

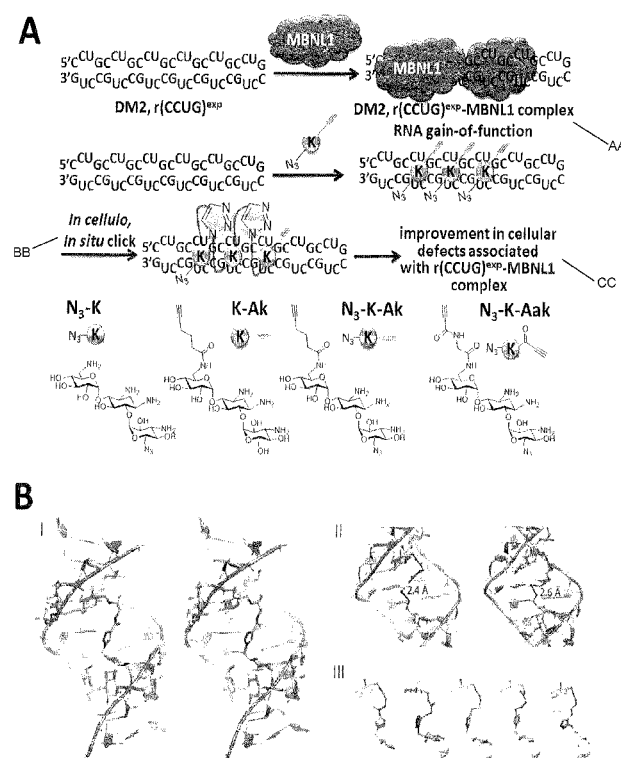


- (51) **International Patent Classification:**  
*A61P 21/00* (2006.01)
- (21) **International Application Number:**  
PCT/US2015/040902
- (22) **International Filing Date:**  
17 July 2015 (17.07.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
62/026,266 18 July 2014 (18.07.2014) US
- (71) **Applicant:** THE SCRIPPS RESEARCH INSTITUTE  
[US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).
- (72) **Inventors:** DISNEY, Matthew, D.; 493 Peacock Lane, North, Jupiter, FL 33458 (US). RZUCZEK, Suzanne, G.; 843 University Blvd., Apt. 308, Jupiter, FL 33458 (US).
- (74) **Agents:** FITTING, Thomas et al.; The Scripps Research Institute, 10550 North Torrey Pines Road, Tpc-8, La Jolla, CA 92037 (US).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

[Continued on next page]

(54) **Title:** TOXIC RNA INHIBITORS SELF-ASSEMBLED IN SITU

Figure 1



**(57) Abstract:** Potent modulators of RNA function can be assembled in cellulo by using the cell as a reaction vessel and a disease-causing RNA as a catalyst. When designing small molecule effectors of function, a balance between permeability and potency must be struck. Low molecular weight compounds are more permeable while higher molecular weight compounds are more potent. The advantages of both types of compounds could be synergized if low molecular weight molecules could be transformed into potent, multivalent ligands via a reaction catalyzed by binding to a target in cells expressing a genetic defect. We demonstrate that this approach is indeed viable in cellulo. Small molecule modules with precisely positioned alkyne and azide moieties bind adjacent internal loops in r(CCUG)exp, the causative agent of myotonic dystrophy type 2 (DM2), and are transformed into oligomeric, potent inhibitors of DM2 RNA dysfunction via a 1,3 Huisgen dipolar cycloaddition reaction, a variant of click chemistry. Additionally, we show that this approach is applicable to the r(CUG) repeating RNA that causes myotonic dystrophy type 1 (DM1). The click chemistry approach also allows for FRET sensors to be synthesized on-site by using r(CUG) repeats as a catalyst. Furthermore it is shown that small molecule binding sites in patient-derived cells can be identified by using reactive approaches termed Chem-CLIP and Chem-CLIP-Map. Lastly, it is shown that small molecules that target r(CUG) expansions can be designed to cleave this RNA by appending a small molecule with a nucleic acid cleaving module.



---

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, **Published:**  
GW, KM, ML, MR, NE, SN, TD, TG).

— *with international search report (Art. 21(3))*

**TOXIC RNA INHIBITORS SELF-ASSEMBLED IN SITU****CROSS-REFERENCE TO RELATED APPLICATION**

This application claims the priority of U.S. provisional application Serial Number 62/026,266, filed July 18, 2014, the disclosure of which is incorporated by reference herein in its entirety.

**BACKGROUND**

RNA dysfunction causes disease through various mechanisms, including microRNA silencing of pro-apoptotic proteins,<sup>[1]</sup> translation of aberrant protein,<sup>[2]</sup> and gain-of-function.<sup>[3]</sup> It has been difficult, however, to design small molecule chemical probes of RNA function or lead therapeutics. If broadly applicable methods were developed to drug non-ribosomal RNAs with small molecules, it could have important applications in chemical biology and medicinal chemistry.<sup>[4]</sup> One class of RNA-mediated diseases is caused by expanded repeating RNAs, or microsatellite disorders. There are >20 known microsatellite disorders, including myotonic dystrophy (DM) and amyotrophic lateral sclerosis (Lou Gehrig's Disease; ALS).<sup>[5]</sup> The cellular consequences of repeats are varied and can include alterations at the protein, RNA, and DNA levels. Myotonic dystrophy type 2 (DM2) is caused by a toxic gain-of-function by a r(CCUG) repeat expansion (r(CCUG)<sup>exp</sup>). Myotonic dystrophy type 1 (DM1) is caused by a toxic gain-of-function by a r(CUG) repeat expansion (r(CUG)<sup>exp</sup>).

**SUMMARY**

We have designed inhibitors or modulators of the RNA that causes myotonic dystrophy type 2 (DM2) based on RNA motif-small molecule interactions.<sup>[6]</sup> DM2 is caused by a toxic gain-of-function by a r(CCUG) repeat expansion (r(CCUG)<sup>exp</sup>) located in intron 1 of the zinc finger protein 9 (*ZNF9*) pre-mRNA.<sup>[7]</sup> The RNA folds into a hairpin structure that contains repeating units of 5'CCUG/3'GUCC (2x2 nucleotide internal loops) in the stem. The loops form high affinity binding sites for muscleblind-like 1 protein (MBNL1), a regulator of alternative pre-mRNA splicing, which is inactivated upon binding (**Fig. 1A**). Our designed inhibitors of r(CCUG)<sup>exp</sup> are based on a kanamycin A derivative, which is acylated at the 6' position, that binds 5'CCUG/3'GUCC with high affinity. Notably, acylation of the 6' amine ablates binding to rRNA.<sup>[8]</sup> Indeed, the kanamycin derivative and modularly assembled (or multivalent) compounds thereof improve DM2-associated defects in a cellular model.<sup>[6a]</sup> As observed for other repeat

expansions, modularly assembled compounds are more potent inhibitors of cellular dysfunction, presumably due to their high affinities, selectivities, and the larger surface areas they sequester on the target.<sup>[6a]</sup> Even though multivalent compounds are more potent and selective *in vitro*, in some cases *in cellulo* potency decreases as a function of valency because they are less cell permeable.<sup>[9]</sup> Thus, a careful balance must be struck between ligand size, potency, and permeability.

We have also used a similar approach to design inhibitors or modulators of the RNA that causes myotonic dystrophy type 1 (DM1). In this case, DM1 is caused by a toxic gain-of-function by a r(CUG) repeat expansion (r(CUG)<sup>exp</sup>) that is located in the 3' untranslated region (UTR) of the dystopia myotonica protein kinase (*DMPK*) mRNA.<sup>[7]</sup> Other aspects of this invention for DM1 include: (i) design of optimized dimeric compounds that target r(CUG)<sup>exp</sup>; (ii) covalent small molecules that target r(CUG)<sup>exp</sup> and allow for target validation (RNAs bound) and the sites in the RNAs that bind the small molecules; (iii) use of *in cellulo* click chemistry to synthesize inhibitors on-site that are highly potent and selective; (iv) fluorescence resonance energy transfer (FRET) approaches to use r(CUG)<sup>exp</sup> as a catalyst to synthesize FRET sensors; and (v) design of small molecules targeting r(CUG)<sup>exp</sup> that cleave r(CUG)<sup>exp</sup> in patient-derived cells.

The invention provides, in various embodiments, a method of forming, within a living cell, a modulator of RNA function, comprising exposing the cell to one or two small molecule modules to which the cell is permeable, a single module bearing both alkyne and azide groups, or the two modules comprising respectively alkyne and azide moieties that bind adjacent internal loops in r(CCUG)<sup>exp</sup> or r(CUG)<sup>exp</sup>, the causative agent of myotonic dystrophy type 2 (DM2) and type 1 (DM1) (Day, J. W., and Ranum, L. P. (2005) RNA pathogenesis of the myotonic dystrophies, *Neuromuscul. Disord.* 15, 5-16.; Miller, J. W., Urbinati, C. R., Teng-Umnay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., and Swanson, M. S. (2000) Recruitment of human muscleblind proteins to (CUG)<sub>(n)</sub> expansions associated with myotonic dystrophy, *EMBO J.* 19, 4439-4448; 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 *Clcn-1* chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy, *Mol. Cell* 10, 35-44), respectively, such that the one module, or the two modules are transformed by condensation of the alkyne and azide group in a 1,3 Huisgen dipolar

cycloaddition reaction into oligomeric, potent inhibitors of DM1 or DM2 RNA dysfunction via a 1,3 Huisgen dipolar cycloaddition reaction. Thus, the monomeric precursors of the oligomeric modulator can either be single compounds that display azide and alkyne units on the same molecule can bind to adjacent sites and undergo a 1,3 Huisgen dipolar cycloaddition reaction, or a pair of molecular structures, one of which bears an alkyne group and the other of which bears an azide group.

The invention also provides, in various embodiments, a method of forming, within a living cell, a modulator of RNA function, comprising exposing the cell to small molecule modules to which the cell is permeable in which the module contains both alkyne and azide moieties. The invention further provides, in various embodiments, modulators of RNA function, formed by exposing a cell containing RNA expanded repeat sequences  $r(\text{CCUG})^{\text{exp}}$  or  $r(\text{CUG})^{\text{exp}}$ , the causative agent of myotonic dystrophy type 2 (DM2) and type 1 (DM1) respectively, to one or more RNA extended repeat binding modules to which the cell is permeable, the two modules comprising respectively alkyne and azide moieties that can condense via a Huisgen 1,3-dipolar cycloaddition reaction to form an oligomeric modulator that interferes with the function of the toxic RNAs. For the  $r(\text{CCUG})^{\text{exp}}$  modulator for treatment of DM2, the cell-permeable modules can comprise kanamycin analogs. For the  $r(\text{CUG})^{\text{exp}}$  modulator for treatment of DM1, the cell-permeable modules can comprise bis-benzimidazoles or other heteroaryl compounds.

### BRIEF DESCRIPTION OF THE FIGURES

**Figure 1.** *In cellulo, in situ click* chemistry to synthesize potent inhibitors of the RNA that causes DM2. A, DM2 is caused by a  $r(\text{CCUG})$  repeat that binds and sequesters muscleblind-like 1 protein (MBNL1). Small molecules that contain azide and alkyne functional groups (**N<sub>3</sub>-K**, **K-Ak**, **N<sub>3</sub>-K-Ak**, and **N<sub>3</sub>-K-Aak**) bind adjacent sites in  $r(\text{CCUG})^{\text{exp}}$  and undergo a Huisgen dipolar cycloaddition reaction. B, Molecular dynamics (MD) simulation models of clickable modules binding to a mimic of  $r(\text{CCUG})^{\text{exp}}$ . BI, Conformational searching reveals close proximity between azide and alkyne groups presented by K modules bound to adjacent sites. BII, A low energy state in MD simulation of 1,4-triazole adduct from **N<sub>3</sub>-K** and **K-Aak** is shown in stereoview. Hydrogen bonds between the RNA and K are shown in dashed lines. BIII, Low energy snapshot of MD simulations showing other linker models.

**Figure 2.** Identifying the extent of *in cellulo* click reactions and the targets of clickable small molecules. A, Schematic of ChemReactBIP, an approach to identify the cellular targets of small molecules. Studies were enabled by using a biotinylated monomer with a single N<sub>3</sub> group to participate in the click reaction, or **N<sub>3</sub>-K-Biotin**, which allows isolation of the clicked oligomer and bound RNA targets by passing cell lysates over streptavidin resin. B, Results of ChemReactBIP mass spectral analysis of pulled down, templated small molecule products that are only synthesized in cells expressing r(CCUG)<sub>300</sub>. C, Results of ChemReactBIP in which the RNA targets of the clicked products were isolated and analyzed via qRT-PCR. “\*” indicates  $p < 0.05$ ; “\*\*” indicates  $p < 0.01$ ; “\*\*\*” indicates  $p < 0.001$  as determined by a two-tailed Student t-test ( $n = 3$ ).

**Figure 3.** Results of cellular studies evaluating the effect of *in cellulo, in situ* click on DM2-associated defects. A, Evaluation of click compounds for inhibiting formation of nuclear foci of r(CCUG)<sub>300</sub> upon treatment with various compounds. B, Schematic of the alternative splicing of bridging integrator 1 (*BINI*) pre-mRNA exon 11 in healthy cells and cells that express r(CCUG)<sub>300</sub>. C, Rescue of the *BINI* splicing defect by templated click reactions. Top, Representative gel image of *BINI* splicing products in the presence and absence of r(CCUG)<sub>300</sub>. Bottom, quantification of *BINI* splicing patterns in treated and untreated cells. The concentration of each compound is indicated above the corresponding bar in  $\mu\text{M}$ . For **N<sub>3</sub>-K + K-Ak**, 5  $\mu\text{M}$  indicates the concentration of each compound. “\*” indicates  $p < 0.05$ ; “\*\*” indicates  $p < 0.01$ ; “\*\*\*” indicates  $p < 0.001$  as determined by a two-tailed Student t-test ( $n \geq 3$ ).

**Figure 4.** Secondary structures of RNAs used for *in vitro* click reactions with **N<sub>3</sub>-K** and **K-Ak**.

**Figure 5.** Results of LC-MS analysis of *in vitro* click reactions between **N<sub>3</sub>-K** and **K-Ak** in the presence and absence of r(CCUG)<sub>12</sub>. These data show that a K dimer is formed in the presence of r(CCUG)<sub>12</sub> but not in its absence.

**Figure 6.** Results of LC-MS analysis of *in vitro* click reactions between **N<sub>3</sub>-K** and **K-Ak** in the presence and absence of r(CCUG)<sub>12</sub>, r(AUUCU)<sub>12</sub>, r(CAG)<sub>12</sub>, r(CGG)<sub>12</sub>, r(CUG)<sub>12</sub>, an RNA hairpin with a fully paired stem, and brewer's yeast tRNA. These data show that the formation of a K dimer is selectively catalyzed by r(CCUG)<sub>12</sub> ( $n=3$ ).

**Figure 7.** Results of LC-MS analysis of *in cellulo* click reactions between **N<sub>3</sub>-K** and **K-Ak** in the presence and absence of r(CCUG)<sub>300</sub>. These data show that a K dimer is formed in the

presence of  $r(\text{CCUG})_{300}$  but not in its absence. “\*” indicates  $p < 0.05$  as determined by a two-tailed Student t-test ( $n=3$ ).

**Figure 8.** Results of LC-MS analysis of *in cellulo* click reactions between **N<sub>3</sub>-K** and **N<sub>3</sub>-K-Ak** in the presence and absence of  $r(\text{CCUG})_{300}$ . These data show that a dimer and a trimer are formed in the presence of  $r(\text{CCUG})_{300}$  but not in its absence. **N<sub>3</sub>-K** was used to limit the extent of polymerization to facilitate LC-MS analysis (see “*Identification of in cellulo clicked products by mass spectrometry*”). “\*” indicates  $p < 0.05$ ; “\*\*\*” indicates  $p < 0.01$  as determined by a two-tailed Student t-test ( $n=3$ ).

**Figure 9.** Results of LC-MS analysis of *in cellulo* click reactions between **N<sub>3</sub>-K** and **N<sub>3</sub>-K-Aak** in the presence and absence of  $r(\text{CCUG})_{300}$ . These data show that a dimer and a trimer are formed in the presence of  $r(\text{CCUG})_{300}$  but not in its absence. **N<sub>3</sub>-K** was used to limit the extent of polymerization to facilitate LC-MS analysis (see “*Identification of in cellulo clicked products by mass spectrometry*”). “\*” indicates  $p < 0.05$ ; “\*\*\*” indicates  $p < 0.01$  as determined by a two-tailed Student t-test ( $n=3$ ).

**Figure 10.** Results of LC-MS analysis of pulled-down, biotin-tagged compounds from *in cellulo* click reactions between **N<sub>3</sub>-K-Biotin** and **K-Ak** in the presence and absence of  $r(\text{CCUG})_{300}$ . These data show that a biotinylated K dimer is formed in the presence of  $r(\text{CCUG})_{300}$  but not in its absence. “\*\*\*” indicates  $p < 0.01$  as determined by a two-tailed Student t-test ( $n=3$ ).

**Figure 11.** Results of LC-MS analysis of pulled-down, biotin-tagged compounds from *in cellulo* click reactions between **N<sub>3</sub>-K-Biotin** and **N<sub>3</sub>-K-Ak** in the presence and absence of  $r(\text{CCUG})_{300}$ . These data show that a dimer and a trimer are formed in the presence of  $r(\text{CCUG})_{300}$  but not in its absence. **N<sub>3</sub>-K-Biotin** was used to limit the extent of polymer formation and facilitate LC-MS analysis (see “*Identification of in cellulo clicked products by mass spectrometry*”). “\*” indicates  $p < 0.05$  as determined by a two-tailed Student t-test ( $n=3$ ).

**Figure 12.** Results of LC-MS analysis of pulled-down, biotin-tagged compounds from *in cellulo* click reactions between **N<sub>3</sub>-K-Biotin** and **N<sub>3</sub>-K-Aak** in the presence and absence of  $r(\text{CCUG})_{300}$ . These data show that a dimer and a trimer are formed in the presence of  $r(\text{CCUG})_{300}$  but not in its absence. **N<sub>3</sub>-K-Biotin** was used to limit the extent of polymer formation and facilitate LC-MS analysis (see “*Identification of in cellulo clicked products by*

mass spectrometry"). "\*" indicates  $p < 0.05$ ; "\*\*\*" indicates  $p < 0.01$  as determined by a two-tailed Student t-test ( $n=3$ ).

**Figure 13. Top,** Representative gel images demonstrating the effect of K derivatives on *BIN1* alternative splicing patterns. A 1:1 mixture of **N<sub>3</sub>-K** and **K-Ak** improves *BIN1* patterns to a similar extent as pre-synthesized dimers. **Bottom,** Quantification of *BIN1* alternative splicing patterns in treated and untreated cells. The activities of monomers **N<sub>3</sub>-K** and **K-Ak** were compared to pre-synthesized dimers as well as an equimolar mixture of each monomer. The pre-synthesized dimers improved splicing to the greatest extent while a mixture of azide and alkyne monomers improved splicing to a greater extent than each monomer alone. "\*" indicates  $p < 0.05$ ; "\*\*\*" indicates  $p < 0.001$  as determined by a two-tailed Student t-test ( $n \geq 3$ ).

**Figure 14. Top,** Representative gel images of *BIN1* splicing patterns in cells treated with **N<sub>3</sub>-K-Ak**. Improvement in splicing patterns was observed when cells were treated with 10 and 1  $\mu$ M **N<sub>3</sub>-K-Ak**. **Middle,** Representative gel images of *BIN1* splicing patterns in cells treated with **N<sub>3</sub>-K-Aak**. Improvement in splicing patterns was observed when cells were treated with 10  $\mu$ M, 1  $\mu$ M, and 100 nM **N<sub>3</sub>-K-Ak**. **Bottom,** Quantification of *BIN1* splicing patterns in untreated cells and cells treated with oligomerizable K compounds. **N<sub>3</sub>-K-Aak** was the most potent compound evaluated. It significantly improves *BIN1* pre-mRNA splicing defects at nanomolar concentrations and is 1000 times more potent than K monomers. "\*" indicates  $p < 0.05$ ; "\*\*\*" indicates  $p < 0.01$ ; "\*\*\*\*" indicates  $p < 0.001$  as determined by a two-tailed Student t-test ( $n \geq 3$ ).

**Figure 15. Top,** Representative gel images showing that K derivatives do not affect *BIN1* splicing patterns in cells that do not express r(CCUG)<sub>300</sub>. **Bottom,** Quantification of *BIN1* splicing patterns in cells that do not express r(CCUG)<sub>300</sub>. None of the compounds evaluated had a statistically significant effect on *BIN1* pre-mRNA splicing patterns in the absence of r(CCUG)<sub>300</sub> as determined by a two-tailed Student t-test ( $n \geq 3$ ). The concentration of compound tested is provided in parentheses (mM).

**Figure 16** shows the scheme for synthesis of the **K 1,4 Dimer**.

**Figure 17** shows the scheme for synthesis of the **K 1,5 Dimer**.

**Figure 18** shows the scheme for synthesis of **N<sub>3</sub>-K-Ak**.

**Figure 19** shows the scheme for synthesis of **N<sub>3</sub>-K-Aak**.

**Figure 20** shows the scheme for synthesis of **N<sub>3</sub>-K-Biotin**.



**Figure 21.** The toxic RNA-protein complex causative of DM1 and designer small molecules that are used to ameliorate and study disease-associated cellular dysfunction and *in cellulo* target selectivity.

**Figure 22.** Lead optimization of compounds for metabolic stability. Compound **2H-K4NMe** is unstable in microsomes (bottom right plot) and instability was traced to the imino proton, which was removed by *N*-methylation to provide **2H-K4NMeS**, which is stable in microsomes.

**Figure 23.** Designer small molecule **2H-K4NMeS** improves DM1-associated pre-mRNA splicing defects in patient-derived cell lines (left) and in a DM1 mouse model (right). Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 289, 1769–1773.

**Figure 24.** The designer small molecule **2H-K4NMeS** is appended with a reactive module (CA) and a purification tag (biotin) to produce **2H-K4NMeS-CA-Biotin**. This compound forms a cross-link (covalent bond) with bound RNAs in cells and allows them to be purified and quantified to validate the targets of the small molecules by using an approach termed Chemical Cross-Linking and Isolation by Pull Down (Chem-CLIP). A, The structure of the probe compound. B, *In vitro* data showing that the probe selectively reacts with r(CUG)<sup>exp</sup>. C, **2H-K4NMeS-CA-Biotin** improves DM1-associated pre-mRNA splicing defects in cells. D, The non-covalent compound **2H-K4NMeS** can compete **2H-K4NMeS-CA-Biotin** from reaction with the r(CUG)<sup>exp</sup> target as determined by Competitive Chem-CLIP (C-Chem-CLIP). E, Pull-down of the RNA targets by **2H-K4NMeS-CA-Biotin** shows the compound binds to disease-causing expanded transcripts and not other RNAs with shorter (non-pathogenic length) r(CUG) repeats.

**Figure 25.** Chemical Cross-Linking and Isolation by Pull Down and Ligand Binding-Site Mapping (Chem-CLIP-Map). The *DMPK* mRNA that reacted with **2H-K4NMeS-CA-Biotin** was site specifically digested with RNase H by using oligonucleotides complementary to different regions in the mRNA. After digestion, regions of the RNA that reacted with **2H-K4NMeS-CA-Biotin** were captured on a streptavidin resin and quantified by using qRT-PCR. Data shows that **2H-K4NMeS-CA-Biotin** binds to r(CUG)<sup>exp</sup> in *DMPK* mRNA.

**Figure 26.** The structures of the compounds that were tested for validating the on-site drug synthesis *in vitro*.

**Figure 27.** Data for the *in vitro* oligomeric drug synthesis by using nucleic acids as catalysts. A, Results of probing linker length to determine the optimal distance for reaction; B Evaluation of *in vitro* drug synthesis between **N<sub>3</sub>-2H-K4NMeS** and **2H-K4NMeS-Aminohexanoate Aak** in the presence of various RNAs.

**Figure 28.** The designer compound **N<sub>3</sub>-2H-K4NMeS-Aak** that can form oligomeric species that has a azide and alkyne site on a single r(CUG)<sup>exp</sup> binding ligand.

**Figure 29.** *In cellulo* click chemistry by using **N<sub>3</sub>-2H-K4NMeS-Aak** (Fig. 28) and scheme of Chem-React-BIP to pull down and quantify targets.

**Figure 30.** Ability of *in cellulo* click compounds to improve A, Pre-mRNA splicing, B, Nuclear foci, and C, A translational defect that is associated with DM1.

**Figure 31.** A **2H-K4NMeS** derivative that is conjugated to bleomycin to allow for the targeted cleavage of r(CUG)<sup>exp</sup> in DM1 patient-derived cells. A, structure of the compound; B, qRT-PCR data to show that the target is cleaved; C, competitive cleaving data in which **2H-K4NMeS-Bleomycin A5** and **2H-K4NMeS** are co-added to cells showing that **2H-K4NMeS** can inhibit targeted r(CUG)<sup>exp</sup> cleavage. D, Improvement of downstream DM1-associated pre-mRNA splicing defects.

