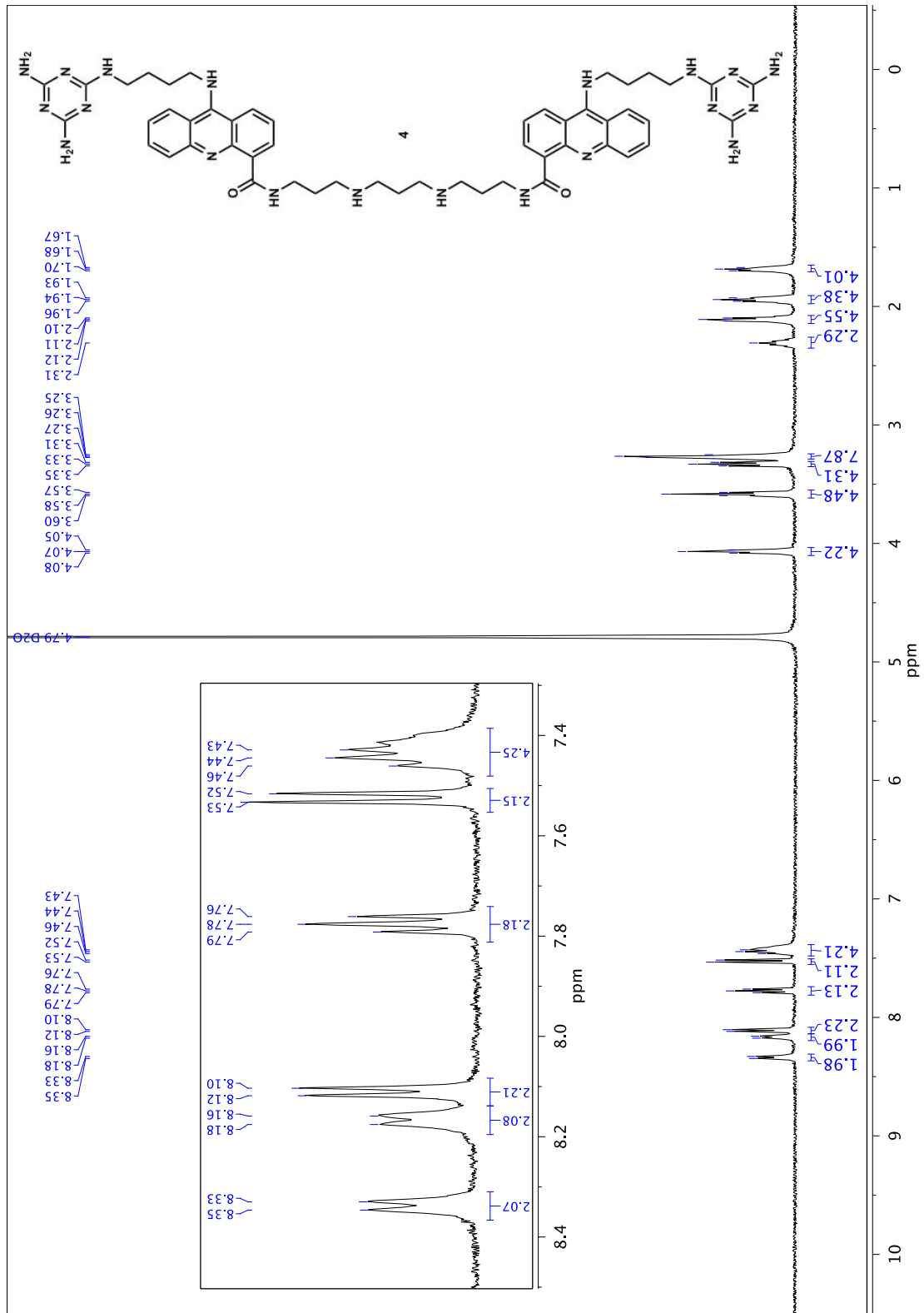












# Chapter 5.<sup>5</sup>

## CUG<sup>exp</sup> Conformation Study

### 5.1 Study of the effect of MBNL1 on CUG<sup>exp</sup> conformation

Muscleblind-like 1 protein (MBNL1) is a necessary factor for mRNA alternative splicing. Sequestration of MBNL1 to an aberrant CUG<sup>exp</sup> is the cause of a disease called myotonic dystrophy type I (DM1).<sup>1</sup> Fu *et al.* have suggested that MBNL1 changes the conformation of a mimic of one of its normal pre-mRNA targets (cTNT21) from double-stranded (ds) to single-stranded (ss) form.<sup>2</sup> Laurent *et al.* have showed that recruitment of p68 RNA helicase to CUG repeat would promote structural changes of the RNA that would facilitate and/or stabilize the binding of the splicing factor MBNL1 to the repeats.<sup>3</sup> On the contrary, Yuan *et al.* suggest that MBNL1 forms a ring-like structure which binds to the stem region of a pathologic ds form of CUG<sup>exp</sup>.<sup>4</sup> On the other hand, in a crystal structure of MBNL1 ZnF3/4 bound to r(CGUCUGU), reported by Teplova *et al.*, it is shown that ZnF3 and ZnF4 target GC steps, rather than UU mismatches, providing the possibility of MBNL1 binding to ss form target.<sup>5</sup>

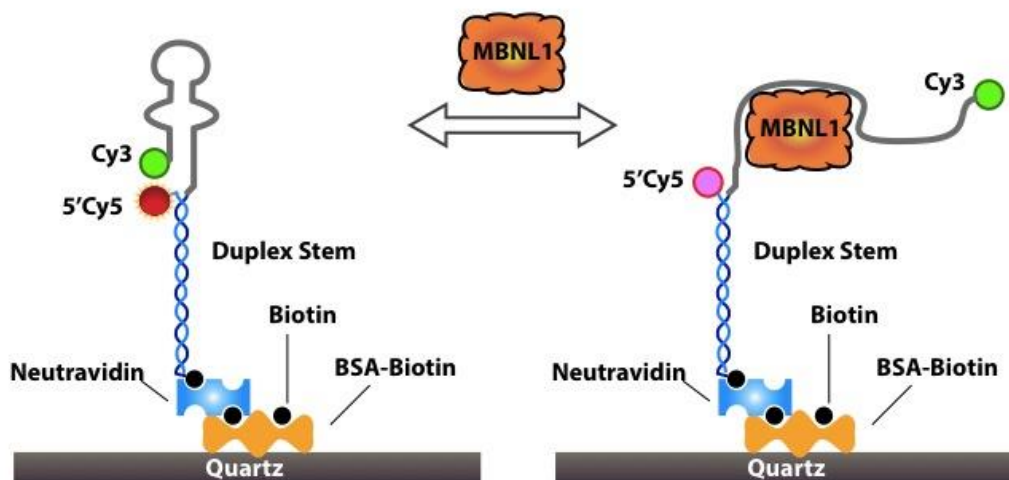
These contradictory precedents prompted us to solve the mystery of CUG<sup>exp</sup> unfolding/folding upon interaction with MBNL1. It is a crucial piece of information for drug discovery efforts against DM1. For example, if MBNL1 prefers the ss CUG<sup>exp</sup>, ligands that stabilize the ds structure of CUG<sup>exp</sup> can inhibit the CUG<sup>exp</sup>·MBNL1 interaction.

To approach this unanswered yet key structural question about how MBNL1 binds CUG<sup>exp</sup>, we did some preliminary bulk FRET (Fluorescence Resonance Energy Transfer) studies, as well as single-molecule FRET studies, in collaboration with Maria Spies lab in University of Iowa. These techniques have been previously used in similar systems.<sup>6,7</sup> Our RNA construct is composed of a (CUG)<sub>4</sub> RNA which is Cy3-labeled at one end and is connected to a Cy5-labeled biotinylated ds-DNA at the other end. This ds-DNA serves as a handle to immobilize (CUG)<sub>4</sub> on the surface and allows us to study the

---

<sup>5</sup> This chapter is a collaborative project with Maria Spies lab. Data collection and analysis was performed by Masayoshi Honda, a postdoctoral researcher in Maria Spies lab.