**Figure 32.** Scheme of the fluorescence resonance energy transfer (FRET) experiment to synthesize a FRET sensor by using r(CUG)<sup>exp</sup> as a catalyst. A, A schematic of the approach. B, Compounds used in these experiments. C, Representative data.

**Figure 33.** Synthetic scheme for **2H-K4NMeS**.

**Figure 34.** Synthetic scheme for **2H-K4NMeS-CA-Biotin**.

**Figure 35.** Synthetic scheme for **2NAc-K4NMeS-CA-Biotin**.

**Figure 36.** Synthetic scheme for **2H-K4NMeS-Bleomycin A5**.

**Figure 37.** Synthetic scheme for **N<sub>3</sub>-2H-K4NMeS**.

**Figure 38.** Synthetic scheme for **N<sub>3</sub>-2H-K4NMeS-Biotin**.

**Figure 39.** Synthetic scheme for **2H-K4NMeS Activated Alkynes**.

**Figure 40.** Synthetic scheme for **N<sub>3</sub>-2H-K4NMeS-Aak**.

**Figure 41.** Synthetic scheme for **Biotin N<sub>3</sub>-2H-K4NMeS-Aak**.

**Figure 42.** Synthetic scheme for **N<sub>3</sub>-2H-K4NMeS-TAMRA**.

**Figure 43.** Synthetic scheme for **FAM-2H-K4NMeS-Aak**.

### DETAILED DESCRIPTION

Given the above considerations, we sought to develop a strategy to target RNA repeat expansions that could exploit the cellular permeability of the small molecule “modules” and the potency and selectivity of higher molecular weight, multivalent compounds. *In situ* click chemistry, via a 1,3 Huisgen dipolar cycloaddition reaction (HDCR), could provide such a strategy. That is, two modules would bind adjacent sites in the target, bringing otherwise unreactive groups into close proximity to form a covalent bond; in particular, azide and alkyne moieties react to form a stable triazole.<sup>[10]</sup> Indeed, this approach has been used *in vitro* to target acetylcholine esterase and the DNA minor groove.<sup>[10-11]</sup> Translating such an approach to cellular systems could be highly impactful, enabling the development of highly selective chemical biology probes yet this has not been previously demonstrated. Expanded repeating RNAs are perhaps ideal targets for this approach because they are modular like the compounds they will template (Figure 1A).

The development of an *in cellulo*, *in situ* click approach for r(CCUG)<sup>exp</sup> was enabled by using a model of the binding of dimeric 6'-N-acylated kanamycin A compound (**2K-4**) to r(CCUG) repeats.<sup>[6a]</sup> Analysis of this model showed that an azido group at 6'' position (**N<sub>3</sub>-K**) and an alkyne group at 6' position (**K-Ak**) could be within close enough proximity to react upon binding to adjacent 2x2 nucleotide internal loops in r(CCUG)<sup>exp</sup> (Figure 1B). Thus, when **N<sub>3</sub>-K** and **K-Ak** are mixed in equal amounts, a dimer could be formed; likewise, a derivative that displays both a 6'' azide and a 6' alkyne (**N<sub>3</sub>-K-Ak**; Figure 1B) could form an oligomer. We also synthesized a compound with an activated, electron deficient alkyne, **N<sub>3</sub>-K-Aak** (Figure 1A), which was employed by the Dervan group to assemble polyamides using DNA as a template.<sup>[11b]</sup> All compounds were synthesized by using variations of known routes (see Examples section).

We first tested the ability of r(CCUG)<sub>12</sub> (Figure 4) to template assembly of K oligomers *in vitro*. After incubation, reaction products were analyzed by mass spectrometry. Indeed, higher valency compounds were formed in the presence of r(CCUG)<sub>12</sub> but not in its absence (Figure 5). To determine if this templated reaction was specific to r(CCUG)<sub>12</sub>, we studied the abilities of other RNAs to catalyze oligomer formation including r(CUG)<sub>12</sub>, r(AUUCU)<sub>12</sub>, r(CGG)<sub>12</sub>, r(CAG)<sub>12</sub>, an RNA hairpin with a fully paired stem, and tRNA (Figure 4).

Importantly, no significant reaction products were observed for all RNAs, except for r(CUG)<sub>12</sub> in which a small amount of dimer was observed (<10%; Figure 6). It is not surprising that r(CUG)<sub>12</sub> templated a small amount of dimer as it binds K, albeit with a much lower affinity than r(CCUG)<sub>12</sub>.<sup>[13]</sup> Further, the optimal distance that separates K RNA-binding modules in a multivalent compound is much shorter for r(CUG)<sub>12</sub> than r(CCUG)<sub>12</sub>.<sup>[13]</sup> Taken together, these studies illustrate that the templated reaction is selective for r(CCUG)<sub>12</sub> and is controlled by the positioning of functional groups, the RNA binding module, and the RNA target (Figure 6). These studies provided impetus for testing this approach in cellular systems and to study the biological impact of the templated compounds on DM2-associated defects, which include alternative pre-mRNA splicing defects and formation of nuclear foci.<sup>[3a]</sup>

To confirm that multivalent compounds are indeed templated *in cellulo*, a cellular model system in which r(CCUG)<sub>300</sub> is expressed was employed.<sup>[14]</sup> Cells were co-treated with **N<sub>3</sub>-K** and **N<sub>3</sub>-K-Ak** or **N<sub>3</sub>-K-Aak**. **N<sub>3</sub>-K** was used to poison the reaction in order to limit the molecular weight of the products, allowing for mass spectral analysis. After treatment, reaction products were partially purified from cell lysates by precipitating cellular material and proteins with organic solvent. Mass spectral analysis of purified fractions showed that oligomerization occurred in cells expressing r(CCUG)<sub>300</sub> as both dimeric and trimeric reaction products are observed when cells were treated with **N<sub>3</sub>-K-Ak** or **N<sub>3</sub>-K-Aak** (Figures 7-9). Importantly, oligomerization is not observed in cells that do not express the RNA (Figures 7-9). Thus, templated synthesis only occurs in disease-affected cells, suggesting that r(CCUG)<sub>300</sub> catalyzes the synthesis of its own inhibitor.

To confirm the extent of templated reaction and that r(CCUG)<sub>300</sub> was indeed the catalyst for oligomerization, we developed and implemented an approach termed ChemReactBIP (Chemical Reactivity and Binding Isolated by Pull Down; Figure 2A). Cells were treated with **K-Ak**, **N<sub>3</sub>-K-Ak**, or **N<sub>3</sub>-K-Aak** in the presence of a kanamycin analogue that contains a biotin moiety at the 6' position (**N<sub>3</sub>-K-Biotin**; Figures 20 & 10-12). **N<sub>3</sub>-K-Biotin** terminates the polymerization reaction *in cellulo* and allows oligomerized compounds and their bound cellular targets to be captured with streptavidin beads after gentle lysis. (Figures 2B & C). In agreement with studies completed with **N<sub>3</sub>-K** described above, oligomerized products were only formed in cells that express r(CCUG)<sub>300</sub> and that the conversion of monomer to products was greater with **N<sub>3</sub>-K-Aak** than **N<sub>3</sub>-K-Ak**, as expected.<sup>[11b]</sup> ChemReactBIP also allowed us to identify the

cellular targets of the templated reaction. A qRT-PCR analysis of the pulled down fraction confirmed that r(CCUG)<sub>300</sub> is a major target of the products of **N<sub>3</sub>-K-Biotin** and **K-Ak**, **N<sub>3</sub>-K-Ak**, or **N<sub>3</sub>-K-Aak** (Figure 2C). The amount of r(CCUG)<sub>300</sub> target pulled down increases as a function of potency: **N<sub>3</sub>-K + K-Ak** < **N<sub>3</sub>-K-Ak** < **N<sub>3</sub>-K-Aak**, *vide infra*. As observed in our other experiments, multivalent compounds were only observed when cells expressed r(CCUG)<sub>300</sub> (Figure 2C).

As observed in other microsatellite disorders, the binding of various proteins to r(CCUG)<sup>exp</sup> causes formation of nuclear foci.<sup>[3a]</sup> Fluorescence *in situ* hybridization (FISH) with a dye-labeled oligonucleotide was employed to determine if our compounds inhibit formation of r(CCUG)<sup>exp</sup>-containing nuclear foci (Figure 3A). In untreated cells, the average number of foci per cell is 9±2. Treatment with **N<sub>3</sub>-K** or **K-Ak** reduces the average number of foci per cell to 6±2 and 6±1, respectively. Treatment with an equimolar mixture of **K-Ak** and **N<sub>3</sub>-K** (**K-Ak + N<sub>3</sub>-K**) reduced foci to 1±1 per cell, similar to pre-synthesized dimers that mimic the reaction products of **K-Ak** and **N<sub>3</sub>-K** (**K 1,4 dimer** and **K 1,5 dimer**; Figures. 16 and 17). **N<sub>3</sub>-K-Ak** and **N<sub>3</sub>-K-Aak**, derivatives that click to self-oligomerize *in cellulo*, are even more potent, with <<1 foci per cell observed. Thus, the potential extent of oligomerization correlates with bioactivity.

Disruption of nuclear foci suggests that proteins sequestered by r(CCUG)<sup>exp</sup> are being freed. Thus, it is likely that the compounds also improve alternative pre-mRNA splicing defects caused by sequestration of MBNL1 in foci.<sup>[3a]</sup> We studied the ability of our compounds to improve dysregulation of the bridging integrator 1 (*BIN1*) pre-mRNA (Figures 3B-C, 13, and 14)<sup>[14]</sup>. When r(CCUG)<sub>300</sub> is present in cells, exon 11 is skipped too frequently, resulting in an inclusion rate of ~20% in the mature mRNA. In contrast, the exon 11 inclusion rate is ~60% in unaffected cells. Interestingly, the trend observed for compound potency as measured by disruption of foci is mirrored by our results for improvement of the *BIN1* alternative splicing defect (Figures 2 & 3). For example, pre-synthesized dimers and **N<sub>3</sub>-K + K-Ak** improve *BIN1* alternative splicing patterns to a similar extent while **N<sub>3</sub>-K-Ak** and **N<sub>3</sub>-K-Aak** are more potent (Figure 3C). Notably, **N<sub>3</sub>-K-Aak** retains the ability to improve the splicing defect at nanomolar concentrations, a >100-fold improvement over our first generation modularly assembled compound<sup>[6a]</sup> and >1000-fold better than monomers **N<sub>3</sub>-K** and **K-Ak**. Importantly, compounds do not affect *BIN1* splicing patterns in cells that do not express r(CCUG)<sub>300</sub> (Figure 15). Collectively, our compounds that participate in a templated click reaction are the most potent

non-covalent compounds that improve DM2-associated spliceopathy known to date. Furthermore, improvements in pre-mRNA splicing are due to binding r(CCUG)<sub>300</sub> and not to a non-specific effect.

In summary, these studies establish that an HDCR can be used to template the synthesis of an RNA inhibitor *in cellulo* and only in disease-affected cells. There may be numerous applications of this technology, as has been previously articulated by Sharpless and colleagues.<sup>[10]</sup> RNA repeating disorders may be a particularly attractive application. The nature of the target, which could have thousands of repeating units, could produce high yields of templated products. Moreover, many RNA gain-of-function disorders such as ALS, DM1, and DM2 cause brain dysfunction, making it important to develop low molecular weight compounds that have potential to cross the blood-brain barrier. Taken together, the click reaction could engender highly permeable low molecular weight monomers with potencies of multivalent compounds in both cellular and tissue models of disease.

This approach was also applied to target the causative agent of DM1 (r(CUG)<sup>exp</sup>), which has a similar disease mechanism as observed for DM2. DM1 is caused by r(CUG)<sup>exp</sup> located in the 3' UTR of *DMPK*, which also binds to and sequesters proteins that are involved in RNA biogenesis such as MBNL1 (Figure 21).

We previously described a dimeric compound that contain a H RNA binding module displayed on a *N*-methyl peptide backbone, **2H-K4NMe**, that binds r(CUG)<sup>exp</sup> and ameliorates DM1 dysfunction in cells and a DM1 mouse model. (Figure 22; Rzuczek SG, Gao Y, Tang ZZ, Thornton CA, Kodadek T, Disney MD. *ACS Chem Biol*. 2013 Oct 18;8(10):2312-21). After exposure to rat microsomes for 2 h, only 35% of the compound was intact. Inspection of the compound suggested that the amide bond could be a metabolic liability; thus the amide was replaced with an *N*-methyl amide to produce **2H-K4NMeS**, which was found to be stable to rat microsomes (Figure 22).

This stable compound, **2H-K4NMeS**, was tested for improving pre-mRNA splicing defects in DM1 patient-derived fibroblasts (see: Konieczny, P.; Stepniak-Konieczna, E.; Sobczak, K. *Nucleic Acids Research* 2014, 42, 10873.) and found to rescue >50% of the defect at concentrations greater than 100 nM (Figure 23). **2H-K4NMeS** (metabolically stable) and **2H-K4NMe** (metabolically unstable) were next tested for improving pre-mRNA splicing defects in the HSA<sup>LR</sup> mouse model (see: Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R.,

Henderson, D., Krym, M., and Thornton, C. A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 289, 1769–1773.). Results show that the **2H-K4NMeS** is >7-fold more potent at improving a *Clnl* pre-mRNA splicing defect in the mouse model compared to **2H-K4NMe** (Figure 23).

Given the good activity of **2H-K4NMeS**, the cellular targets of this molecule were probed by using chemical cross-linking and isolation by pull down (Chem-CLIP). In this approach, a reactive derivative of the parent compound **2H-K4NMeS-CA-Biotin** that contains a reactive module (CA) and a purification tag (Biotin). The structure of the Chem-CLIP probe is shown in Figure 24A. Testing of this compounds for *in vitro* reaction with RNA targets was completed and showed that the compound most effectively reacted with r(CUG)<sub>109</sub> (Figure 24B).

Application of this compound to DM1 patient-derived fibroblasts was next completed to profile the RNA targets of the compound. After compound exposure and cell lysis, the total RNA was captured onto streptavidin resin. The bound, biotinylated materials were released from the resin and quantified by using qRT-PCR. These studies showed that *DMPK* mRNA, which contained r(CUG)<sup>exp</sup> in the 3'UTR is enriched significantly in the pull down (Figure 24C). Addition of unreactive and excess of **2H-K4NMeS** over **2H-K4NMeS-CA-Biotin** and repeating Chem-CLIP – this experiment is termed competitive chemical cross-linking and isolation by pull down (C-Chem-CLIP\_ shows that the unreactive compound depletes the amount of *DMPK* in the pulled down fraction (Figure 24D).

Next, Chemical Crosslinking and Isolation by Pull Down to Map Ligand Binding Sites (Chem-CLIP-Map) was performed to identify the binding site for **2H-K4NMeS-CA-Biotin** in the *DMPK* mRNA (Yang, W.-Y., Wilson, H. D., Velagapudi, S. P., and Disney, M. D. (2015) Inhibition of non-ATG translational events in cells via covalent small molecules targeting RNA, *J. Am.Chem. Soc.* 137, 5336-5345.). This was completed by precisely cleaving the *DMPK* mRNA pulled down by **2H-K4NMeS-CA-Biotin** with antisense oligonucleotides and RNase H. The cleaved fragments that were biotinylated by reaction with **2H-K4NMeS-CA-Biotin** were captured onto streptavidin resin and purified. Quantification showed that the 3' UTR r(CUG)<sup>exp</sup>-containing fragment of the *DMPK* mRNA was significantly enriched in the pulled-down fraction, validating the expected binding site of the compound.

Next, a derivative of **2H-K4NMeS** was synthesized such that it could be used in a cellular click reaction to synthesize dimeric compounds targeting r(CUG)<sup>exp</sup> on-site, by analogy to the

experiments presented herein for the DM2 RNA. A series of compounds were synthesized in which: (i) a single azide was displayed (**N<sub>3</sub>-2H-K4NMeS**); and (ii) activated alkynes were displayed at varying distances from the *N*-methyl backbone (**2H-K4NMeS-Activated Alkynes**). These structures are shown in Figure 26.

Initially, we studied if r(CUG)<sub>12</sub> could catalyze oligomerization of the structures shown in Figure 26 via an HDCR between two compounds bound to adjacent sites. Results show that **2H-K4NMeS-Aak** provided the best alkyne spacer unit to react with **N<sub>3</sub>-2H-K4NMeS** in the presence of r(CUG)<sub>12</sub> (Figure 27A). The optimal compounds were then tested for their union being catalyzed by using other nucleic acids (Figure 27B) and it was shown that only r(CUG)<sub>12</sub> provided significant reaction yield.

Next, the lessons in spacing developed above were used to design a small molecule that had reactive azide and alkyne moieties in a single small molecule. The compound **N<sub>3</sub>-2H-K4NMeS-Aak** was thus synthesized (Figure 28). This compound should be capable of forming larger oligomers in the presence of r(CUG)<sup>exp</sup> than a mixture of the corresponding azide and alkyne tagged compounds.

The extent of **N<sub>3</sub>-2H-K4NMeS-Aak** oligomerization in patient-derived cells was tested by using ChemReactBIP. In this approach, a version of **N<sub>3</sub>-2H-K4NMeS-Aak** was synthesized (**Biotin-N<sub>3</sub>-2H-K4NMeS-Aak**, Figure 29A) that allowed for both starting materials and products of a reaction to be purified from cells via streptavidin capture (Figure 29B). These studies showed that the oligomeric product was only observed in DM1-derived fibroblasts (Figure 29C). Since the above data showed that on-site drug synthesis for DM1 was viable (Figure 29), the biological activity of this compound (**N<sub>3</sub>-2H-K4NMeS-Aak**) was assessed. Addition of the **N<sub>3</sub>-2H-K4NMeS-Aak** improved DM1-associated pre-mRNA splicing defects by >50% at 10 nM dosage, much more potent than treatment with **N<sub>3</sub>-2H-K4NMeS** alone or co-treatment of **N<sub>3</sub>-2H-K4NMeS** and **2H-K4NMeS-Aak** (Figure 30A). Next, the effects of the compounds on RNA foci were measured, showing that far fewer foci are observed when the clickable compound **N<sub>3</sub>-2H-K4NMeS-Aak** is added to cells (Figure 30B). Lastly, r(CUG)<sup>exp</sup>-containing *DMPK* mRNA is translated at much lower levels as it is retained in the nucleus (Childs-Disney, J. L.; Hoskins, J.; Rzuczek, S. G.; Thornton, C. A.; Disney, M. D. *ACS Chemical Biology* 2012.). Thus, **N<sub>3</sub>-2H-K4NMeS-Aak** may alter in the subcellular localization of *DMPK* by alleviating sequestration in foci and stimulate translation. Addition of **N<sub>3</sub>-2H-K4NMeS-Aak** improved the *DMPK*



translational defect at nanomolar concentrations in cells with long, toxic r(CUG) repeats, but has no effect on cells that did not have pathogenic repeats.

In summary, these data show that lead optimization of compounds can be completed by varying the backbone to make them more stable. Equipping these compounds with azide and alkyne moieties allows for on-site drug synthesis in DM1 patient-derived cells and affords very high potency.

Next, an approach to cleave r(CUG)<sup>exp</sup> was developed. It has been shown that the natural product bleomycin can cleave DNA but has also been shown to cleave RNA targets *in vitro* (Carter, B. J.; de Vroom, E.; Long, E. C.; van der Marel, G. A.; van Boom, J. H.; Hecht, S. M. *Proceedings of the National Academy of Sciences* 1990, 87, 9373.). Thus appendage of bleomycin to **2H-K4NMeS** could equip these compounds with the ability to cleave r(CUG)<sup>exp</sup> in patient-derived cells. Synthesis of **2H-K4NMeS-Bleomycin A5** was thus completed (Figure 31A). Application of this compound to DM1 patient-derived fibroblasts results in reduced levels of *DMPK* mRNA as assessed by qRT-PCR but has no effect on *DMPK* mRNA levels in healthy fibroblasts (Figure 31B). To further test the ability of the compound to selectively cleave *DMPK* mRNA, 4-fold excess of **2H-K4NMeS** vs. **2H-K4NMeS-Bleomycin A5** was added to cells. As expected, the unreactive compound competed with the reactive compound for binding, reducing the amount of cleavage and suggesting selectivity (Figure 31C). These cleavage experiments also correlated with improvement of pre-mRNA splicing defects (Figure 31D).

Fluorescent reporters were also synthesized on-site based on the RNA-catalyzed click reaction. In this approach, compounds that can undergo a click reaction are tagged with FRET pairs (Figure 32A) such that FRET can be observed upon a click reaction. FAM (fluorescein) and TAMARA (5-Carboxytetramethylrhodamine) FRET dye pairs were used, affording **FAM-2H-K4NMeS-Aak** and **N<sub>3</sub>-2H-K4NMeS-TAMARA**. Addition of the compounds showed enhancement in FRET only in the presence of r(CUG)<sub>12</sub>. Thus, the approach allows for the on-site synthesis of FRET reporters that will have broad applicability.

#### Documents Cited

- [1] M. L. Yeung, Y. Bennasser, K. T. Jeang, *Curr. Med. Chem.* 2007, 14, 191-197.
- [2] M. E. MacDonald, J. F. Gusella, *Curr. Opin. Neurobiol.* 1996, 6, 638-643.

- [3] a) A. Mankodi, P. Teng-Umnuaay, M. Krym, D. Henderson, M. Swanson, C. A. Thornton, *Ann. Neurol.* 2003, *54*, 760-768; b) H. T. Orr, H. Y. Zoghbi, *Annu. Rev. Neurosci.* 2007, *30*, 575-621.
- [4] a) J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* 2008, *108*, 1171-1224; b) L. Guan, M. D. Disney, *ACS Chem. Biol.* 2012, *7*, 73-86.
- [5] J. R. Brouwer, R. Willemsen, B. A. Oostra, *BioEssays* 2009, *31*, 71-83.
- [6] a) J. L. Childs-Disney, I. Yildirim, H. Park, J. R. Lohman, L. Guan, T. Tran, P. Sarkar, G. C. Schatz, M. D. Disney, *ACS Chem. Biol.* 2014, *9*, 538-550; b) M. D. Disney, L. P. Labuda, D. J. Paul, S. G. Poplawski, A. Pushechnikov, T. Tran, S. P. Velagapudi, M. Wu, J. L. Childs-Disney, *J. Am. Chem. Soc.* 2008, *130*, 11185-11194.
- [7] C. L. Liquori, K. Ricker, M. L. Moseley, J. F. Jacobsen, W. Kress, S. L. Naylor, J. W. Day, L. P. Ranum, *Science* 2001, *293*, 864-867.
- [8] B. Llano-Sotelo, E. F. Azucena, Jr., L. P. Kotra, S. Mobashery, C. S. Chow, *Chem. Biol.* 2002, *9*, 455-463.
- [9] J. L. Childs-Disney, J. Hoskins, S. G. Rzuczek, C. A. Thornton, M. D. Disney, *ACS Chem. Biol.* 2012, *7*, 856-862.
- [10] W. G. Lewis, L. G. Green, F. Grynszpan, Z. Radic, P. R. Carlier, P. Taylor, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed. Engl.* 2002, *41*, 1053-1057.
- [11] a) A. Krasinski, Z. Radic, R. Manetsch, J. Raushel, P. Taylor, K. B. Sharpless, H. C. Kolb, *J. Am. Chem. Soc.* 2005, *127*, 6686-6692; b) A. T. Poulin-Kerstien, P. B. Dervan, *J. Am. Chem. Soc.* 2003, *125*, 15811-15821.
- [12] M. D. Disney, J. L. Childs-Disney, *ChemBiochem* 2007, *8*, 649-656.
- [13] M. M. Lee, J. L. Childs-Disney, A. Pushechnikov, J. M. French, K. Sobczak, C. A. Thornton, M. D. Disney, *J. Am. Chem. Soc.* 2009, *131*, 17464-17472.
- [14] C. Fugier, A. F. Klein, C. Hammer, S. Vassilopoulos, Y. Ivarsson, A. Toussaint, V. Tosch, A. Vignaud, A. Ferry, N. Messaddeq, Y. Kokunai, R. Tsuburaya, P. de la Grange, D. Dembele, V. Francois, G. Precigout, C. Boulade-Ladame, M. C. Hummel, A. L. de Munain, N. Sergeant, A. Laquerriere, C. Thibault, F. Deryckere, D. Auboeuf, L. Garcia, P. Zimmermann, B. Udd, B. Schoser, M. P. Takahashi, I. Nishino, G. Bassez, J. Laporte, D. Furling, N. Charlet-Berguerand, *Nat. Med.* 2011, *17*, 720-725.