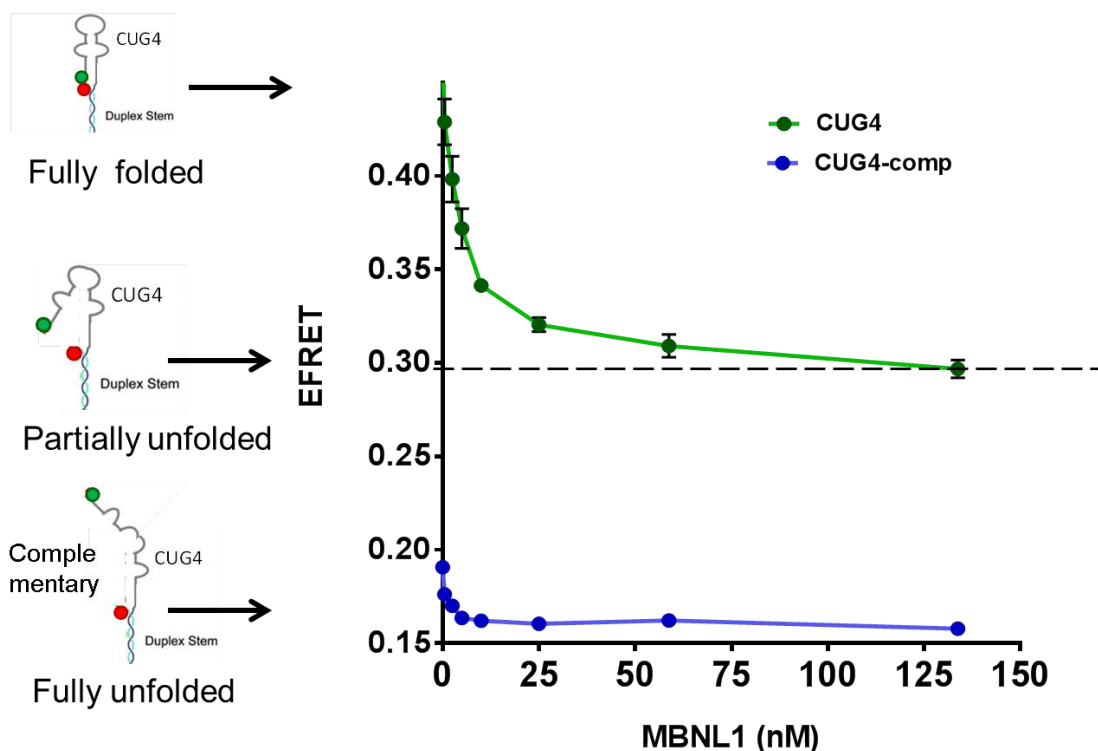
distance of Cy3 label from Cy5 by FRET experiments as shown in figure 5.1. The unfolding can be induced by the addition of the DNA oligo - d(GAC)<sub>n</sub> and this serves as a positive control on the FRET experiment.



**Figure 5.1.** Schematic representation of the FRET pair used in the TIRFM study. CUG repeat is immobilized on the surface *via* neutravidin-biotin interaction through a DNA oligonucleotide. Cy3 is located at the 3' end of CUG repeat whereas Cy5 is on the 5' end of the DNA sequence.

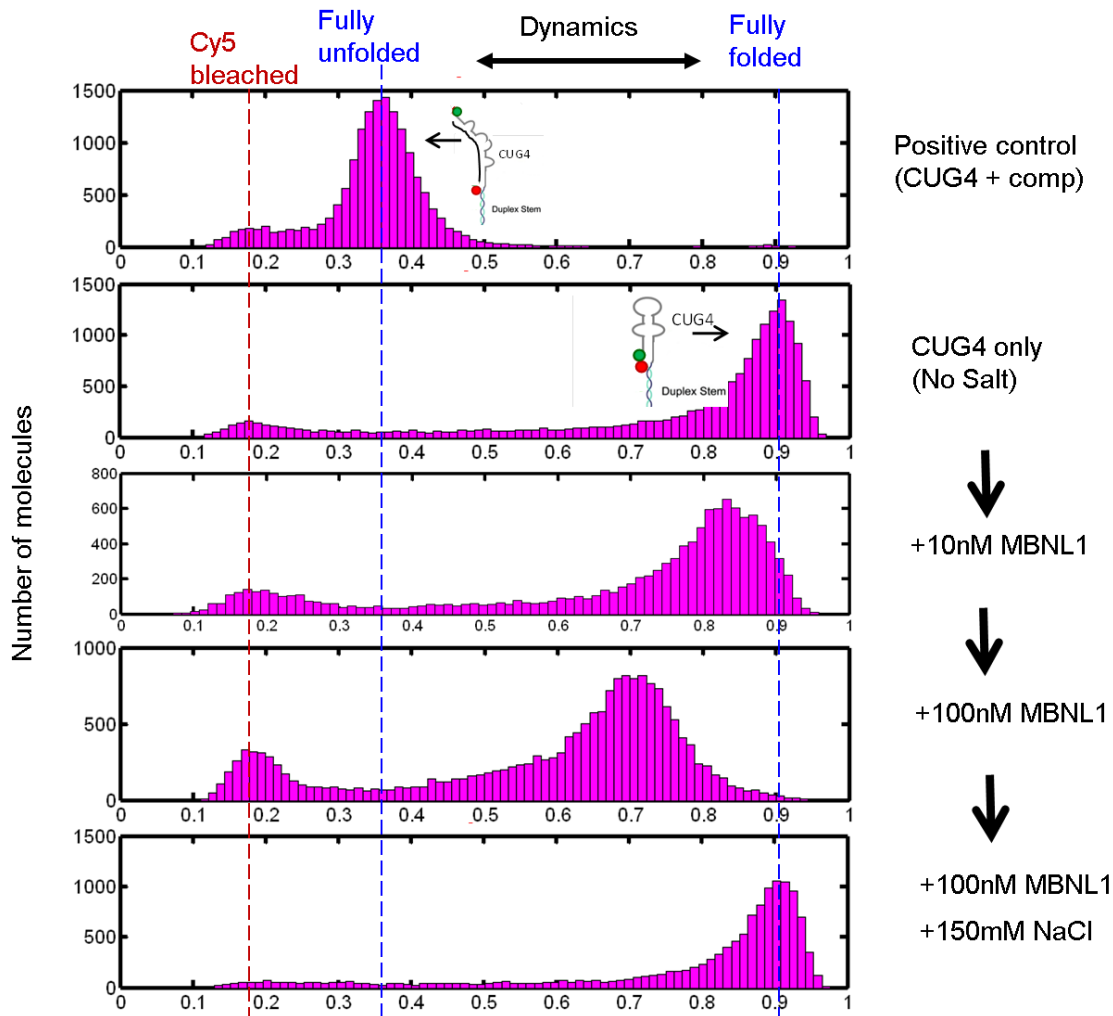
Bulk (ensemble) FRET Experiments suggested that MBNL1 induces partial unfolding of (CUG)<sub>4</sub> hairpin structure as we observed a partial decrease in FRET efficiency of (CUG)<sub>4</sub> from 0.46 to 0.30, by increasing [MBNL1] from 0 to 125 nM. Complementary DNA to (CUG)<sub>4</sub> serves as our positive control for fully unfolded state. This construct represents FRET efficiency as low as 0.19 and was not affected by MBNL1 in the concentration range from 0 to 125 nM. This preliminary result is encouraging and paved the way for studying this system at the single-molecule level. Using Total Internal Reflection Fluorescence (TIRF) microscopy we did some preliminary smFRET experiments on the (CUG)<sub>4</sub>:MBNL1 interaction at the single molecule level. We are also working on the dynamics of (CUG)<sub>4</sub> conformation and the effect of various cations such as magnesium on it.

smFRET experiments allow us to study the conformational dynamics of single (CUG)<sub>4</sub> molecules in the presence and absence of MBNL1. We obtained (CUG)<sub>4</sub> FRET efficiency histograms and studied the effect of MBNL1 on it (Figure 5.3). We observed that MBNL1 shifted the (CUG)<sub>4</sub> populations toward partial unfolded states and induced a fast transition between folded and partially unfolded state.