## **Examples**

## Methods for computational analysis

We previously published a model of the binding of **2K-4**, two **K** RNA-binding modules linked by a peptoid, to an RNA containing two 5'CCUG/3'GUCC loops<sup>[1]</sup>. The peptoid was removed and the **K** RNA-binding modules were edited to include linkers: (i) the C6' of ring I in one **K** contained either hex-5-ynamide or N-(2-amino-2-oxoethyl)propiolamide; and (ii) an azido group was added to C6'' of ring III of the other **K**. Conformational scanning of the linkers was conducted to check whether the reactive termini of the two **Ks** come within close proximity of each other. The exhaustive non-redundant conformational scanning was conducted with MacroModel (Schrödinger, LLC) using the OPLS\_2005 force field<sup>[2]</sup>. The result showed multiple conformations bring the two reactive groups in close proximity with the shortest ethynyl C to azido N distances of 2.61 Å and 2.37 Å in hex-5-ynamide and N-(2-amino-2-oxoethyl)propiolamide linkers, respectively (Figure 1, BI).

The linker sets with the closest reactive groups in each structure were used to generate models for cycloaddition end products, **K 1,4-dimer** and **K 1,5-dimer**. The **K 1,5-dimer** for each linker combination (hex-5-ynamide + N<sub>3</sub> or N-(2-amino-2-oxoethyl)propiolamide + N<sub>3</sub>) has one conformation. The **K 1,4-dimer** for each linker combination showed two major conformations were possible, either the triazole ring C<sub>5</sub>H pointing toward solvent (C<sub>5</sub>H-out) or toward RNA major groove (C<sub>5</sub>H-in). A total of six models were thus produced (Figure 1, BII & BIII). Each structure was briefly energy minimized followed by molecular dynamics (MD) simulations in order to observe the stability of the complex and the effects of the linkers on the bound **Ks**. MD simulations were conducted using Desmond<sup>[3]</sup> with the OPLS\_2005 force field<sup>[2]</sup>. Each model structure was solvated in an orthorhombic box with TIP3 water and 150 mM NaCl, which buffers 10 Å in each direction. Additional ions were added to neutralize the system. Before MD simulations, the systems were pre-equilibrated using the NPT relaxation protocol, which consists of restrained/unrestrained minimizations and short simulations with isothermal and isobaric ensemble. The 10 ns MD simulations were performed at constant temperature (300K) and pressure (1.01325 bar). Positional restraints were applied to the RNA throughout the simulation. Short- and long-range Coulombic interactions were set to Cutoff method with 9 Å radius (short) and smooth particle mesh Ewald tolerance method with a tolerance of  $1 \times 10^{-9}$  (long). Analyses of the simulations were completed with Maestro graphics interface.

No significant fluctuations of system volume, pressure, temperature and potential energy were observed. Total energies of the systems were stabilized after about ~4 – 6 ns in most of the simulations except in two cases where the value continuously increased. However, the slopes were less than 0.05 kcal/ps during the simulation periods. Hydrogen bonds between **K** moieties and the RNA were preserved throughout the simulations. Figures were prepared using PyMol (Schrödinger, LLC).

### Methods for Compound Synthesis & Compound Characterization

*Abbreviations.* DIC, *N,N'*-Diisopropylcarbodiimide; DIEA, *N,N*-Diisopropylethylamine; DMF, *N,N*-dimethylformamide; HPLC, high performance liquid chromatography; HRMS, high resolution mass spectrometry; LC-MS, liquid chromatography-mass spectrometry; MeOH, methanol; MALDI ToF/ToF, matrix-assisted laser desorption/ionization time of flight/time of flight; MS, mass spectrometry; NBD, 7-nitrobenz-2-oxa-1,3-diazole-4-yl; TFA, trifluoroacetic acid

*Synthesis.* Fmoc-Rink amide resin (0.59 mmol/g) was purchased from Advanced ChemTech. *N,N*-dimethylformamide (DMF, anhydrous) was purchased from EMD and used without further purification. Piperidine, trifluoroacetic acid (TFA), *N,N*-diisopropylethyl amine (DIEA), and 2-bromoacetic acid were purchased from Sigma Aldrich. *N,N'*-diisopropylcarbodiimide (DIC), 1-*H*-hydroxy-7-azabenzotriazole (HOAt), and Fmoc- $\beta$ -alanine were purchased from Advanced ChemTech. Fmoc-*N*-methyl-L-alanine and *N*-(4-aminobutyl)-*N*-methyl carbamic acid tert-butyl ester were purchased from Combi-Blocks. *N*-(4-aminoethyl)-*N*-methyl carbamic acid tert-butyl ester was purchased from Oakwood Products. Chlorambucil was purchased from MP Biomedicals. Bleomycin A5 was purchased from LKT Laboratories. Hoechst carboxylate (Pushechnikov A, Lee MM, Childs-Disney JL, Sobczak K, French JM, Thornton CA, Disney MD. *J Am Chem Soc.* 2009 Jul 22;131(28):9767-79), 2H-K4NMe (Rzuczek SG, Gao Y, Tang ZZ, Thornton CA, Kodadek T, Disney MD. *ACS Chem Biol.* 2013 Oct 18;8(10):2312-21), and biotin amine (Yamada, M.; Harada, K.; Maeda, Y.; Hasegawa, T. *New Journal of Chemistry* 2013, 37, 3762.) were synthesized as reported previously. *N*-methyl peptides were synthesized using a Biotage Initiator+ SP Wave microwave.

*Compound purification and analysis.* Preparative HPLC was performed using a Waters 1525 Binary HPLC pump equipped with a Waters 2487 dual absorbance detector system and a Waters Sunfire C18 OBD 5  $\mu$ m 19 x 150 mm column. Absorbance was monitored at 280 and

220 nm. A gradient of 0-100% MeOH in H<sub>2</sub>O with 0.1% TFA over 60 min was used for compound purification. Analytical HPLC was performed using a Waters Symmetry C18 5  $\mu$ m 4.6 x 150 mm column. All compounds evaluated had  $\geq$ 95% purity as determined by analytical HPLC. Mass spectrometry was performed with an Applied Biosystems MALDI ToF/ToF Analyzer 4800 Plus using an  $\alpha$ -hydroxycinnamic acid matrix. All microwave reactions were performed using a Biotage initiator+ SP Wave microwave. High resolution mass spectral analysis was performed by the University of Illinois Urbana-Champaign School of Chemical Sciences Mass Spectrometry Center.

*Synthesis of **K 1, 4 Dimer**. See Figure 16.* **N<sub>3</sub>-K** (50 mg, 50  $\mu$ mol) and **K-Ak** (45 mg, 50  $\mu$ mol) were dissolved in a 1:1 mixture of DMF and water (3 mL) and treated with polytriazole Cu(I) catalyst (10 mg, 16  $\mu$ mol)<sup>[4]</sup> and Et<sub>3</sub>N (500  $\mu$ L, 3.6 mmol). The reaction mixture was microwaved at 100 °C for 2.5 h, after which the reaction mixture was concentrated and purified by reverse phase HPLC as described above. Yield: 11%; 6.2 mg of white solid as a TFA salt. <sup>1</sup>H NMR (700 MHz, D<sub>2</sub>O, TSP):  $\delta$  4.31 (m, 2H), 3.99 (m, 2H), 3.92 (m, 2H), 3.75 (m, 9H), 3.54 (m, 7H), 3.41 (t, 2H,  $J$ =10 Hz), 3.31 (m, 2H), 3.24 (dd, 2H,  $J$ =7, 13), 2.73 (t, 2H,  $J$ =8 Hz), 2.52 (m, 2H), 2.34 (t, 2H,  $J$ =8 Hz), 1.92 (m, 4H), 1.25 (m, 4H); <sup>13</sup>C NMR (175 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 179.8, 166.2, 166.0, 165.8, 165.6, 150.3, 132.1, 127.8, 121.8, 120.1, 118.4, 116.8, 103.7, 103.5, 101.2, 100.3, 86.6, 82.4, 81.7, 76.0, 75.7, 75.0, 74.9, 74.1, 74.0, 73.7, 73.6, 72.9, 72.5, 71.6, 71.1, 70.9, 69.5, 68.2, 62.6, 57.9, 53.1, 52.9, 52.6, 51.2, 50.9, 43.2, 42.1, 38.0, 31.2, 30.7, 30.6, 28.1, 27.0 ppm; HRMS (ESI)  $m/z$  calculated for C<sub>42</sub>H<sub>79</sub>N<sub>12</sub>O<sub>11</sub> (M+H) 1087.5483; found 1087.5453.

*Synthesis of **K 1,5 Dimer**. See Figure 17.* **Boc-N<sub>3</sub>-K** (20 mg, 20  $\mu$ mol) and **Boc-K-Ak** (23 mg, 20  $\mu$ mol) were dissolved in toluene (2 mL) and placed under argon. The sample was treated with pentamethylcyclopentadienylbis(triphenylphosphine)ruthenium(II) chloride (6 mg, 0.8  $\mu$ mol) and microwaved at 130 °C for 3 h. The reaction mixture was then filtered and concentrated. The residue was treated with a 1:1 mixture of TFA in CH<sub>2</sub>Cl<sub>2</sub> at 0 °C for 2 h. The reaction mixture was concentrated and purified by reverse phase HPLC as described above. Yield: 0.5%; 1 mg of white solid as a TFA salt. <sup>1</sup>H NMR (700 MHz, D<sub>2</sub>O, TSP):  $\delta$  7.66 (s, 1H), 3.52 (m, 32H), 2.53 (m, 2H), 2.37 (m, 2H), 2.24 (m, 2H), 1.98 (m, 2H), 1.87 (m, 4H), 1.34 (m, 4H); <sup>13</sup>C NMR (175 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 179.8, 166.2, 166.0, 165.8, 165.6, 121.8, 120.1, 118.4, 116.8, 103.5, 103.4, 101.0, 98.4, 87.6, 86.5, 82.3, 81.2, 76.0, 75.9, 75.8, 75.7, 75.1, 74.1, 74.0, 73.6, 73.0, 72.9, 71.7, 71.6, 71.1, 68.2, 62.6, 57.9, 57.3, 52.6, 51.1, 51.0, 50.8, 50.5, 45.5,

43.2, 42.1, 37.5, 30.9, 30.6, 27.0, 20.6, 19.9, 19.2, 15.1 ppm; HRMS (ESI)  $m/z$  calculated for  $C_{42}H_{78}N_{11}O_{22}$  (M+H) 1088.5323; found 1088.5316.

*Synthesis of N<sub>3</sub>-K-Ak.* See Figure 18. **N<sub>3</sub>-K** (25 mg, 49  $\mu$ mol) was dissolved in a 1:1 mixture of acetone and water (2 mL) and NBD-activated 6-aminohexynoic acid (6 mg, 49  $\mu$ mol) was added. The reaction was stirred at room temperature overnight and then the solvent was removed *in vacuo*. The resulting residue was purified by reverse phase HPLC as described above. Yield: 20%; 5 mg of white solid as a TFA salt.  $^1H$  NMR (700 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 4.05 (m, 1H), 3.91 (m, 4H), 3.78 (t, 2H,  $J=9$ ), 3.74 (dd, 2H,  $J=4, 14$ ), 3.70 (m, 4H), 3.66 (m, 2H), 3.61 (dd, 1H,  $J=5, 15$ ), 3.55 (dd, 2H,  $J=5, 12$ ), 3.48 (m, 2H), 3.34 (t, 1H,  $J=10$ ), 2.71 (s, 1H), 2.50 (d, 2H,  $J=12$ ), 1.86 (m, 2H), 1.44 (t, 1H,  $J=7$ );  $^{13}C$  NMR (175 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 166.2, 166.0, 165.8, 165.6, 162.2, 121.8, 120.1, 119.6, 118.4, 118.0, 116.8, 103.6, 102.0, 86.5, 76.1, 75.2, 74.3, 73.9, 73.5, 71.1, 69.1, 57.7, 53.3, 52.9, 43.2, 41.6, 19.4, 19.0 ppm; HRMS (ESI)  $m/z$  calculated for  $C_{24}H_{42}N_7O_{11}$  (M+H) 604.2942; found 604.2938.

*Synthesis of N<sub>3</sub>-K-Aak.* See Figure 19. 2-Propiolamidoacetic acid (175 mg, 1.36 mmol) was dissolved in anhydrous DMF and was treated with DIC (213  $\mu$ L, 1.4 mmol), *N*-hydroxy-5-norborene-2,3-dicarboximide (243 mg, 1.36 mmol), and DIEA (600  $\mu$ L, 3.4 mmol). The reaction mixture was stirred under argon at room temperature overnight. Next, **N<sub>3</sub>-K** (20 mg, 40  $\mu$ mol) was dissolved in a 1:1 mixture of acetone and water (2 mL), and 500  $\mu$ L of the NBD-activated acid was added gradually over 4 h. The solvent was then removed *in vacuo*, and the resulting residue was purified by reverse phase HPLC as described above. Yield: 2%, 600  $\mu$ mol.  $^1H$  NMR (700 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 4.06 (m, 1H), 3.98 (d, 1H,  $J=5$ ), 3.94 (dd, 1H,  $J=3, 11$ ), 3.83 (t, 2H,  $J=10$ ), 3.52 (m, 14H), 2.43 (d, 2H,  $J=11$ ), 1.73 (m, 2H);  $^{13}C$  NMR (175 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 169.8, 166.3, 166.0, 165.8, 165.6, 131.9, 126.0, 121.8, 120.1, 118.4, 116.8, 103.4, 76.5, 75.2, 74.2, 74.2, 74.1, 73.8, 73.7, 73.4, 71.6, 71.2, 69.4, 69.2, 57.7, 53.6, 53.3, 53.2, 53.0, 51.5, 45.5, 43.2, 42.7 ppm; HRMS (ESI)  $m/z$  calculated for  $C_{23}H_{39}N_8O_{12}$  (M+H) 619.2687; found 619.2680.

*Synthesis of N<sub>3</sub>-K-Biotin.* See Figure 20. **N<sub>3</sub>-K** (17 mg, 30  $\mu$ mol) was dissolved in a 1:1 mixture of acetone and water (2 mL) and NBD-biotin (12 mg, 30  $\mu$ mol) was added. The reaction was stirred at room temperature overnight and then the solvent was removed *in vacuo*. The resulting residue was purified by reverse phase HPLC as described above. Yield: 7%; 1.6 mg of white solid as a TFA salt.  $^1H$  NMR (700 MHz, D<sub>2</sub>O, TSP):  $\delta$  = 4.60 (m 1H), 4.41 (m, 1H), 4.05

(m, 1H), 3.91 (dd, 2H,  $J=4$ , 11), 3.83 (m, 2H), 3.67 (m, 5H), 3.55 (dd, 2H,  $J=5$ , 14), 3.45 (m, 2H), 3.31 (m, 2H), 3.19 (m, 1H), 2.98 (dd, 1H,  $J=8$ , 13), 2.77 (d, 1H,  $J=13$ ), 2.38 (m, 1H), 2.28 (t, 2H,  $J=7$ ), 1.65 (m, 5H), 1.39 (m, 2H), 1.32 (m, 2H);  $^{13}\text{C}$ NMR (175 MHz,  $\text{D}_2\text{O}$ , TSP):  $\delta$  = 177.4, 177.0, 165.3, 163.3, 162.9, 162.5, 162.2, 134.3, 120.5, 117.6, 114.7, 111.8, 100.5, 98.3, 83.4, 79.9, 73.0, 72.2, 71.4, 71.2, 71.1, 70.1, 68.1, 66.2, 62.0, 60.2, 55.4, 54.7, 51.1, 50.4, 49.8, 48.1, 44.3, 42.6, 39.6, 39.1, 35.3, 27.8, 27.6, 25.1 ppm; HRMS (ESI)  $m/z$  calculated for  $\text{C}_{28}\text{H}_{50}\text{N}_9\text{O}_{12}\text{S}$  (M+H) 736.3300; found 736.3297.

### Synthesis of Small Molecule Dimers.

*Synthesis of 2H-K4NMeS.* See Figure 33. Rink amide resin (500 mg, 0.3 mmol) was swollen in DMF at room temperature for 10 min and then deprotected with a solution of 20% piperidine in DMF (5 mL, 2 x 20 min). The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250  $\mu\text{L}$ , 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminobutyl)-*N*-methyl carbamic acid tert-butyl ester (121 mg, 0.6 mmol) in DMF (4 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Then a solution of Fmoc-*N*-methyl-L-alanine (264 mg, 0.81 mmol), DIC (250  $\mu\text{L}$ , 1.6 mmol), HOAt (110 mg, 0.81 mmol), and DIEA (141  $\mu\text{L}$ , 0.81 mmol) in DMF (4 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). This cycle was repeated a total of four times. Next the resin was washed with DMF and DCM and then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (80 mg, 0.16 mmol), HOAt (22 mg, 0.16 mmol), DIC (25  $\mu\text{L}$ , 0.16 mmol) and DIEA (100  $\mu\text{L}$ ) in DMF (2 mL) and heated via microwave to 75 °C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/ $\text{H}_2\text{O}$  + 0.1% (v/v) TFA over 1 h. Isolated 640 nanomoles of 2H-K4NMeS; 0.2%. 2H-K4NMeS ( $\text{C}_{81}\text{H}_{101}\text{N}_{19}\text{O}_9$ ) MS calculated (M+H) 1484.81, MS found 1484.93 (M+H);  $t_{\text{R}}$  = 32 min.

*Synthesis of 2H-K4NMeS intermediate for functionalization.* Rink amide resin (500 mg, 0.3 mmol) was swollen in DMF at room temperature for 10 min and then deprotected with a

solution of 20% piperidine in DMF (5 mL, 2 x 20 min). The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250  $\mu$ L, 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminobutyl)-*N*-methyl carbamic acid tert-butyl ester (121 mg, 0.6 mmol) in DMF (4 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Then a solution of Fmoc-*N*-methyl-L-alanine (264 mg, 0.81 mmol), DIC (250  $\mu$ L, 1.6 mmol), HOAt (110 mg, 0.81 mmol), and DIEA (141  $\mu$ L, 0.81 mmol) in DMF (4 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). This cycle was repeated a total of three times. The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250  $\mu$ L, 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminoethyl)-*N*-methyl carbamic acid tert-butyl ester (105 mg, 0.6 mmol) in DMF (4 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL).

*Synthesis of 2H-K4NMeS-CA-Biotin.* See Figure 34. The **2H-K4NMeS** intermediate resin prepared above was reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250  $\mu$ L, 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of biotin amine (200 mg, 0.6 mmol) in DMF (4 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Then a solution of Fmoc- $\beta$ -alanine (252 mg, 0.81 mmol), DIC (250  $\mu$ L, 1.6 mmol), HOAt (110 mg, 0.81 mmol), and DIEA (141  $\mu$ L, 0.81 mmol) in DMF (4 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and DCM and then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (80 mg, 0.16 mmol), HOAt (22 mg, 0.16 mmol), DIC (25  $\mu$ L, 0.16 mmol) and DIEA (100  $\mu$ L) in DMF (2 mL) and heated



via microwave to 75 ° C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. Then the Fmoc was removed with 20% piperidine/DMF (1 x 10 min) and then concentrated *in vacuo*. Isolated 143 nanomoles (0.05%) of free amine dimer which was treated with a solution of chlorambucil (1 mg, 3 μmoles), DIC (25 μL, 0.16 mmol) and DIEA (25 μL) in DMF (200 μL) at room temperature for 3 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. Isolated 119 nanomoles of **2H-K4NMeS-CA-Biotin**; 0.04%. **2H-K4NMeS-CA-Biotin** (C<sub>113</sub>H<sub>149</sub>Cl<sub>2</sub>N<sub>26</sub>O<sub>14</sub>S) MS calculated 2196.08 (M+H), MS found 2195.96 (M+H); t<sub>R</sub> = 30 min.

*Synthesis of 2NAc-K4NMeS-CA-Biotin.* See Figure 35. Synthesized as described above for **2H-K4NMeS-CA-Biotin** but instead of coupling with Hoechst carboxylate the peptide was treated with a 1:1 solution of acetic anhydride and DIEA (400 μL total volume) at room temperature for 1 h. Isolated 800 nanomoles of **2NAc-K4NMeS**; 0.3%). **2NAc-K4NMeS** (C<sub>59</sub>H<sub>96</sub>Cl<sub>2</sub>N<sub>14</sub>O<sub>12</sub>SNa) MS calculated 1317.63 (M+Na), MS found 1317.28 (M+Na); t<sub>R</sub> = 18 min.

*Synthesis of 2H-K4NMeS-Bleomycin A5.* See Figure 36. The **2H-K4NMeS** intermediate resin prepared above was reacted with a solution of 6-(Fmoc-amino)hexanoic acid (286 mg, 0.81 mmol) DIC (250 μL, 1.6 mmol), HOAt (110 mg, 0.81 mmol), and DIEA (141 μL, 0.81 mmol) in DMF (4 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and DCM and then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (80 mg, 0.16 mmol), HOAt (22 mg, 0.16 mmol), DIC (25 μL, 0.16 mmol) and DIEA (100 μL) in DMF (2 mL) and heated via microwave to 75 ° C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. Then the Fmoc was removed with 20% piperidine/DMF (1 x 10 min) and then concentrated *in vacuo*. Isolated 450 nmol (0.2%) of free amine dimer which was treated with a solution of *N,N'*-disuccinimidyl carbonate (450 nmol, LKT Labs) and DIEA (10 μL) in dry DMF (50 μL). The reaction stirred at room temperature and additional, *N,N'*-disuccinimidyl carbonate was added periodically until 80-90% of the starting material was consumed. Then a solution of bleomycin A5-copper complex (450 nmol) in dry DMF was added and the reaction stirred at

room temperature overnight. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC by first using 0.1 mM EDTA in water (pH 6.3) for 15 min and then 0-100% CH<sub>3</sub>CN/H<sub>2</sub>O + 0.1% TFA over 1 h. Isolated 52 nanomoles of **2H-K4NMeS-Bleomycin A5** (0.02%). **2H-K4NMeS-Bleomycin A5** (C<sub>146</sub>H<sub>203</sub>N<sub>40</sub>O<sub>32</sub>S<sub>2</sub>) MS calculated 3092.49 (M+H), MS found 3092.18 (M+H); *t<sub>R</sub>* = 35 min.

*Synthesis of N<sub>3</sub>-2H-K4NMeS.* See Figure 37. Rink amide resin (200 mg, 0.12 mmol) was swollen in DMF at room temperature for 10 min and then deprotected with a solution of 20% piperidine in DMF (3 mL, 2 x 20 min). The resin was washed with DMF (3 x 3 mL) and reacted twice with a solution of 1M bromoacetic acid (2 mL) in DMF and DIC (125 μL, 0.8 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 3 mL) and reacted with a solution of 3-azidopropylamine (60 mg, 0.6 mmol) in DMF (2 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 3 mL). The resin was again reacted twice with a solution of 1M bromoacetic acid (2 mL) in DMF and DIC (125 μL, 0.8 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 3 mL) and reacted with a solution of *N*-(4-aminobutyl)-*N*-methyl carbamic acid tert-butyl ester (60 mg, 0.3 mmol) in DMF (2 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 3 mL). Then a solution of Fmoc-*N*-methyl-L-alanine (132 mg, 0.4 mmol), DIC (125 μL, 0.8 mmol), HOAt (55 mg, 0.4 mmol), and DIEA (70 μL, 0.4 mmol) in DMF (2 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). This cycle was repeated a total of four times. Next the resin was washed with DMF and DCM and then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (40 mg, 0.08 mmol), HOAt (11 mg, 0.08 mmol), DIC (25 μL, 0.16 mmol) and DIEA (50 μL) in DMF (1 mL) and heated via microwave to 75 °C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. Isolated 640 nanomoles of **N<sub>3</sub>-2H-K4NMeS**;

0.5%. **N<sub>3</sub>-2H-K4NMeS** (C<sub>85</sub>H<sub>108</sub>N<sub>23</sub>O<sub>10</sub>) MS calculated (M+H) 1610.87, MS found 1610.96 (M+H); t<sub>R</sub> = 34 min.

*Synthesis of N<sub>3</sub>-2H-K4NMeS Biotin.* See Figure 38. Rink amide resin (500 mg, 0.3 mmol) was swollen in DMF at room temperature for 10 min and then deprotected with a solution of 20% piperidine in DMF (3 mL, 2 x 20 min). The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250 µL, 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of 3-azidopropylamine (120 mg, 1.2 mmol) in DMF (5 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). The resin was again reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250 µL, 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminobutyl)-*N*-methyl carbamic acid tert-butyl ester (121 mg, 0.6 mmol) in DMF (4 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Then a solution of Fmoc-*N*-methyl-L-alanine (264 mg, 0.81 mmol), DIC (250 µL, 1.6 mmol), HOAt (110 mg, 0.81 mmol), and DIEA (140 µL, 0.81 mmol) in DMF (5 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). This cycle was repeated a total of three times. The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (4 mL) in DMF and DIC (250 µL, 1.6 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminoethyl)-*N*-methyl carbamic acid tert-butyl ester (105 mg, 0.6 mmol) in DMF (4 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Next the resin was treated with a solution of D-biotin (195 mg, 0.81 mmol), DIC (250 µL, 1.6 mmol), HOAt (110 mg, 0.81 mmol), and DIEA (140 µL, 0.81 mmol) in DMF (5 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. Next the resin was washed with DMF and DCM and then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was

concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (40 mg, 0.08 mmol), HOAt (11 mg, 0.08 mmol), DIC (25  $\mu$ L, 0.16 mmol) and DIEA (50  $\mu$ L) in DMF (1 mL) and heated via microwave to 75 °C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. Isolated 640 nanomoles of **N<sub>3</sub>-2H-K4NMeS-Biotin**; 0.2%. **N<sub>3</sub>-2H-K4NMeS-Biotin** (C<sub>97</sub>H<sub>127</sub>N<sub>26</sub>O<sub>12</sub>S) MS calculated (M+H) 1879.98, MS found 1879.87 (M+H); t<sub>R</sub> = 31 min.

*Synthesis of 2H-K4NMeS Activated Alkynes.* See Figure 39. For each alkyne derivative 300 mg of resin (0.18 mmol) with **2H-K4NMeS** intermediate was treated with a solution of *N*-Fmoc carboxylate linker (0.6 mmol), DIC (125  $\mu$ L, 0.8 mmol), HOAt (72 mg, 0.6 mmol), and DIEA (80  $\mu$ L, 0.6 mmol) in DMF (2 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). Then the resin was washed with DMF and then suspended in 2 mL of DMF with DIEA (80  $\mu$ L, 0.6 mmol). The suspension was flushed with argon and cooled to 0 °C. Then DIC (125  $\mu$ L, 0.8 mmol) was added followed by slow addition of a solution of propiolic acid (37  $\mu$ L, 0.6 mmol) in DMF (500  $\mu$ L). The reaction stirred at low temperature for 1 h and then was washed with DMF and DCM. The resin was then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (40 mg, 0.08 mmol), HOAt (11 mg, 0.08 mmol), DIC (25  $\mu$ L, 0.16 mmol) and DIEA (50  $\mu$ L) in DMF (1 mL) and heated via microwave to 75 °C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h.

**2H-K4NMeS Propiolate** (n=0) 900 nmoles (0.5%) (C<sub>85</sub>H<sub>105</sub>N<sub>20</sub>O<sub>10</sub>) MS calculated (M+H) 1565.83, MS found 1565.61 (M+H); t<sub>R</sub> = 31 min. **2H-K4NMeS Glycine** (n=1) 600 nmoles (0.3%) (C<sub>87</sub>H<sub>108</sub>N<sub>21</sub>O<sub>11</sub>) MS calculated (M+H) 1622.854, MS found 1622.854 (M+H); t<sub>R</sub> = 30 min. **2H-K4NMeS  $\beta$ -Alanine** (n=2) 750 nmoles (0.4%) (C<sub>88</sub>H<sub>110</sub>N<sub>21</sub>O<sub>11</sub>) MS calculated (M+H) 1636.87, MS found 1636.7 (M+H); t<sub>R</sub> = 31 min. **2H-K4NMeS Aminohexanoate** (n=5, Aak) 4.93  $\mu$ moles (3%)(C<sub>91</sub>H<sub>116</sub>N<sub>21</sub>O<sub>11</sub>) MS calculated (M+H) 1678.92, MS found 1678.7 (M+H); t<sub>R</sub> = 31 min. **2H-K4NMeS Amino-octanoate** (n=7) 1.6  $\mu$ moles (0.9%) (C<sub>93</sub>H<sub>120</sub>N<sub>21</sub>O<sub>11</sub>) MS calculated (M+H) 1706.95, MS found 1707.0 (M+H); t<sub>R</sub> = 33 min. **2H-K4NMeS**

**Aminododecanoate** (n=11) 4.49  $\mu$ moles (3%) ( $C_{97}H_{128}N_{21}O_{11}$ ) MS calculated (M+H) 1763.01, MS found 1763.2 (M+H);  $t_R$  = 34 min.

*Synthesis of N<sub>3</sub>-2H-K4NMeS-Aak.* See Figure 40. Rink amide resin (800 mg, 0.47 mmol) was swollen in DMF at room temperature for 10 min and then deprotected with a solution of 20% piperidine in DMF (5 mL, 2 x 20 min). The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (5 mL) in DMF and DIC (500  $\mu$ L, 3.2 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of 3-azidopropylamine (300 mg, 3 mmol) in DMF (5 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). The resin was again reacted twice with a solution of 1M bromoacetic acid (5 mL) in DMF and DIC (500  $\mu$ L, 3.2 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminobutyl)-*N*-methyl carbamic acid tert-butyl ester (240 mg, 1.2 mmol) in DMF (5 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Then a solution of Fmoc-*N*-methyl-L-alanine (528 mg, 1.6 mmol), DIC (500  $\mu$ L, 3.2 mmol), HOAt (220 mg, 1.6 mmol), and DIEA (280  $\mu$ L, 1.6 mmol) in DMF (5 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). This cycle was repeated a total of three times. The resin was washed with DMF (3 x 5 mL) and reacted twice with a solution of 1M bromoacetic acid (5 mL) in DMF and DIC (500  $\mu$ L, 3.2 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 5 mL) and reacted with a solution of *N*-(4-aminoethyl)-*N*-methyl carbamic acid tert-butyl ester (210 mg, 1.2 mmol) in DMF (5 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). Then the resin was treated with a solution of *N*-Fmoc aminohexanoic acid (565 mg, 1.6 mmol), DIC (500  $\mu$ L, 3.2 mmol), HOAt (220 mg, 1.6 mmol), and DIEA (280  $\mu$ L, 1.6 mmol) in DMF (5 mL) was added and the reaction was heated via microwave to 75 °C for 30 min. The resin was washed with DMF and the Fmoc was removed with 20% piperidine/DMF (2 x 10 min). Then the resin

was washed with DMF and then suspended in 5 mL of DMF with DIEA (280  $\mu$ L, 1.6 mmol). The suspension was flushed with argon and cooled to 0 ° C. Then DIC (500  $\mu$ L, 3.2 mmol) was added followed by slow addition of a solution of propionic acid (148  $\mu$ L, 3.2 mmol) in DMF (500  $\mu$ L). The reaction stirred at low temperature for 1 h and then was washed with DMF and DCM. The resin was then treated with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (40 mg, 0.08 mmol), HOAt (11 mg, 0.08 mmol), DIC (25  $\mu$ L, 0.16 mmol) and DIEA (50  $\mu$ L) in DMF (1 mL) and heated via microwave to 75 ° C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. Isolated 523 nmoles (0.1%) of **N<sub>3</sub>-2H-K4NMeS-Aak**. **N<sub>3</sub>-2H-K4NMeS Aak** (C<sub>96</sub>H<sub>124</sub>N<sub>25</sub>O<sub>12</sub>) MS calculated (M+H) 1818.9861, MS found 1819.0081 (M+H); t<sub>R</sub> = 32 min.

*Synthesis of Biotin-N<sub>3</sub>-2H-K4NMeS-Aak*. See Figure 41. Rink amide resin (250 mg, 0.15 mmol) was swollen in DMF at room temperature for 10 min and then deprotected with a solution of 20% piperidine in DMF (3 mL, 2 x 20 min). The resin was washed with DMF (3 x 3 mL) and reacted twice with a solution of 1M bromoacetic acid (2 mL) in DMF and DIC (125  $\mu$ L, 0.8 mmol) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was washed with DMF (3 x 3 mL) and reacted with a solution of *N*-[2-(D-Biotinylamino)ethyl]amine<sup>3</sup> (100 mg, 0.3 mmol) in DMF (3 mL) via microwave irradiation (3 x 15 s) using a 700 W microwave set to 10% power. The resin was shaken with this solution at room temperature for 1 h and then washed with DMF (3 x 5 mL). The remaining synthesis was conducted as described above for **N<sub>3</sub>-2H-K4NMeS-Aak**. Isolated 230 nmoles of **Biotin-N<sub>3</sub>-2H-K4NMeS-Aak**; 0.2%). **Biotin-N<sub>3</sub>-2H-K4NMeS-Aak** (C<sub>110</sub>H<sub>146</sub>N<sub>29</sub>O<sub>15</sub>S) MS calculated (M+H) 2145.13, MS found 2145.07 (M+H); t<sub>R</sub> = 34 min.

*Synthesis of N<sub>3</sub>-2H-K4NMeS-TAMRA (Carboxytetramethylrhodamine)*. See Figure 42. 250 mg of resin containing **N<sub>3</sub>-2H-K4NMeS** with the Fmoc 6-amino hexanoate was cleaved from the beads with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (40 mg, 0.08 mmol), HOAt (11 mg, 0.08 mmol), DIC (25  $\mu$ L, 0.16 mmol) and DIEA (50  $\mu$ L) in DMF (1 mL) and heated via microwave to 75 ° C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-

100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. 2.95  $\mu$ moles of product was isolated and the Fmoc was removed by treating with 1 mL of 20% piperidine/DMF for 20 min at room temperature and the solution was concentrated *in vacuo*. The residue was again purified by reverse phase HPLC using the above conditions. A portion of the dimer (315 nmol) was coupled with 5-TAMRA (0.4 mg, 0.95  $\mu$ mol) with DIC (25  $\mu$ L, 0.16 mmol) in DMF (500  $\mu$ L) by heating to 75 ° C for 1.5 h with microwave irradiation. The reaction mixture was concentrated and purified as described above to afford 216 nmoles of **N<sub>3</sub>-2H-K4NMeS-TAMRA**; (0.1%). **N<sub>3</sub>-2H-K4NMeS-TAMRA** (C<sub>118</sub>H<sub>144</sub>N<sub>27</sub>O<sub>15</sub>) MS calculated (M+H) 2179.1335, MS found 2179.0483 (M+H);  $t_R$  = 34 min.

*Synthesis of FAM-2H-K4NMeS-Aak (Fluorescein).* See Figure 43. 500 mg of resin containing **N<sub>3</sub>-2H-K4NMeS** with the Fmoc 6-amino hexanoate was cleaved from the beads with 30% TFA/DCM (5 mL) for 10 min. The solution was concentrated *in vacuo* and azeotroped with toluene three times. The resulting pale yellow oil was treated with a solution of Hoechst carboxylate (40 mg, 0.08 mmol), HOAt (11 mg, 0.08 mmol), DIC (25  $\mu$ L, 0.16 mmol) and DIEA (50  $\mu$ L) in DMF (1 mL) and heated via microwave to 75 ° C for 1.5 h. The solution was then concentrated *in vacuo* and purified using reverse phase HPLC with 20-100% MeOH/H<sub>2</sub>O + 0.1% (v/v) TFA over 1 h. The product was isolated and the Fmoc was removed by treating with 1 mL of 20% piperidine/DMF for 20 min at room temperature and the solution was concentrated *in vacuo*. The residue was again purified by reverse phase HPLC using the above conditions to give 8  $\mu$ moles of dimer. This was then clicked with FAM alkyne (25 mg, 6  $\mu$ mol) and copper catalyst (1 mg, 1.6  $\mu$ mol) by heating in DMF (2 mL) and H<sub>2</sub>O (1 mL) with trimethylamine (250  $\mu$ L) at 120 ° C for 3 h. The reaction mixture was then concentrated and purified as described above to afford 550 nmoles of FAM dimer. This was azeotroped with toluene 3X and was then treated with a solution of propiolic acid succinimidyl ester<sup>4</sup> (1 mg, 6  $\mu$ mol) and DIEA (25  $\mu$ L) in DMF (250  $\mu$ L). The reaction stirred at room temperature for 4 h and was then purified as described above. Isolated 5.9 nmoles of **FAM-2H-K4NMeS-Aak**; (0.002%). **FAM-2H-K4NMeS-Aak** (C<sub>120</sub>H<sub>139</sub>N<sub>26</sub>O<sub>18</sub>) MS calculated (M+H) 2232.0761, MS found 2232.0827 (M+H);  $t_R$  = 35 min.

## METHODS FOR BIOCHEMICAL & CELLULAR ASSAYS

*Evaluation of In Vitro Click by Liquid Chromatography-Mass Spectrometry (LC-MS).* In order to determine if the repeating RNA that causes DM2 serves as a template for

oligomerization via HDCR, alkyne- and azide-containing compounds were incubated with an RNA containing 12 CCUG repeats. rCCUG<sub>12</sub> (50  $\mu$ M final concentration) was folded in 1 $\times$  Folding Buffer (8 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.0, 185 mM NaCl, and 1 mM EDTA) at 60 °C for 5 min. After cooling to room temperature, **N<sub>3</sub>-K** and **K-Ak** (500  $\mu$ M final concentration each) were added and the reaction mixtures were incubated at 37 °C for 48 h. Each sample was analyzed by LC-MS using a Thermo Scientific LTQ-ETD mass spectrometer. A gradient of 0-100% acetonitrile in water plus 0.1% formic acid over 10 min was used for analysis. Total ion counts for each component were normalized to the percent ionization of each component as measured by using a control injection containing an equimolar mixture of **N<sub>3</sub>-K**, **K-Ak**, and **K 1,4 dimer**. Control reactions using r(CUG)<sub>12</sub>, r(AUUCU)<sub>12</sub>, r(CAG)<sub>12</sub>, r(CGG)<sub>12</sub>, an RNA hairpin with a fully paired stem, and brewer's yeast tRNA (Roche) were also evaluated analogously.

*Evaluation of DM2-associated pre-mRNA splicing defects in a DM2 cell culture model.*

In order to determine the effects of clickable compounds on DM2-associated alternative pre-mRNA splicing defects, a previously reported cell culture model was employed<sup>[5]</sup>. Briefly, C2C12 cells were maintained as monolayers in growth medium (1 $\times$  DMEM, 10% FBS, and 1 $\times$  Glutamax (Invitrogen)) at 37 °C and 5% CO<sub>2</sub>. Once cells in 96-well plates reached 60-70% confluency, each sample was transfected with 200 ng of total plasmid using 1  $\mu$ L of Lipofectamine 2000 (Invitrogen) according to the manufacturer's standard protocol. Equal amounts of plasmid expressing a DM2 mini-gene with 300 CCTG repeats<sup>[1]</sup> and a *BIN1* reporter mini-gene were used<sup>[5]</sup>. After 5 h, the transfection cocktail was removed and replaced with differentiation medium (1 $\times$  DMEM, 2% horse serum, and 1 $\times$  Glutamax) containing the compound of interest. After 72 h, total RNA was harvested using a Zymo Quick RNA miniprep kit. Approximately 150 ng of total RNA was subjected to RT-PCR. The RT-PCR primers for the *BIN1* mini-gene were 5'CATTACACATTGGTGTGC (forward) and 5'AAGTGATCCTAGACTAGCCGCC (reverse). RT-PCR products were separated using a denaturing 8% polyacrylamide gel run at 200 V for 90 min in 1 $\times$  TBE buffer. The products were visualized by staining with SYBR Gold (Molecular Probes) and scanned using a Bio-Rad Gel Doc XR+ imaging system.

*Evaluation of in cellulo click on the disruption of nuclear foci by fluorescence in situ hybridization (FISH).* FISH was utilized to determine the effects of clickable compounds on the formation of nuclear foci. C2C12 were sub-cultured into Mat-Tek 96-well glass bottom plates in



growth medium. After the cells reached 60-70% confluency, each well was transfected with 200 ng of total plasmid using 1  $\mu$ L of Lipofectamine 2000 according to the manufacturer's standard protocol. Equal amounts of plasmids expressing the DM2 mini-gene<sup>[1]</sup> and EGFP, used as a positive marker of transfection, were used. After 5 h, the transfection cocktail was removed and replaced with differentiation medium containing the compound of interest. The cells were treated with compound for 72 h followed by FISH as previously described using 1 ng/ $\mu$ L DY547-2'OMe-(CAGG)<sub>5</sub><sup>[6]</sup>. Cells were imaged in 1 $\times$  DPBS using an Olympus FluoView 1000 confocal microscope at 100 $\times$  magnification.

*Identification of in cellulo clicked products by mass spectrometry.* C2C12 cells were grown in T-75 dishes in growth medium. Once the cells reached 60-70% confluency, each dish was transfected with 10  $\mu$ g of a plasmid expressing the DM2 mini-gene using Lipofectamine 2000 according to the manufacturer's standard protocol. After 5 h, the transfection cocktail was removed and replaced with differentiation medium containing the compounds of interest. Cells were treated with a mixture of 12.5  $\mu$ M **N<sub>3</sub>-K** and either 12.5  $\mu$ M **K-Ak**, 12.5  $\mu$ M **N<sub>3</sub>-K-Ak**, or 12.5  $\mu$ M **N<sub>3</sub>-K-Aak** immediately after transfection. **N<sub>3</sub>-K** was added to limit the molecular weight of the oligomeric products in order to enable detection by LC-MS. After 72 h, the cells were lysed by freezing and thawing with 10% water in acetonitrile. The thawed lysate was concentrated and re-suspended in 1 mL of 10% water in acetonitrile. Insoluble cellular debris was pelleted, and the supernatant was used for mass spectral analysis. Approximately 20  $\mu$ L of each sample was analyzed by LC-MS using a Thermo Scientific LTQ-ETD mass spectrometer. A gradient of 0-100% acetonitrile in water plus 0.1% formic acid over 10 min was used for analysis. Total ion counts for each component were normalized to the percent ionization of each component in a control injection containing an equimolar mixture of **N<sub>3</sub>-K**, **K-Ak**, and **K 1,4 dimer**.

*Analysis of Target Pull-Down by qRT-PCR.* C2C12 cells were grown in T-25 flasks as monolayers in growth medium and were transfected at 60-70% confluency. Each dish was transfected with 3.4  $\mu$ g of a plasmid expressing the DM2 mini-gene using Lipofectamine 2000 according to the manufacturer's standard protocol. After 5 h, the transfection cocktail was removed and replaced with differentiation medium containing the compounds of interest. Cells were treated with a mixture of 5  $\mu$ M **N<sub>3</sub>-K-Biotin** and either 5  $\mu$ M **K-Ak**, 5  $\mu$ M **N<sub>3</sub>-K-Ak**, or 5  $\mu$ M **N<sub>3</sub>-K-Aak** immediately after transfection. **N<sub>3</sub>-K-Biotin** was added to limit the molecular

weight of the oligomeric products in order to enable detection by LC-MS. After 72 h, the cells were washed with 1× DPBS and trypsinized. The pelleted cells were lysed by treating with 500 µL of Lysis Buffer (2% Triton X-100, 2% NP40, 1/25 RNaseq (Ambion), and 1 u/µL RQ1 DNase (Promega))<sup>[7]</sup> for 5 min at room temperature and then incubated at 75 °C for 5 min.

Approximately 90% of the cell lysate was used for pull-down using 15 nmoles of streptavidin-agarose beads (Sigma, 15 µg/mL biotin loading). The lysate was incubated with the beads for 1 h at room temperature with shaking at 500 rpm. The solution was removed, and the beads were washed with 500 µL aliquots of 1× PBS. Bound material was released from the beads by heating at 95 °C for 5 min in 30 µL of 95% formamide containing 10 mM EDTA, pH 8.2.

Reverse transcription reactions were carried out using qScript cDNA synthesis kit by adding approximately 10% volume of either cell lysate or eluted material according to the manufacturer's protocol. Then, 30% of each cDNA sample was used for real time PCR (qPCR) analysis for each primer set. qPCR was performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems) using SYBR Green I. The PCR primers for the r(CCUG)<sup>exp</sup>-containing mRNA were 5' GTGAGTTTGGGGACCCTTGA (forward) and 5' CACCCTGAAACTTTGCCCC (reverse). The PCR primers for 18S ribosomal RNA were 5' GTAACCCGTTGAACCCCAT (forward) and 5' CCATCCAATCGGTAGTAGCG (reverse).

*Analysis of Reaction Products Pulled Down by LC-MS.* In order to detect pulled down products, ~4 µL of eluted material was diluted in 5 volumes of water containing 0.1% TFA and analyzed by LC-MS using a Thermo Scientific LTQ-ETD mass spectrometer. These samples were compared to pulled-down samples from compound-treated cells that did not express r(CCUG)<sub>300</sub>. A gradient of 0-100% acetonitrile in water plus 0.1% formic acid over 10 min was used for analysis.

*Evaluation of DM1 Splicing Defects in Patient Derived Fibroblasts.* Bioactivity of small molecule dimers was assessed by using DM1 patient derived fibroblasts containing 500 CTG repeats (GM03987) and healthy fibroblasts (GM07492). Cells were grown as monolayers in 12 well plates in growth medium (1X EMEM (Lonza), 10% FBS, 1X glutagro (Corning), 1X MEM non-essential amino acids (Corning) and 1X antibiotic/antimycotic (Corning)). Once cells were ~80% confluent, they were treated with growth medium containing the compound of interest (10, 1, 0.1 and 0.01 µM **2H-K4NMeS**; 100, 10 and 1 nM **2H-K4NMeS-CA-Biotin**; 1 µM **N<sub>3</sub>-2H-**

**K4NMeS**; 500 nM **N<sub>3</sub>-2H-K4NMeS** + 500 nM **2H-K4NMeS-Aminohexanoate**; 1000, 10 and 0.1 nM **N<sub>3</sub>-2H-K4NMeS-Aak**; 250, 100, and 50 nM **2H-K4NMeS-Bleomycin A5**). After 48 h the cells were lysed and the total RNA was harvested using a Zymo Quick RNA miniprep kit. An on-column DNA digestion was completed per the manufacturer's recommended protocol. Approximately 150 ng of total RNA was reverse transcribed at 50 °C using 100 units of SuperScript III reverse transcriptase (Life Technologies). Then 20% of the RT reaction was subjected to PCR using GoTaq DNA polymerase (Promega). RT-PCR products were observed after 25 cycles of 95 °C for 30 s, 58 °C for 30 s, 72 °C for 1 min and a final extension at 72 °C for 1 min. The products were separated on a 2 % agarose gel ran at 100 V for 1 h in 1X TAE buffer. The products were visualized by staining with ethidium bromide and scanned using a Bio-Rad Gel Doc XR+ imaging system. The RT-PCR primers for the *MBNL1* were 5'GCTGCCCAATACCAGGTCAAC (forward) and 5'TGGTGGGAGAAATGCTGTATG (reverse).

*Evaluation of Splicing Defects Using a DM1 Mouse Model.* All experimental procedures, mouse handling, and husbandry were completed in accordance with the Association for Assessment and Accreditation of Laboratory Animal Care. A mouse model for DM1<sup>[8]</sup>, HSA<sup>LR</sup> in line 20b, was used. HSA<sup>LR</sup> mice express human skeletal actin RNA with 250 CUG repeats in the 3' UTR. Age- and gender-matched HSA<sup>LR</sup> mice were injected intraperitoneally with either 100 mg/kg **2H-K4NMe** in water or 13.3 mg/kg **2H-K4NMeS** for treatment and 0.9% NaCl for control once per day for 7 days. Mice were sacrificed one day after the last injection, and the vastus muscle was obtained. RNA was extracted from the vastus tissue, and cDNA was synthesized as previously described.

*In Vitro Reaction and Pull Down of Small Molecule Dimers with RNA by using compounds to target r(CUG)<sup>exp</sup>.* Growth medium ((1X EMEM (Lonza), 10% FBS, 1X glutagro (Corning), 1X MEM non-essential amino acids (Corning) and 1X antibiotic/antimycotic (Corning) 2mL) was heat inactivated at 95 °C for 15 min and then slowly cooled to room temperature. Then ~10,000 counts of  $\gamma$  <sup>32</sup>P labeled RNA (rCUG<sub>109</sub>, GC-based paired RNA and tRNA) was added and folded at 95 °C for 1 min. Then 2-fold dilutions of compound (800, 400, 200 and 100 nM final concentrations) were prepared in 50  $\mu$ L of RNA solution and incubated at 37 °C overnight. Then a 400  $\mu$ L slurry of streptavidin-agarose beads (Sigma) were washed with 1X PBS and then resuspended in 2 mL of 1X PBS. Then 30  $\mu$ L of this slurry were added to each

sample and incubated at room temperature for 1 h. Then the samples were centrifuged and the supernatant was transferred to a separate tube. The beads were then washed with 1X PBST and centrifuged. The supernatant was transferred to the tube containing the unbound RNA. The total radioactive counts of bound and unbound RNA were measured by scintillation counting.