**Figure 5.2.** The effect of MBNL1 on the FRET efficiency of (CUG)<sub>4</sub> and (CUG)<sub>4</sub>-(GAC)<sub>4</sub>. The curve of (CUG)<sub>4</sub> is shown in green and The curve of (CUG)<sub>4</sub>-(GAC)<sub>4</sub> is shown in blue. MBNL1 lowers the FRET efficiency of (CUG)<sub>4</sub> whereas has no significant effect on the FRET efficiency of (CUG)<sub>4</sub>-(GAC)<sub>4</sub>. Proposed folding status of (CUG)<sub>4</sub> corresponding to each state is shown in left.

Cross-correlation/HaMMY analysis as well as study of (CUG)<sub>8</sub>, a longer CUG repeat, is under way to study the folding kinetics of RNA molecules. Addition of more CUG repeats to the construct not only makes the study more biologically relevant, but also removes the effect, non-native stable UUCG hairpin loop of (CUG)<sub>4</sub> might have affect on the dynamic of (CUG)<sub>4</sub>. The presence of stable UUCG hairpin loop is required for proper folding of (CUG)<sub>4</sub> but not (CUG)<sub>8</sub>. The loop does not bind to MBNL1 and may trap (CUG)<sub>4</sub> in a semi-folded state, even in the presence of MBNL1. This way, it may contribute as a hairpin swinging back and forth between a closed and semi-closed form.



**Figure 5.3.** FRET histogram revealed the distribution of  $(CUG)_4$  hairpin folding states and the effect of MBNL1 on them. X axis shows the FRET efficiency.  $(CUG)_4-(GAC)_4$  is in the unfolded state ( $E_{FRET} = 0.35$ ).  $(CUG)_4$  is in folded state ( $E_{FRET} = 0.9$ ). MBNL1 shifts this peak to semi-folded state in the absence of NaCl, however in the presence of 150 mM NaCl, it has no effect on the  $E_{FRET}$  (courtesy of Dr. Masayoshi Honda).

## 5.2 Acknowledgements

I would like to thank Dr. Masayoshi Honda and Professor Maria Spies (University of Iowa) for the single molecule collaboration. I thank Dr. Yuan Fu and Kali A. Miller for their assistance with MBNL1 expression, purification and the synthetic procedures.



### 5.3 References

- (1) Cooper, T. A. (2006) A reversal of misfortune for myotonic dystrophy?, *New Engl. J. Med.* 355, 1825-1827.
- (2) Fu, Y., Ramisetty, S. R., Hussain, N., and Baranger, A. M. (2012) MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation, *ChemBioChem* 13, 112-119.
- (3) Laurent, F. X., Sureau, A., Klein, A. F., Trouslard, F., Gasnier, E., Furling, D., and Marie, J. (2012) New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats, *Nucleic Acids Res.* 40, 3159-3171.
- (4) Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., and Swanson, M. S. (2007) Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs, *Nucleic Acids Res.* 35, 5474-5486.
- (5) Teplova, M., and Patel, D. J. (2008) Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1, *Nat. Struct. Mol. Biol.* 15, 1343-1351.
- (6) Zhao, R., and Rueda, D. (2009) RNA folding dynamics by single-molecule fluorescence resonance energy transfer, *Methods* 49, 112-117.
- (7) Grimme, J. M., and Spies, M. (2011) FRET-based assays to monitor DNA binding and annealing by Rad52 recombination mediator protein, *Methods Mol. Biol.* 745, 463-483.