*Target Identification and Pull Down by Chem-CLIP by using the DM1 system.* Target identification of small molecule dimers was assessed using DM1 patient derived fibroblasts containing 500 CTG repeats (GM03987) and healthy fibroblasts (GM07492). Cells were grown as monolayers in 100 mm<sup>2</sup> in growth medium (1X EMEM (Lonza), 10% FBS, 1X glutagro (Corning), 1X MEM non-essential amino acids (Corning) and 1X antibiotic/antimycotic (Corning)). Once cells were ~80% confluent, they were treated with growth medium containing the compound of interest (100 nM **2H-K4NMeS-CA-Biotin**). After 48 h the cells were lysed and the total RNA was harvested using Trizol reagent (Life Technologies). Approximately 10 µg of total RNA was incubated with streptavidin-agarose beads (100 µL, Sigma) for 1 h at room temperature. Then the beads were washed with 1X PBS and the bound RNA was eluted by adding 100 µL of 95% formamide, 10 mM EDTA pH 8.2 for 10 min at 60 °C. The bound RNA was cleaned up using a Zymo Quick RNA miniprep kit. Approximately 100 ng of RNA was used for RT qScript cDNA synthesis kit (Quanta BioSciences). 10% of the RT reaction was used for real time PCR (qPCR) with SYBR green master mix (Life Technologies) performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems). Monitored rCUG-containing mRNAs *DMPK* (500 repeats), *SUPT20HL1* (17 repeats), *CASK* (16 repeats), *LRP8* (11 repeats), *MAP3K4* (11 repeats), *SCUBE* (7 repeats) and *SORCS2* (7 repeats). Quantified by  $\Delta\Delta C_t$  relative to *GAPDH*

*Target Identification and Pull Down by Chem-CLIP-Map.* Target identification of the binding sites of small molecule dimers was assessed using DM1 patient derived fibroblasts containing 500 CTG repeats (GM03987). Cells were grown as monolayers in 100 mm<sup>2</sup> in growth medium (1X EMEM (Lonza), 10% FBS, 1X glutagro (Corning), 1X MEM non-essential amino acids (Corning) and 1X antibiotic/antimycotic (Corning)). Once cells were ~80% confluent, they were treated with growth medium containing the compound of interest (100 nM **2H-K4NMeS-CA-Biotin**). After 48 h the cells were lysed and the total RNA was harvested using Trizol reagent (Life Technologies). Approximately 6 µg of total RNA was folded with 8 µM antisense oligonucleotide in 1X RNase H buffer by heating to 95 °C for 1 min and then

cooling on ice. Next 5 units of RNase H (Life Technologies) were added and the reaction was incubated at 37 ° C for 1.5 h followed by heat inactivation at 65 ° C for 20 min. Then the cut RNA was treated with DNase at 37 ° C for 30 min and then stop buffer was added and incubated at 65 ° C for 10 min. The cut RNA solution was incubated with streptavidin-agarose beads (100  $\mu$ L, Sigma) for 1 h at room temperature. Then the beads were washed with 1X PBS and the bound RNA was eluted by adding 100  $\mu$ L of 95% formamide, 10 mM EDTA pH 8.2 for 10 min at 60 ° C. The bound RNA was cleaned up using a Zymo Quick RNA miniprep kit.

Approximately 150 ng of RNA was used for RT qScript cDNA synthesis kit (Quanta BioSciences). 40% of the RT reaction was used for real time PCR (qPCR) with SYBR green master mix (Life Technologies) performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems). Quantified by  $\Delta\Delta C_t$  relative to *GAPDH*.

*Evaluation of In Vitro Click by Liquid Chromatography-Mass Spectrometry (LC-MS).* In order to determine if the repeating RNA that causes DM1 serves as a template for oligomerization via HDCR, **2H-K4NMeS Azide** and each **2H-K4NMeS Activated Alkyne** derivative were incubated with an RNA containing 12 CUG repeats. rCUG<sub>12</sub> (100  $\mu$ M final concentration) was folded in 1 $\times$  Folding Buffer (8 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.0, 185 mM NaCl, and 1 mM EDTA) at 60 °C for 5 min. After cooling to room temperature, azide and alkyne dimers (25  $\mu$ M final concentration each) were added and the reaction mixtures were incubated at 37 °C for 24 h. Each sample was analyzed by LC-MS using a Thermo Scientific LTQ-ETD mass spectrometer. A gradient of 0-100% acetonitrile in water plus 0.1% formic acid over 10 min was used for analysis. The alkyne which gave the most reaction with r(CUG)<sub>12</sub>, **2H-K4NMeS-Aminohexanoate Aak**, was next evaluated for selectivity by analyzing the amount of dimer formed in the presence of other RNA targets. Other RNAs evaluated were r(CCUG)<sub>12</sub>, r(CAG)<sub>12</sub>, r(CGG)<sub>12</sub>, and brewer's yeast tRNA (Roche) each at 100  $\mu$ M final concentration.

*Identification of In Cellulo Clicked Products by Mass Spectrometry using DM1 fibroblasts.* In order to determine if the repeating RNA that causes DM1 serves as a template for oligomerization within cells DM1 patient derived fibroblasts containing 500 CTG repeats (GM03987) and healthy fibroblasts (GM07492) were grown to 80% confluence in growth medium (1X EMEM (Lonza), 10% FBS, 1X glutagro (Corning), 1X MEM non-essential amino acids (Corning) and 1X antibiotic/antimycotic (Corning)) in T25 dishes and treated with equimolar amounts of **N<sub>3</sub>-2H-K4NMeS Biotin** and **2H-K4NMeS-Ahx Alkyne** (500 nM each)

for 2 days. The cells were washed with 1X DPBS and then trypsinized. Then quenched with growth medium and spun down. Cells were washed 2X with 1X DPBS and then lysed by adding 0.25 mL of lysis buffer (2 mL total = 2% Triton X 100, 2% NP40, 80 uL RNasecure (1/25) and 50 uL DNase) for 5 min at room temperature and then heated to 75 ° C for 5 min. 200 µL of lysate was incubated with streptavidin-agarose beads (100 µL, Sigma) for 1 h at room temperature. Then the beads were washed with 1X PBST and the bound RNA was eluted by adding 20 µL of 95% formamide, 10 mM EDTA pH 8.2 for 5 min at 60 ° C. Approximately 4 µL of each sample was diluted in 20 µL of water plus 0.1% formic acid and analyzed by LC-MS using a Thermo Scientific LTQ-ETD mass spectrometer. A gradient of 0-100% acetonitrile in water plus 0.1% formic acid over 10 min was used for analysis.

*Evaluation of Nuclear Foci Using Fluorescence In Situ Hybridization (FISH).* FISH was utilized to determine the effects of small molecule dimers on the formation and disruption of nuclear foci. DM1 patient derived fibroblasts containing 500 CTG repeats (GM03987) were grown to ~80 % confluence in a Mat-Tek 96-well glass bottom plate in growth medium. Cells were treated with the compound of interest for 48 h in growth medium followed by FISH as previously described<sup>6</sup> using 1 ng/µL DY547 -2'OMe-(CAG)<sub>6</sub>. Immunostaining of MBNL1 was completed as previously described<sup>7</sup> using the MB1a antibody (diluted 1:4), which was generously supplied by Prof. Glenn E. Morris (Wolfson Centre for Inherited Neuromuscular Disease<sup>[9]</sup>). This was fluorescently labeled using a 1:200 dilution of goat anti-mouse IgG DyLight 488 conjugate. Untreated controls were stained using a 1 µg/µL solution of DAPI in 1X DPBS. Cells were imaged in 1× DPBS using an Olympus FluoView 1000 confocal microscope at 100× magnification.

*Evaluation of Translational Defects Using a Luciferase Reporter Assay.* C2C12 cell lines expressing 800 or 0 CTG repeats in the 3' UTR of luciferase were grown as monolayers in 96-well plates in growth medium (1X DMEM, 10% FBS, 1X glutagro, (Corning) and 1X antibiotic/antimycotic (Corning))<sup>[10]</sup>. Once the cells were 70% confluent, the click functionalized dimers were added in 100 µL of growth medium (1000 nM **N<sub>3</sub>-2H-K4NMeS** and 1000, 10 and 0.1 nM **N<sub>3</sub>-2H-K4NMeS-Aak**). Cells were treated with compound for 48 h and then the cell count was normalized using WST-1 reagent (Roche). Then cells were washed with 1X DPBS and lysed by treating with 50 µL of PPBT lysis buffer at room temperature for 10 minutes. Then 50 µL of luciferase substrate was added and luminescence was measured.

*Evaluation of In Vitro Click as Measured by FRET.* In order to determine if the repeating RNA that causes DM1 serves as a template for oligomerization via HDCR, **FAM-2H-K4NMeS Aak** and **N<sub>3</sub>-2H-K4NMeS Aak** were incubated with an RNA containing 12 CUG repeats. rCUG<sub>12</sub> (80  $\mu$ M final concentration) was folded in 1 $\times$  Folding Buffer (20 mM HEPES, pH 7.5, 100 M KCl, and 10 mM NaCl). After cooling to room temperature, **FAM-2H-K4NMeS Aak** (60 nM final concentration) and **N<sub>3</sub>-2H-K4NMeS Aak** (40 nM final concentration) were added and the reaction mixtures were incubated at 37 °C for a total of 48 h. FRET was measured by exciting at 485 nm and measuring emission at 590 nm. Enhancement in FRET was quantified by comparing to controls with **FAM-2H-K4NMeS Aak** (60 nM final concentration) and **N<sub>3</sub>-2H-K4NMeS Aak** (40 nM final concentration) in the absence of RNA. Also FRET was measured using a base-paired control RNA (r(GC)<sub>20</sub>) as a negative control.

Documents cited for Examples section

- [1] J. L. Childs-Disney, I. Yildirim, H. Park, J. R. Lohman, L. Guan, T. Tran, P. Sarkar, G. C. Schatz, M. D. Disney, *ACS Chem. Biol.* 2014, 9, 538-550.
- [2] J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio, R. M. Levy, *J. Comput. Chem.* 2005, 26, 1752-1780.
- [3] K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossv  ry, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, D. E. Shaw, *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*, Tampa, FL 2006.
- [4] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* 2004, 6, 2853-2855.
- [5] C. Fugier, A. F. Klein, C. Hammer, S. Vassilopoulos, Y. Ivarsson, A. Toussaint, V. Tosch, A. Vignaud, A. Ferry, N. Messaddeq, Y. Kokunai, R. Tsuburaya, P. de la Grange, D. Dembele, V. Francois, G. Precigout, C. Boulade-Ladame, M. C. Hummel, A. L. de Munain, N. Sergeant, A. Laquerriere, C. Thibault, F. Deryckere, D. Auboeuf, L. Garcia, P. Zimmermann, B. Udd, B. Schoser, M. P. Takahashi, I. Nishino, G. Bassez, J. Laporte, D. Furling, N. Charlet-Berguerand, *Nat. Med.* 2011, 17, 720-725.
- [6] M. B. Warf, M. Nakamori, C. M. Matthys, C. A. Thornton, J. A. Berglund, *Proc. Natl. Acad. Sci. U. S. A.* 2009, 106, 18551-18556.
- [7] Y. K. Ho, W. T. Xu, H. P. Too, *PLoS ONE* 2013, 8, e72463.

- [8] Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000) Myotonicdystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 289, 1769–1773.
- [9] Holt, I.; Mittal, S.; Furling, D.; Butler-Browne, G. S.; David Brook, J.; Morris, G. E. *Genes to Cells* 2007, 12, 1035.
- [10] Childs-Disney, J. L.; Hoskins, J.; Rzuczek, S. G.; Thornton, C. A.; Disney, M. D. *ACS Chemical Biology* 2012.

While the invention has been described and exemplified in sufficient detail for those skilled in this art to make and use it, various alternatives, modifications, and improvements will be apparent to those skilled in the art without departing from the spirit and scope of the claims.

All patents and publications referred to herein are incorporated by reference herein to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference in its entirety.

The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.



## CLAIMS

What is claimed is:

1. A modulator of the function of a toxic disease-inducing RNA, wherein the RNA comprises expanded repeat sequences, the modulator comprising an RNA-binding oligomer that is formed *in situ* within a living cell from a condensation reaction of a cell-permeable RNA expanded repeat sequence binding module, wherein the cell-permeable module comprises both an alkyne and an azide moiety, or from a condensation of a pair of cell-permeable RNA expanded repeat sequence binding modules, wherein a first module comprises an alkyne moiety and a second module comprises an azide moiety, the condensation reaction taking place between the RNA expanded repeat bound modules via an a 1,3 Huisgen dipolar cycloaddition reaction between the alkyne and azide moieties.
2. The modulator of claim 1, wherein the toxic RNA comprises adjacent internal loops formed by RNA expanded repeat sequences r(CCUG)<sup>exp</sup>,
3. The modulator of claim 2, wherein each cell-permeable module comprises a modified kanamycin moiety.
4. The modulator of claim 3, wherein the cell-permeable module is **N<sub>3</sub>-K-Ak**, or **N<sub>3</sub>-K-AaK**, or a pair of cell-permeable modules used together are **N<sub>3</sub>-K** and **K-Ak**.
5. The modulator of any one of claims 2-4 for treatment of myotonic dystrophy type 2 (DM2).
6. A method of forming the modulator of claim 2 bound to the toxic RNA, comprising exposing the living cell comprising the toxic RNA to a cell-permeable RNA expanded repeat sequence binding module, wherein the cell-permeable module comprises both an alkyne and an azide moiety, or to a pair of cell-permeable RNA expanded repeat sequence binding modules,

wherein a first module comprises an alkyne moiety and a second module comprises an azide moiety.

10. The modulator of claim 1, wherein the toxic RNA comprises adjacent internal loops formed by RNA expanded repeat r(CUG)<sup>exp</sup>.
11. The modulator of claim 10, wherein each cell-permeable module comprises a bis-heteroaryl moiety, preferably a bis-benzimidazole moiety
12. The modulator of claim 11, wherein the cell permeable module is **N<sub>3</sub>-2H-K4NMeS-Aak**.
13. The modulator of any one of claims 10-12 for treatment of myotonic dystrophy type 1 (DM1).
14. A method of forming the modulator of claim 11 bound to the toxic RNA, comprising exposing the living cell comprising the toxic RNA to a cell-permeable RNA expanded repeat sequence binding module, wherein the cell-permeable module comprises both an alkyne and an azide moiety, or to a pair of cell-permeable RNA expanded repeat sequence binding modules, wherein a first module comprises an alkyne moiety and a second module comprises an azide moiety.
15. A method of cleaving a toxic disease-inducing RNA, wherein the RNA comprises expanded repeat sequences, the modulator comprising a modulator of any one of claims 1-4 or 10-12 further comprising a nucleic acid cleaving moiety.
16. The method of claim 15 wherein the nucleic acid cleaving moiety is a derivative of bleomycin.
17. A method of treatment of a disease in a patient wherein the disease is induced by the presence of a toxic RNA having expanded repeat sequences, comprising forming within the patient's cells a modulator of claim 1.

18. The method of claim 17, wherein the disease is myotonic dystrophy type 2 (DM2), the toxic RNA comprises r(CCUG)<sup>exp</sup> expanded repeat sequences, and the modulator is formed *in situ* from one cell-permeable module comprising a kanamycin derivative bearing an alkyne and an azide group, or a pair of cell-permeable modules, each comprising a kanamycin derivative, a first module bearing an alkyne group and a second module bearing an azide group.
19. The method of claim 17, wherein the disease is myotonic dystrophy type 1 (DM1), the toxic RNA comprises or r(CUG)<sup>exp</sup> expanded repeat sequences, and the modulator is formed *in situ* from one cell-permeable module bearing an alkyne and an azide group, or a pair of cell-permeable modules each comprising a bis-benzimidazole derivative, a first module bearing an alkyne group and a second module bearing an azide group.
20. The method of claim 18 or 19, wherein the modulator further comprises a nucleic acid cleaving moiety.

## FIGURES

Figure 1

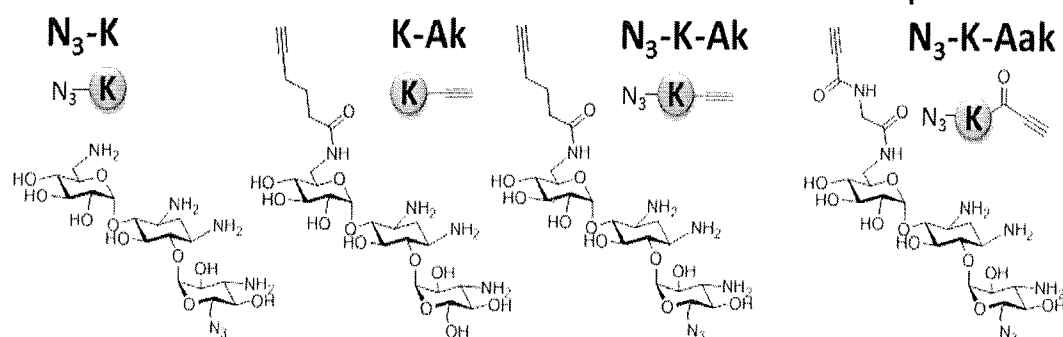
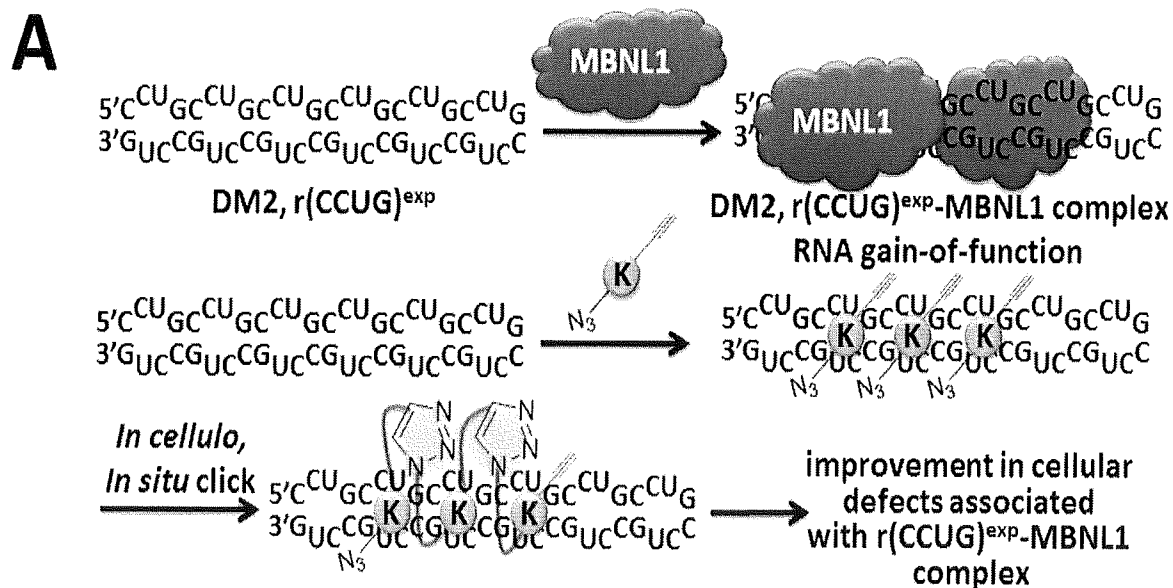
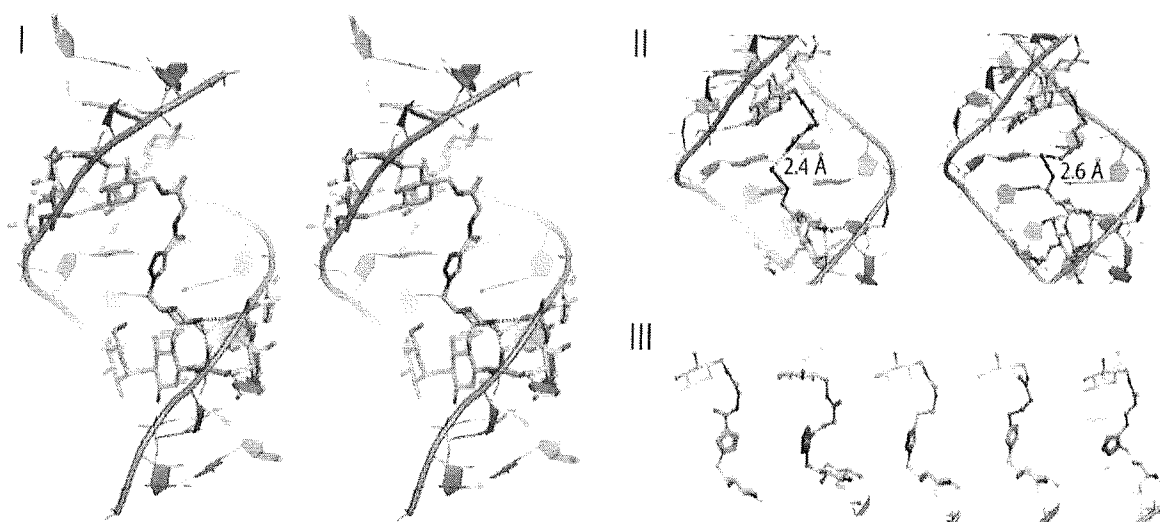
**B**

Figure 2

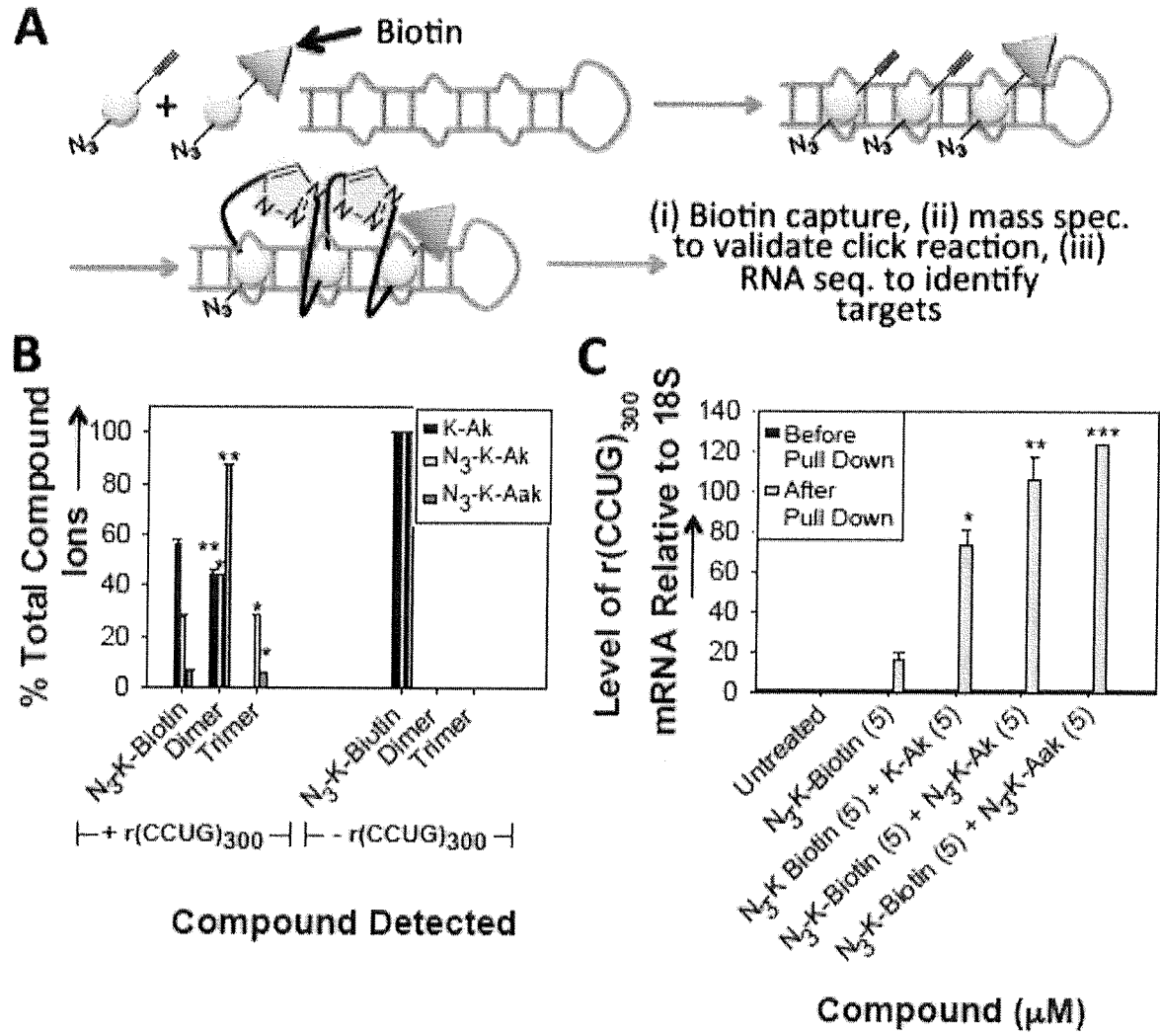


Figure 3

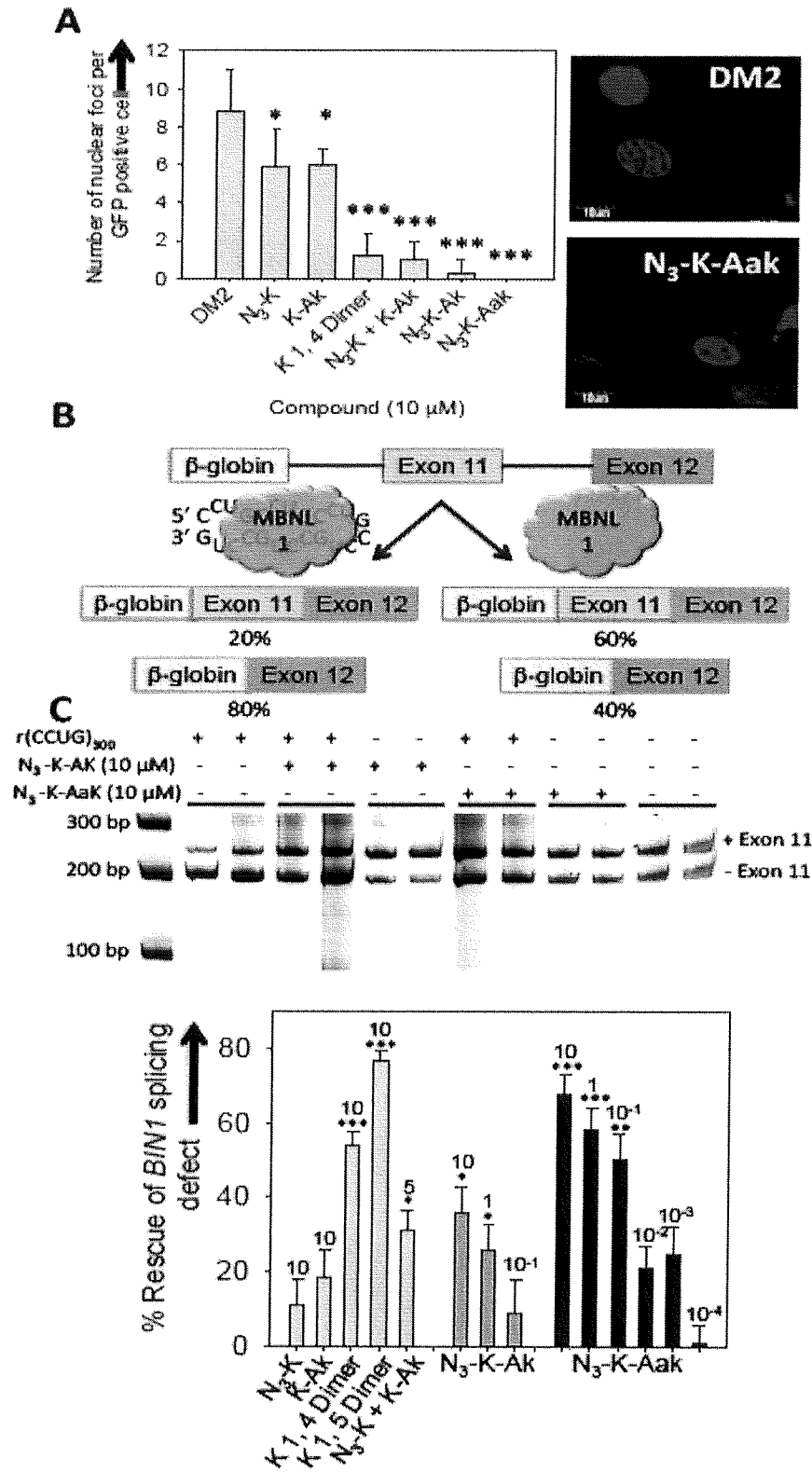


Figure 4

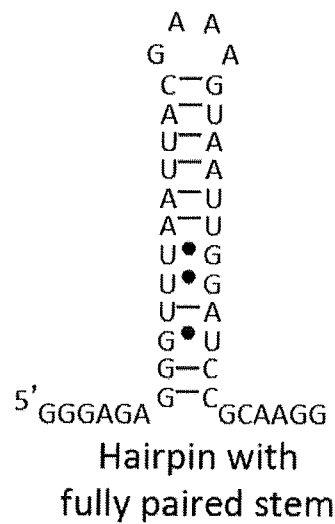
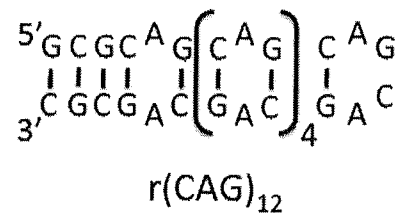
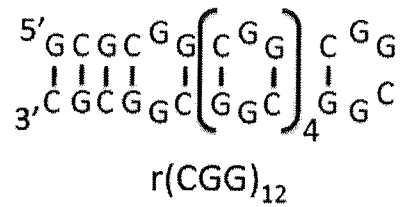
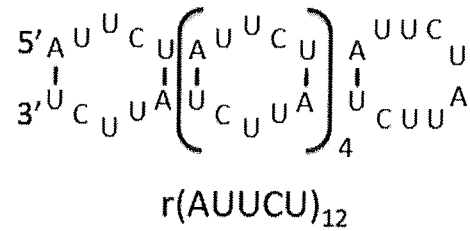
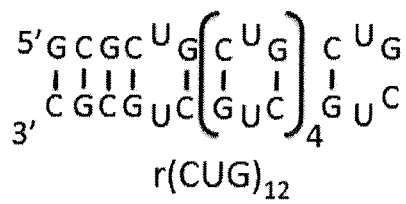
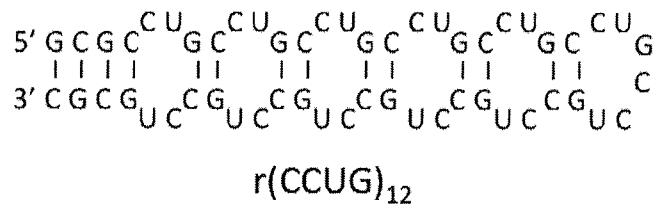


Figure 5

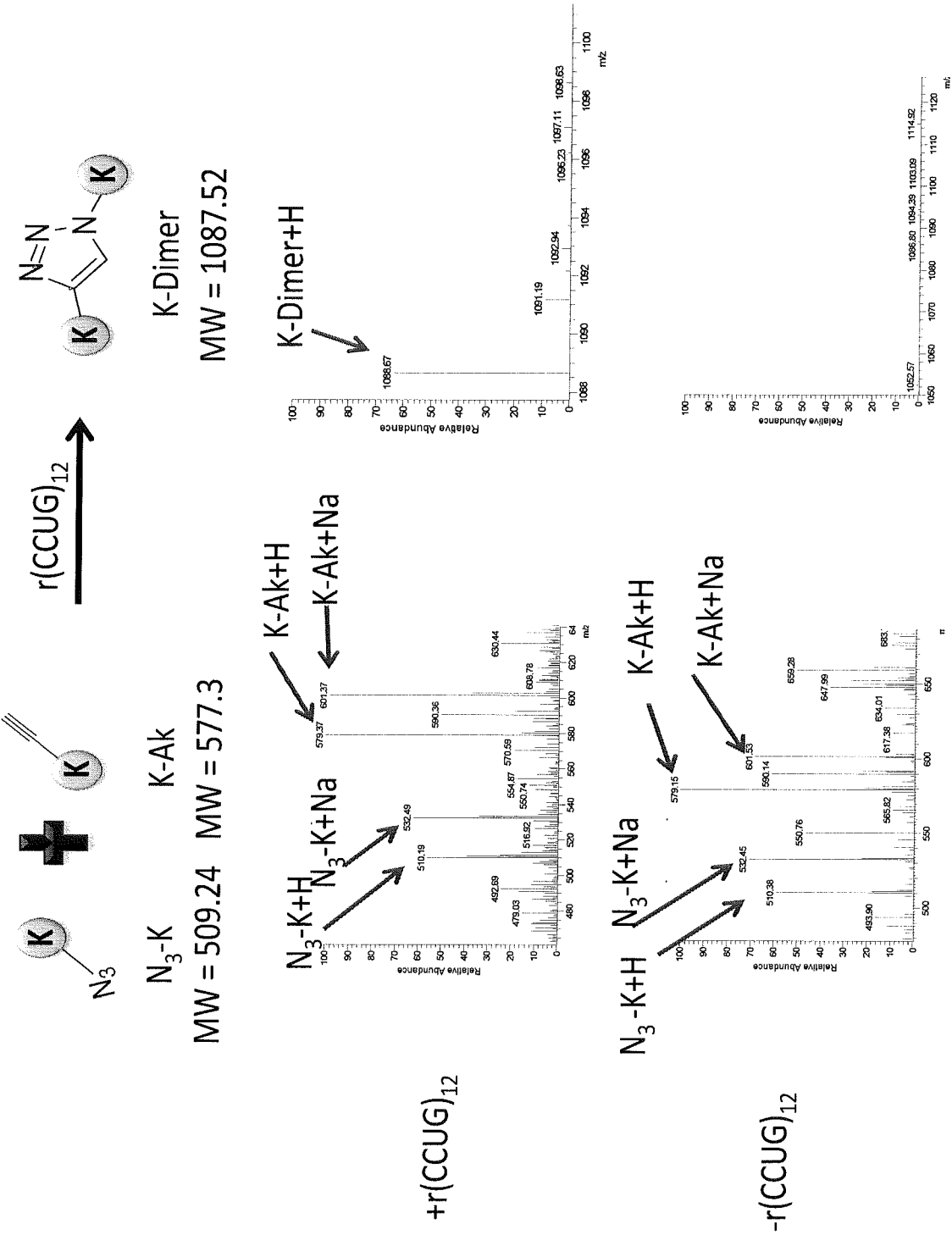




Figure 6

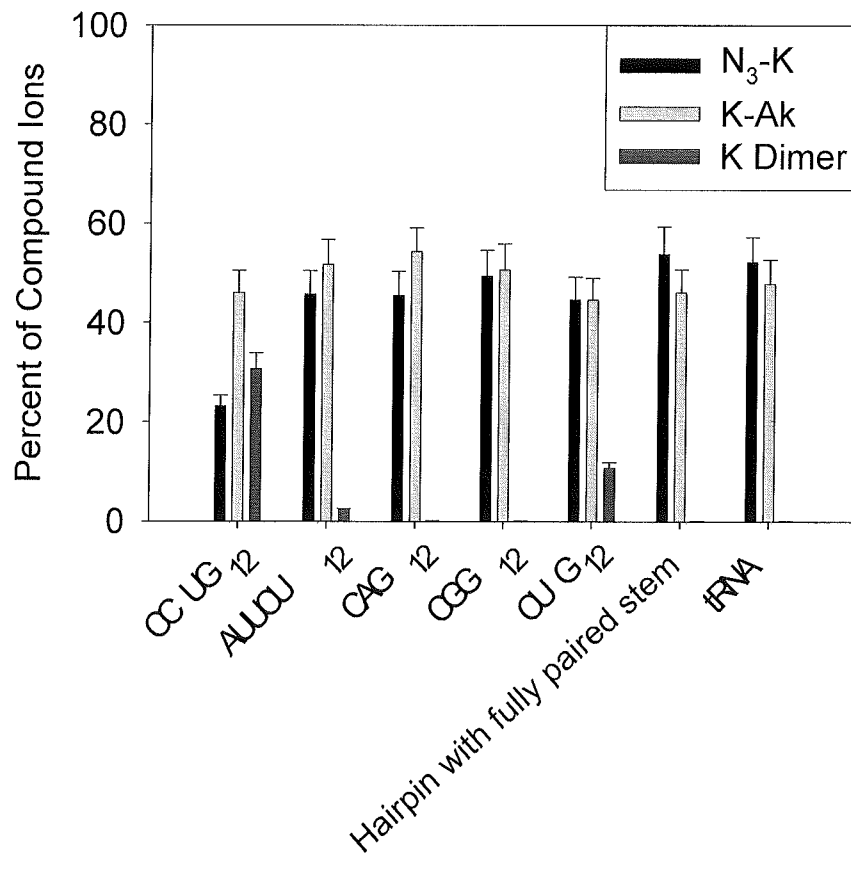


Figure 7

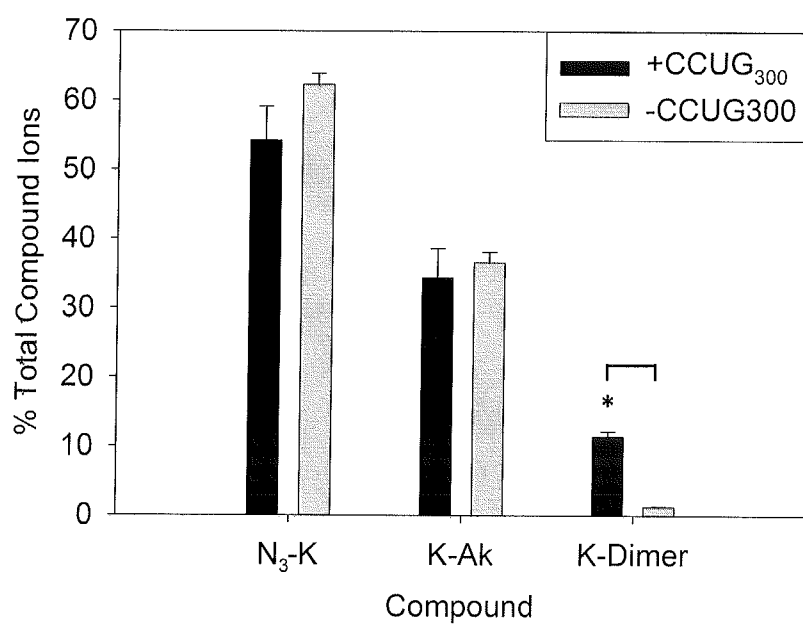
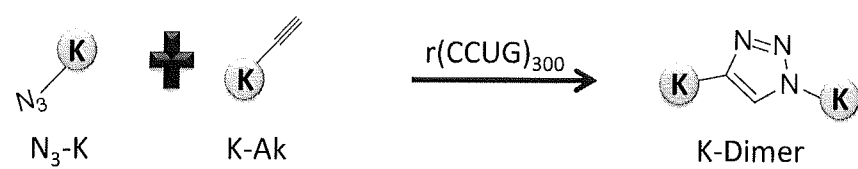


Figure 8

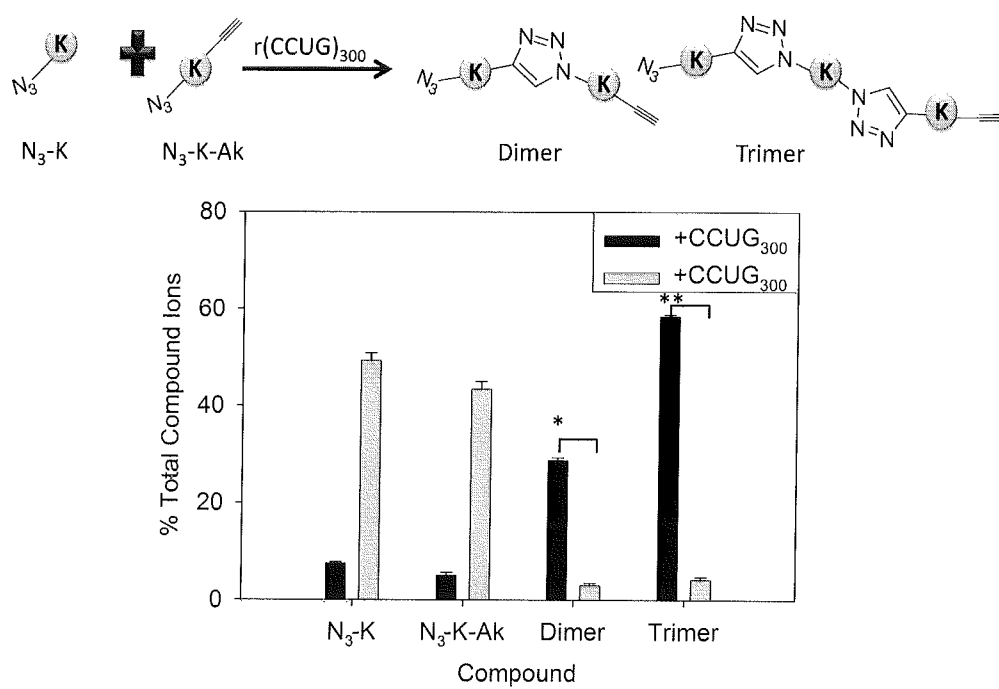


Figure 9

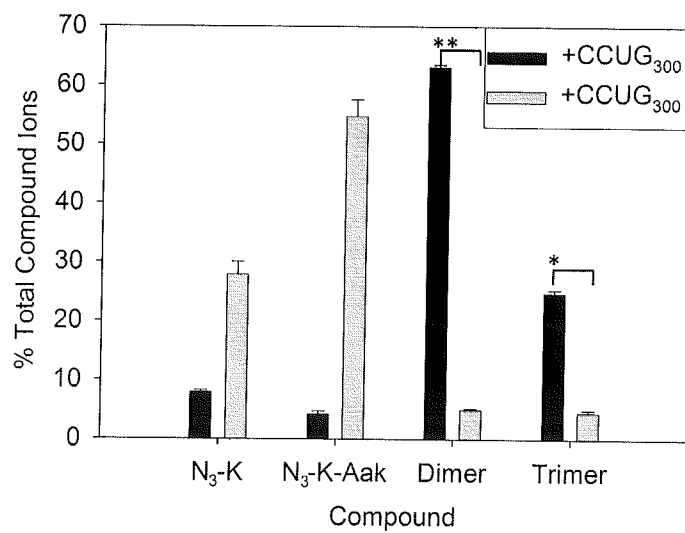
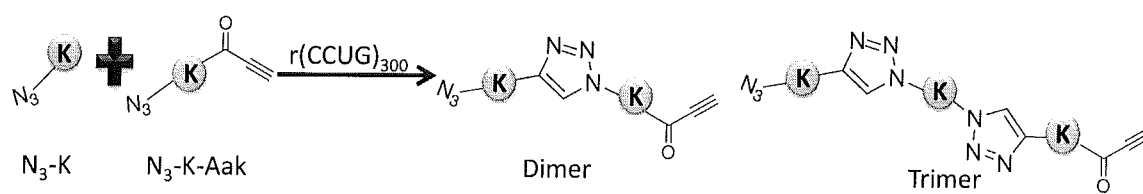


Figure 10

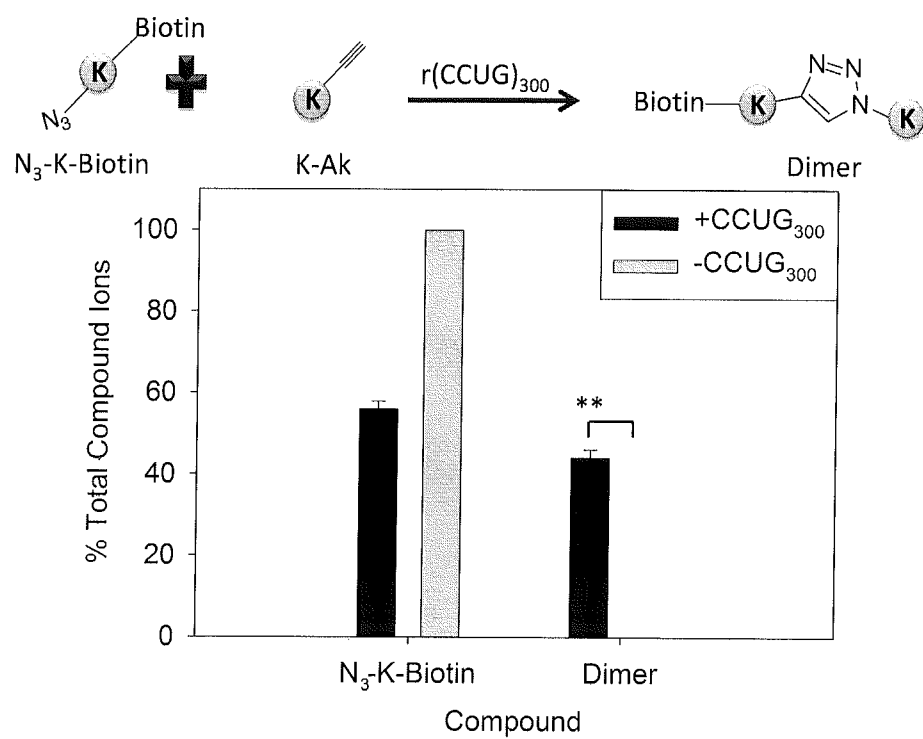


Figure 11

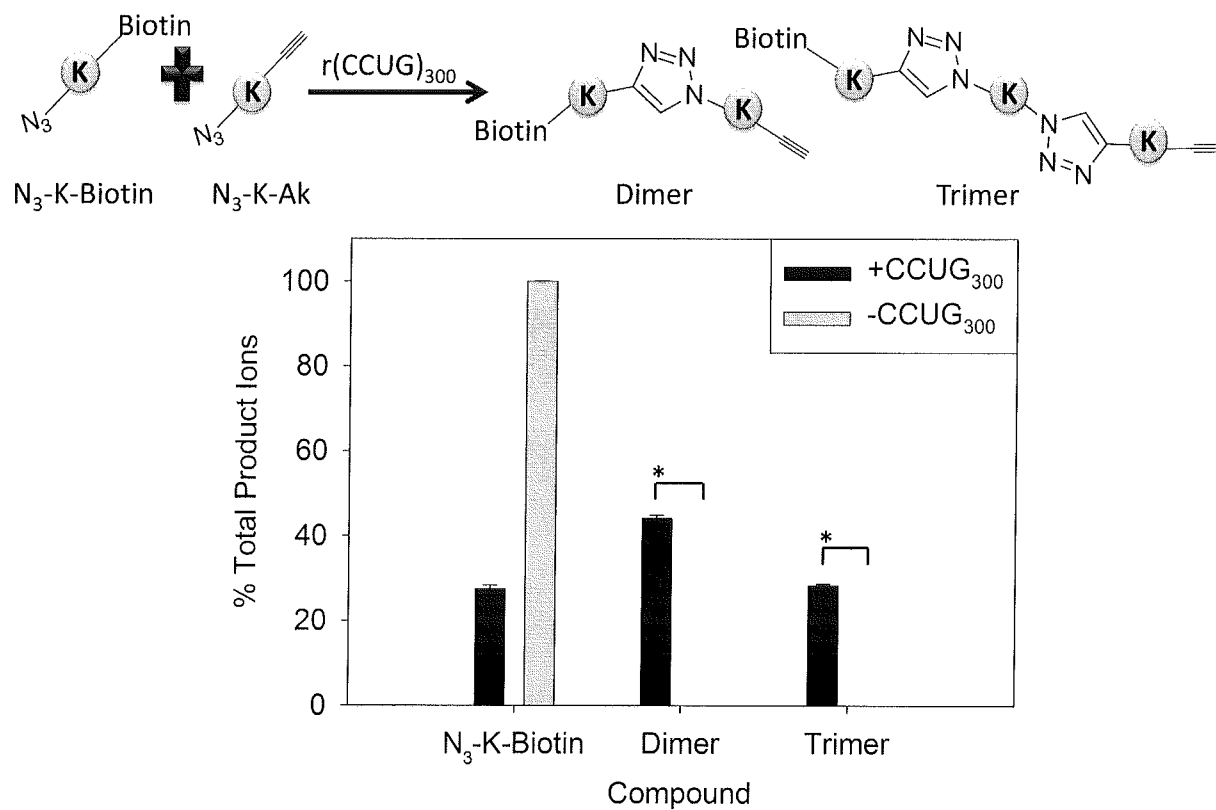


Figure 12

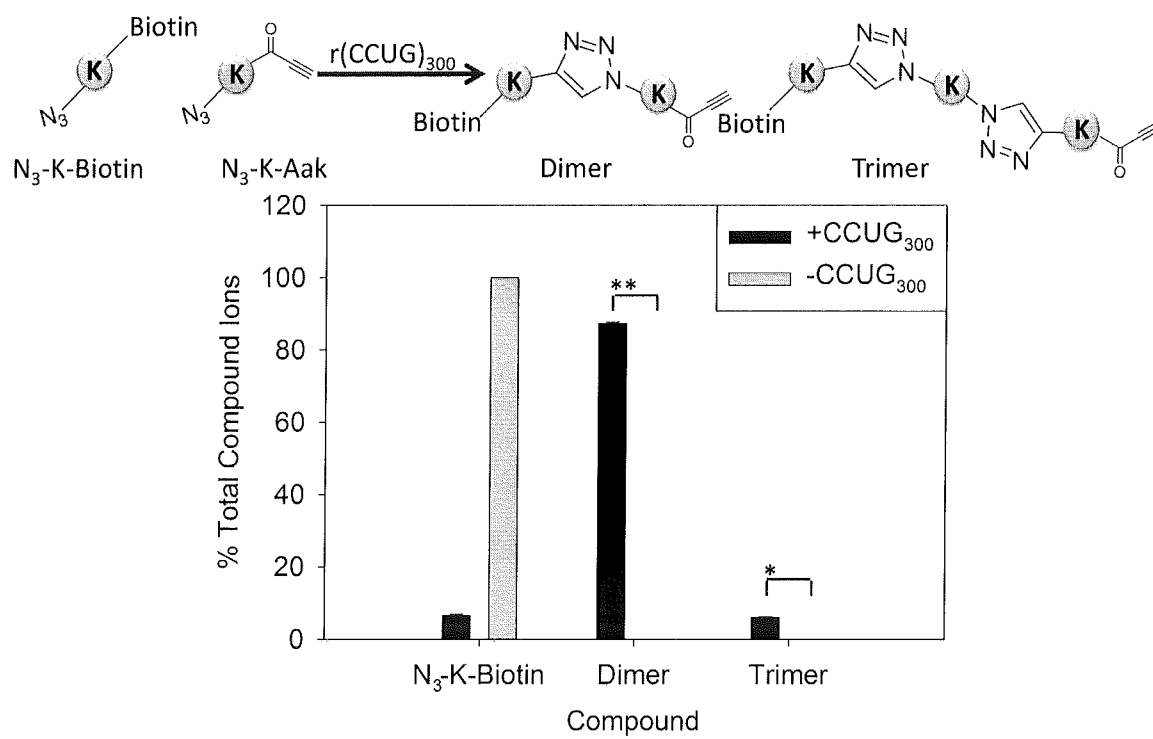


Figure 13

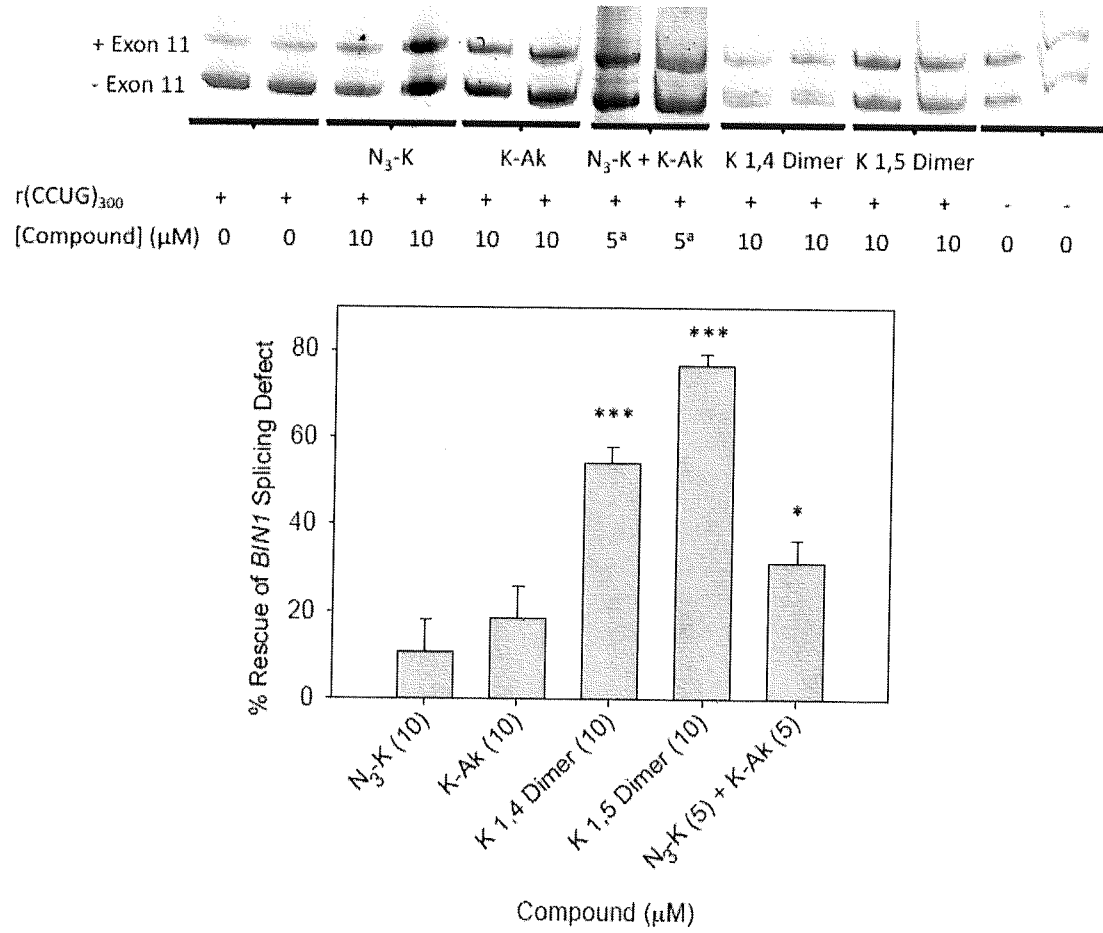
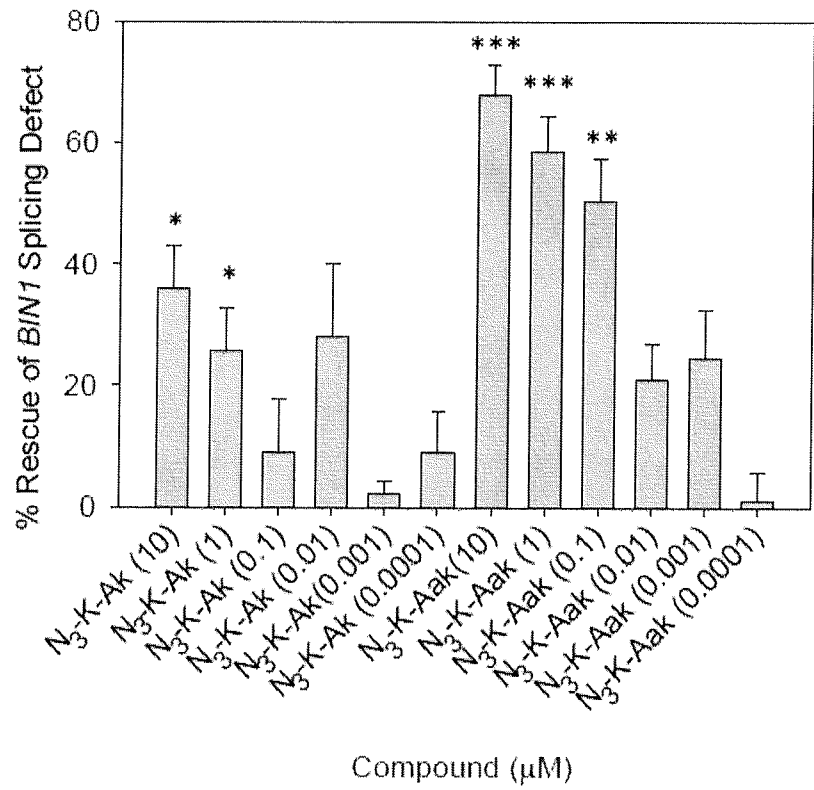
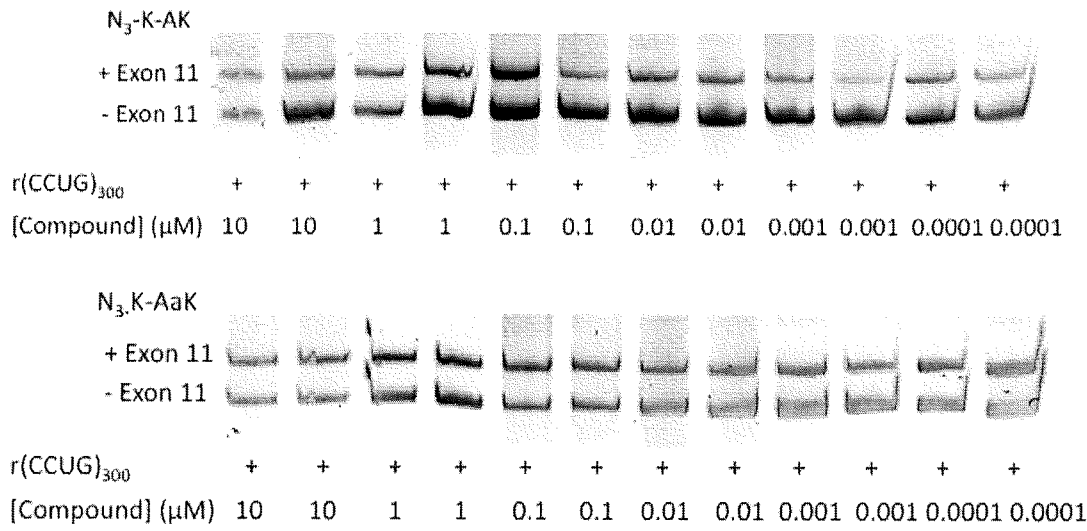




Figure 14





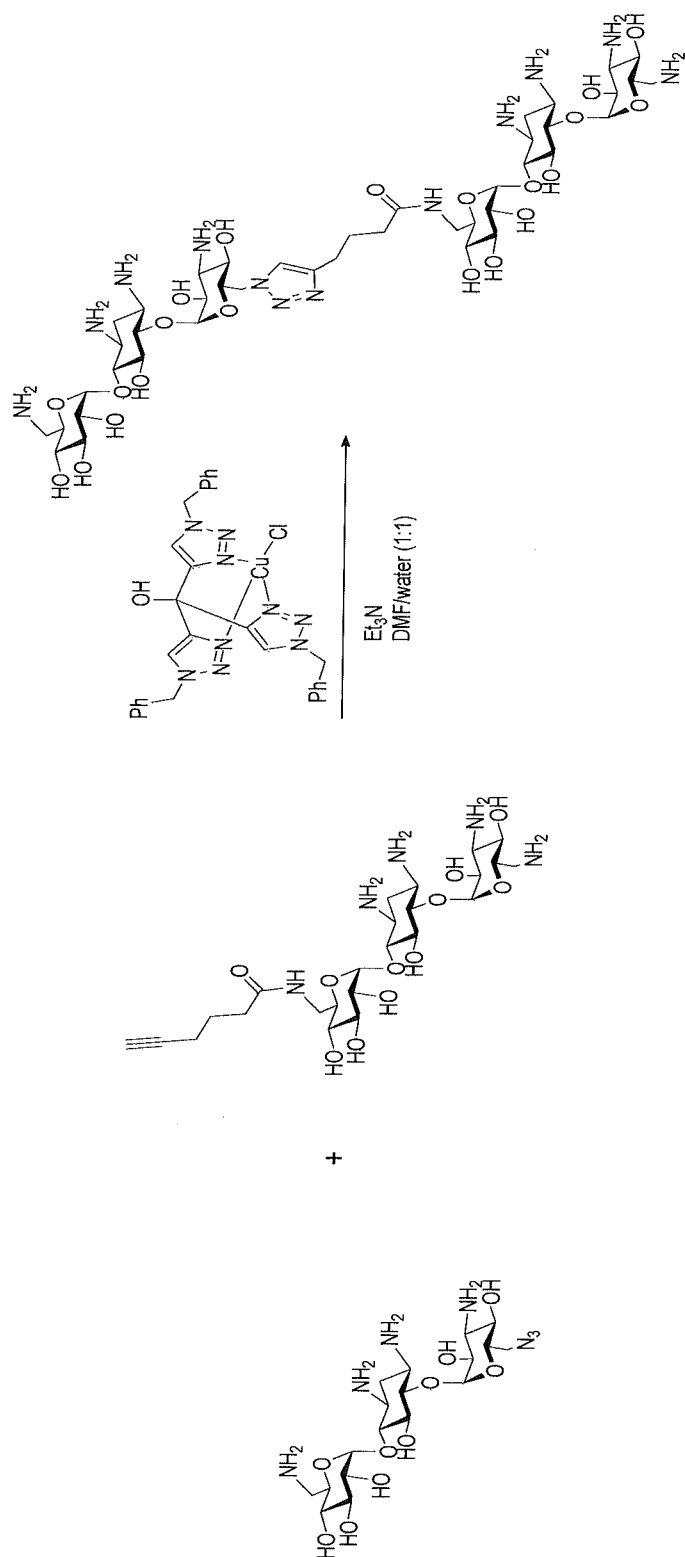


Figure 16

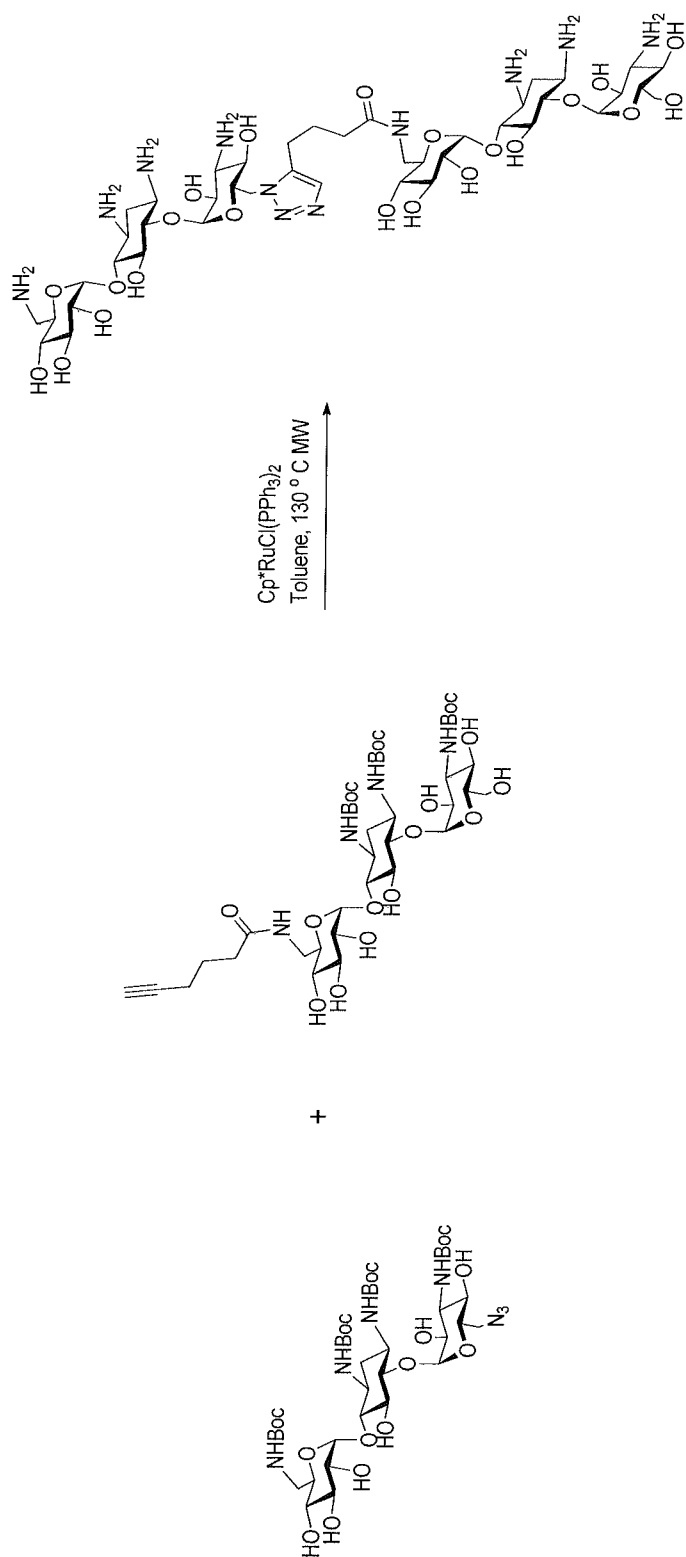


Figure 17

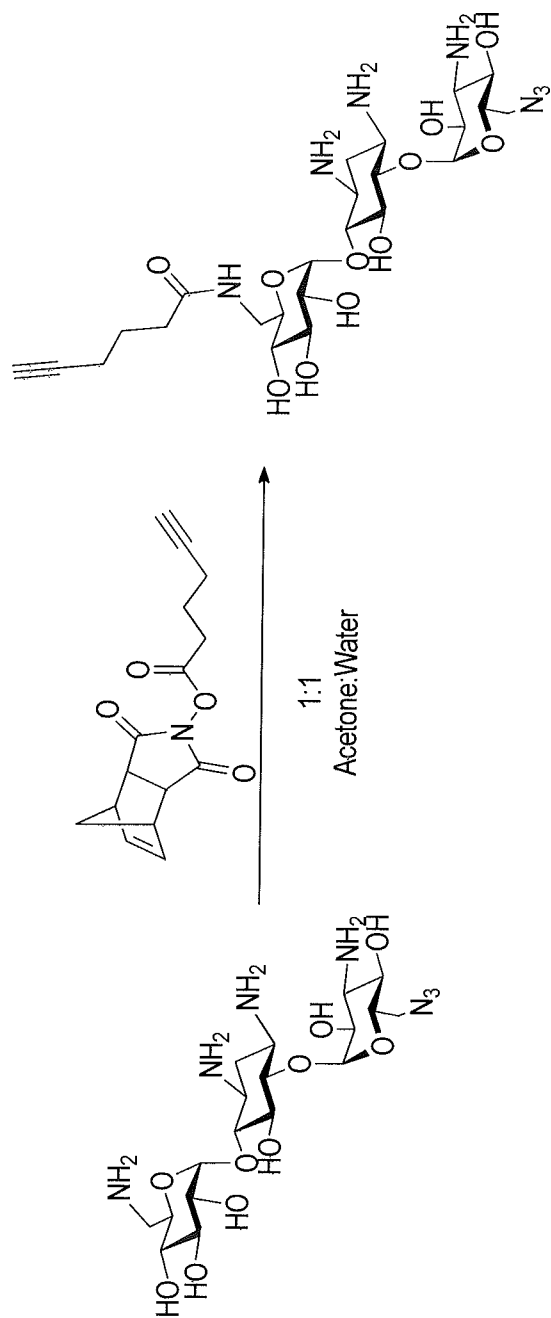


Figure 18

Figure 19

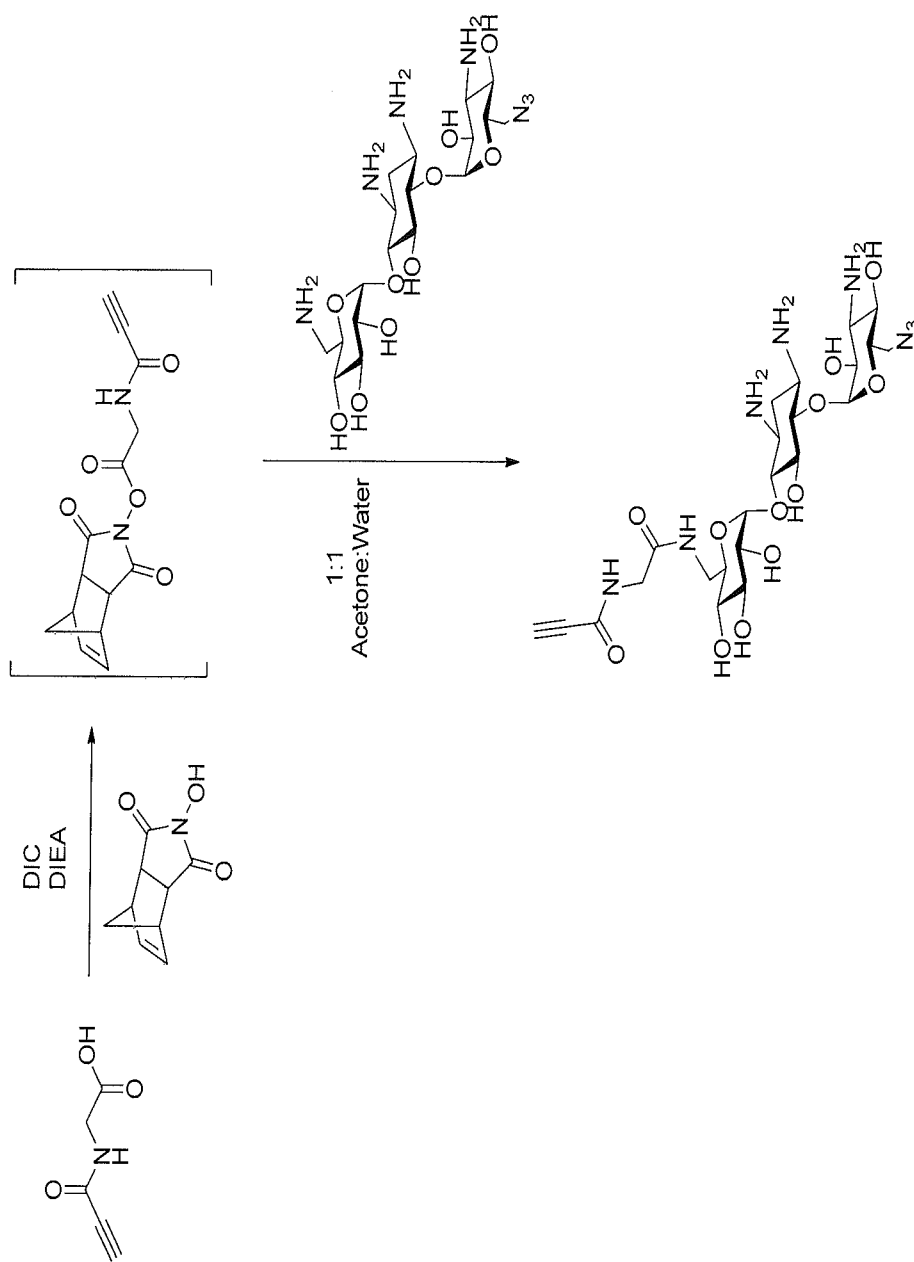
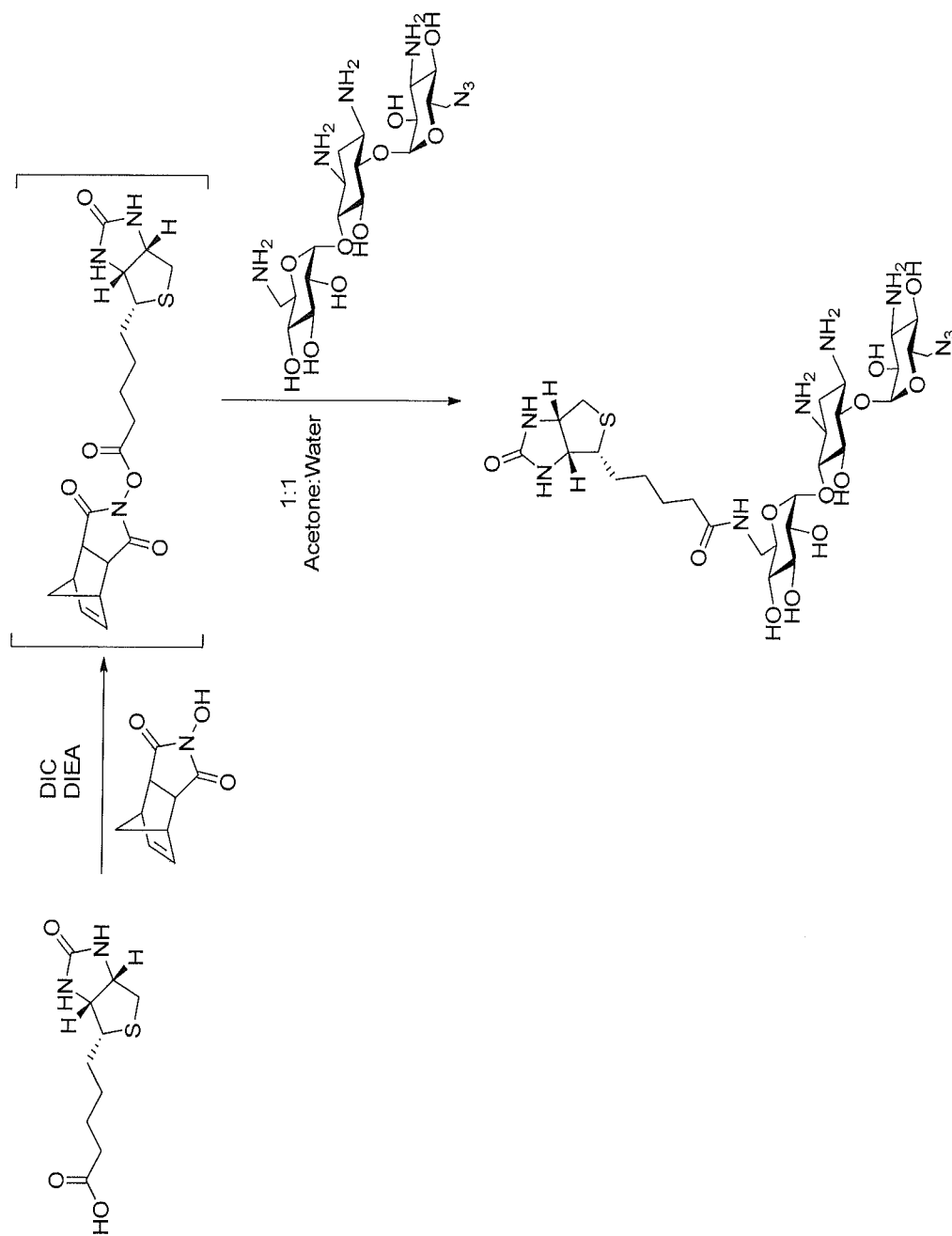


Figure 20







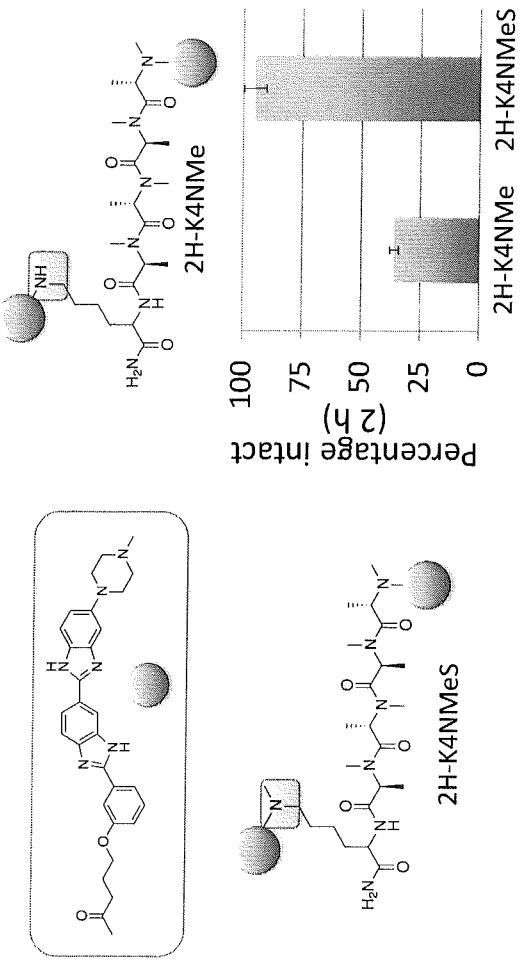


Figure 22



Figure 23

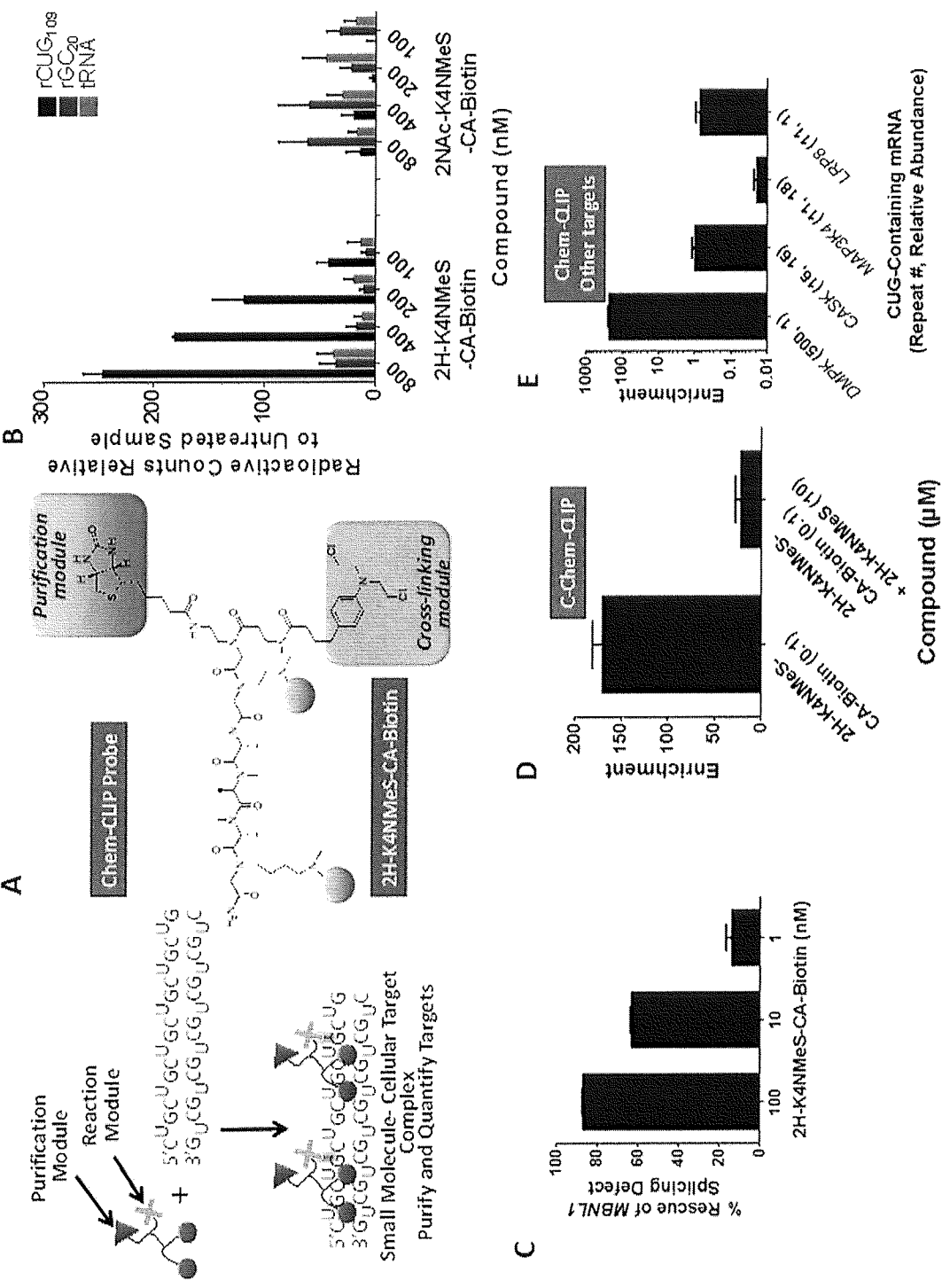


Figure 24

Figure 25

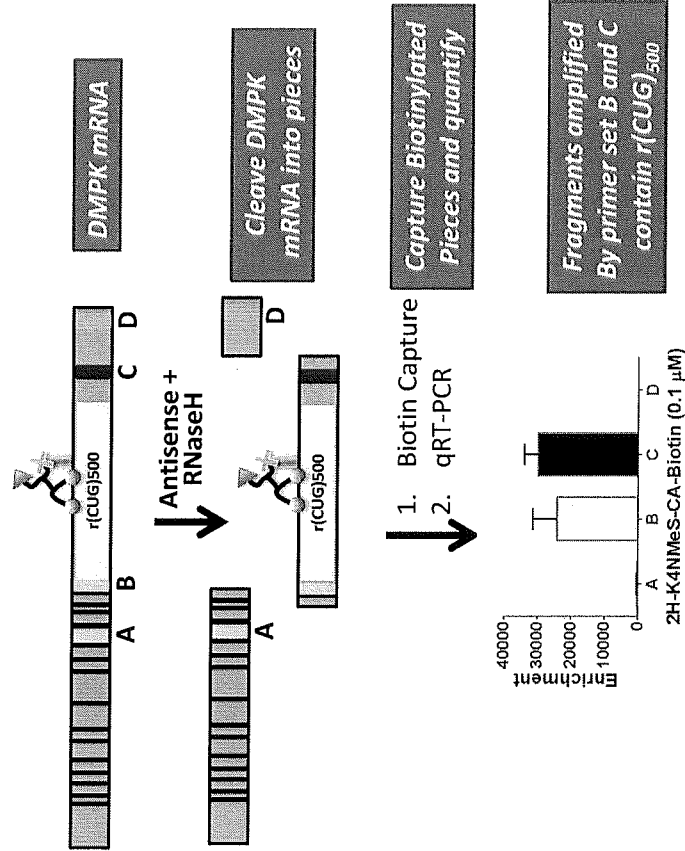


Figure 26

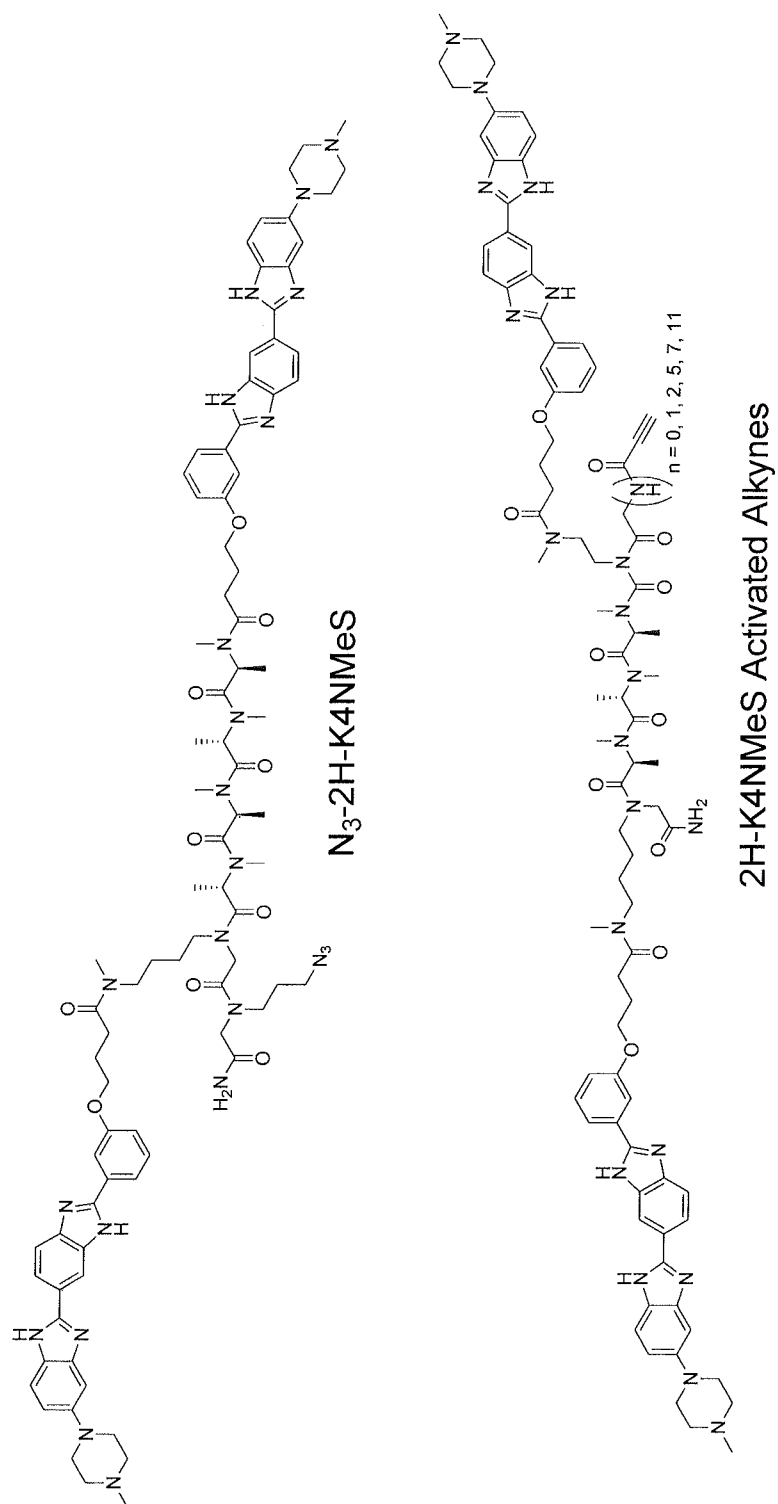
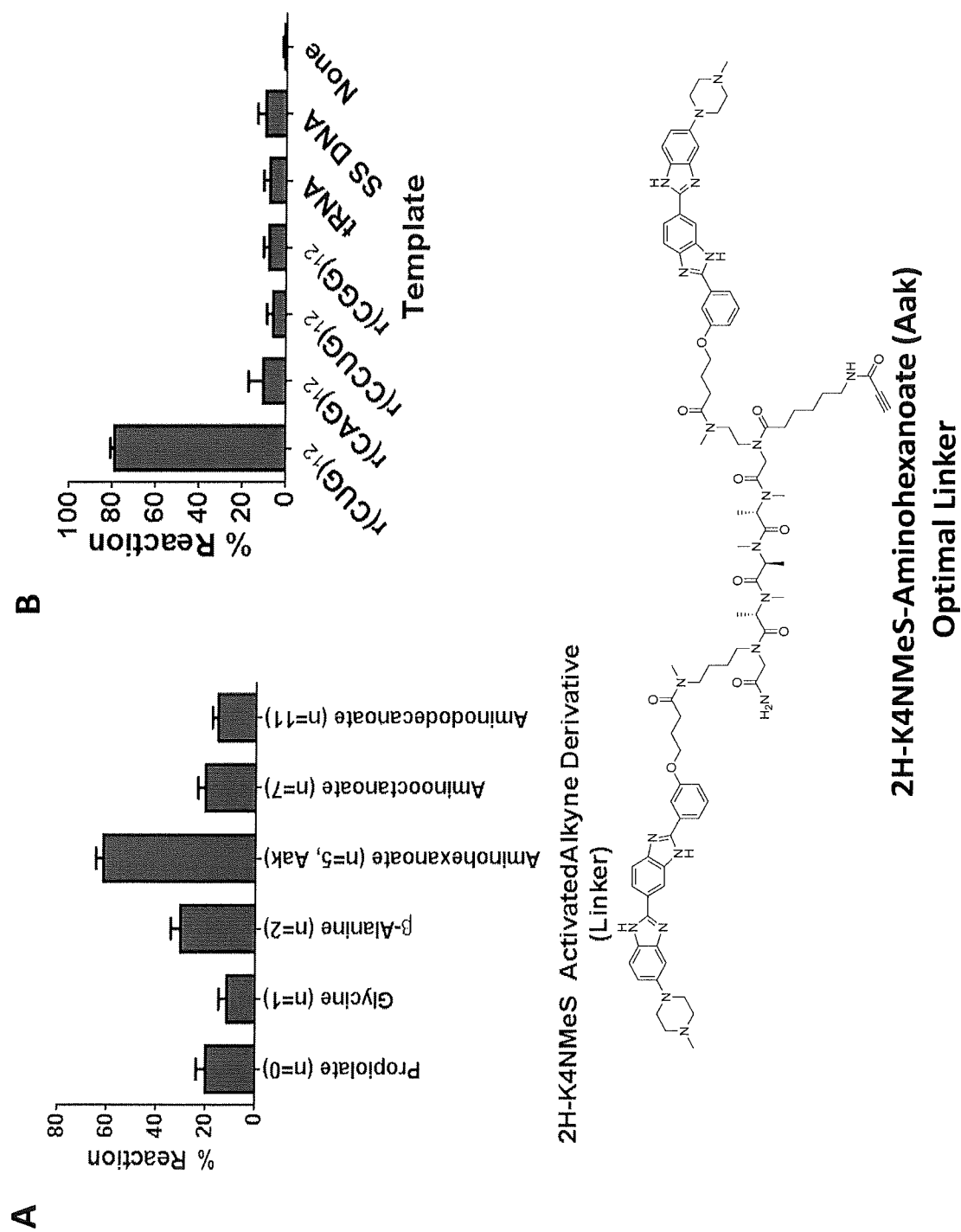


Figure 27



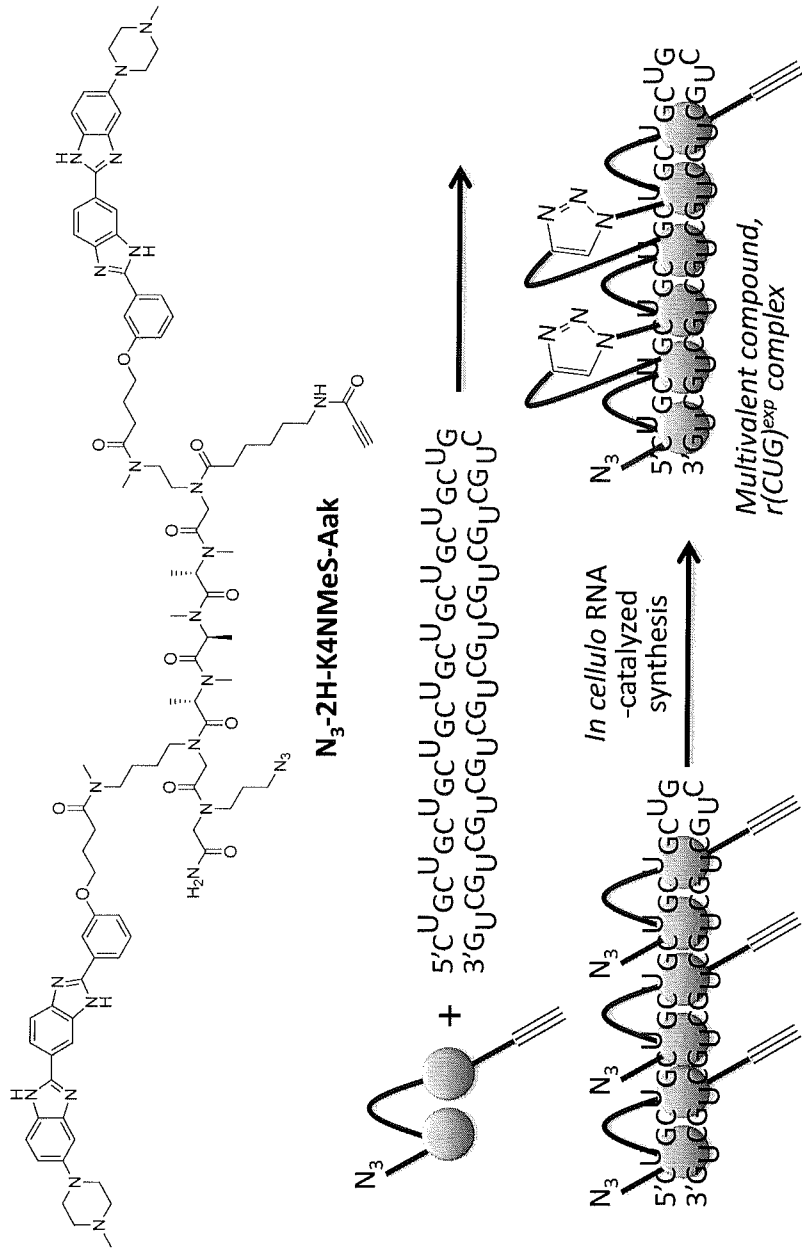


Figure 28

Figure 29

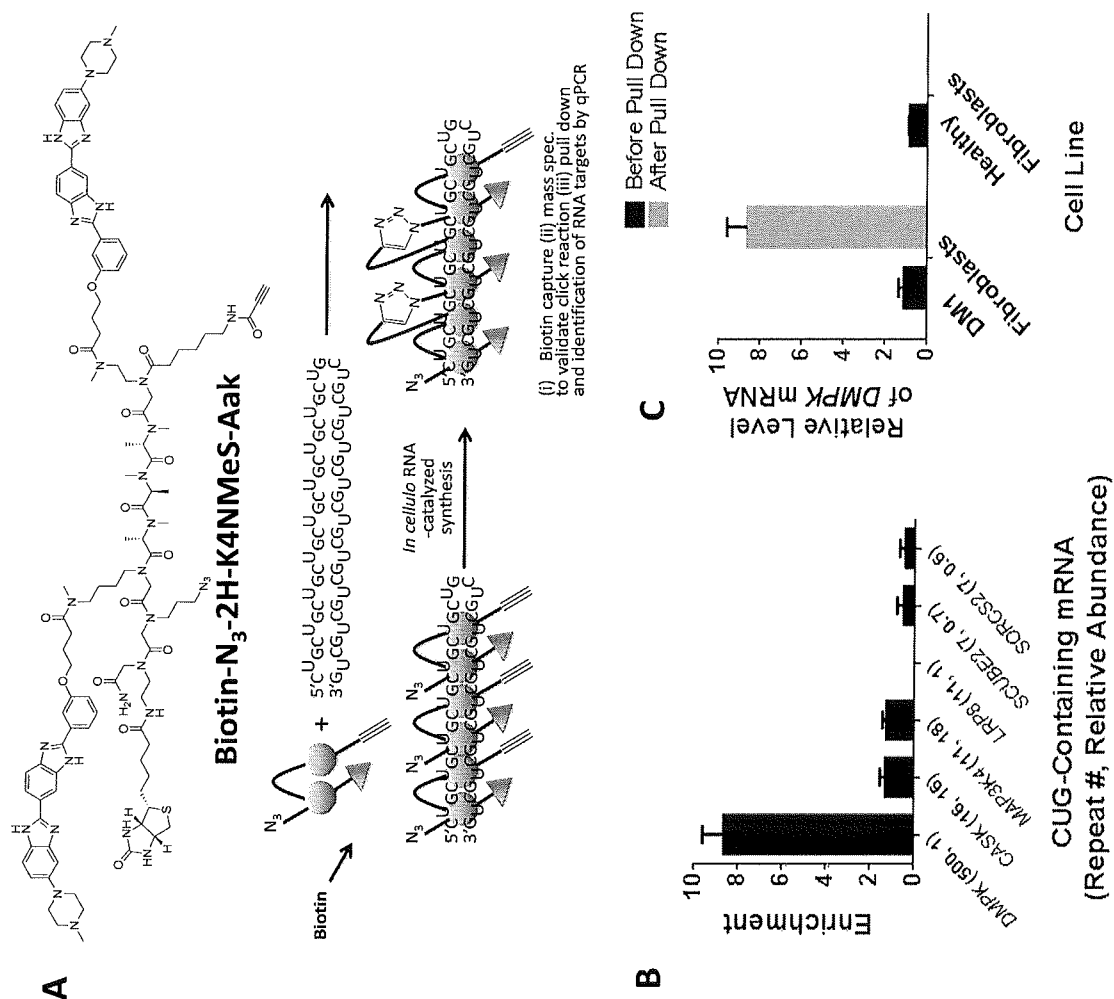




Figure 30

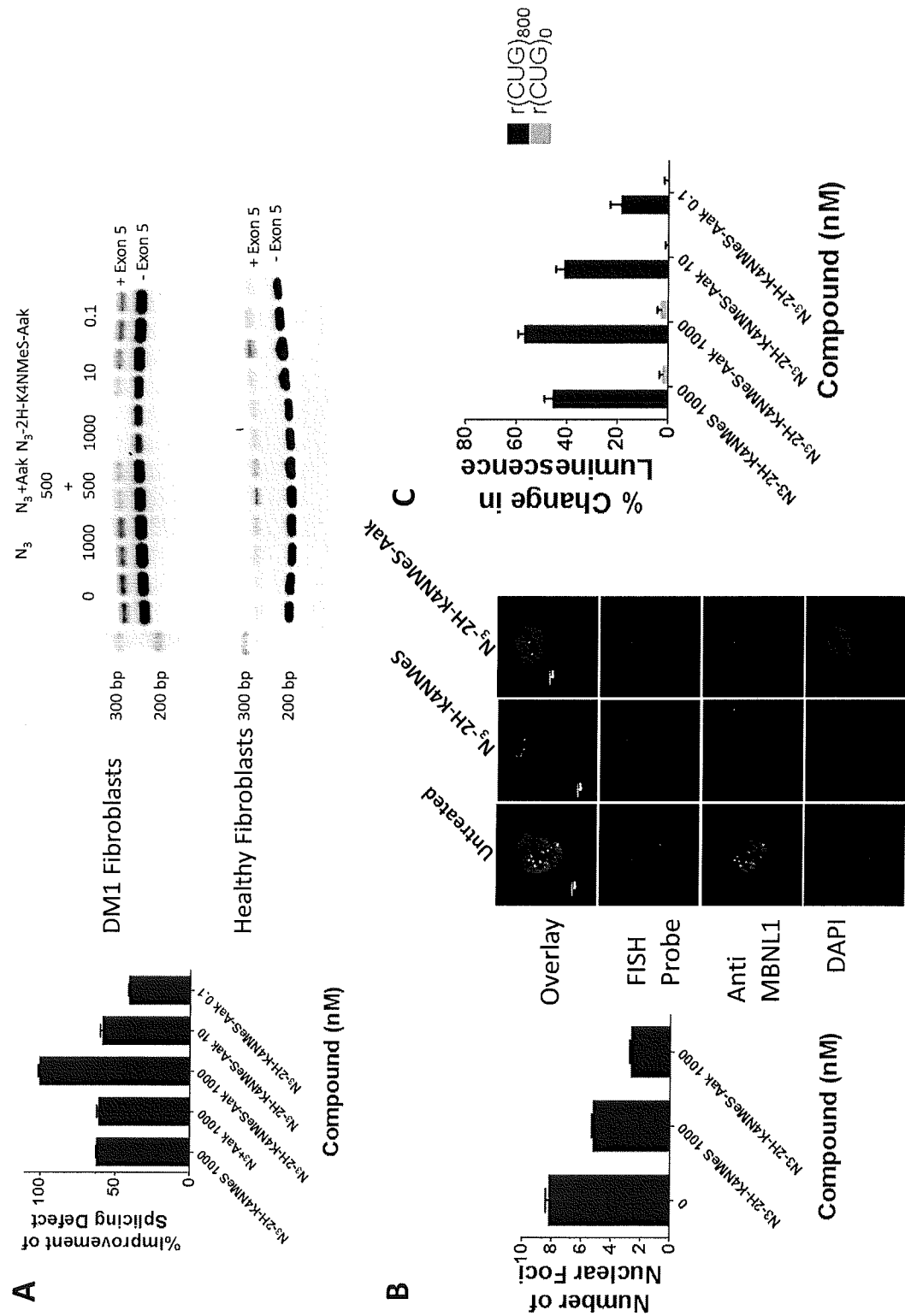


Figure 31

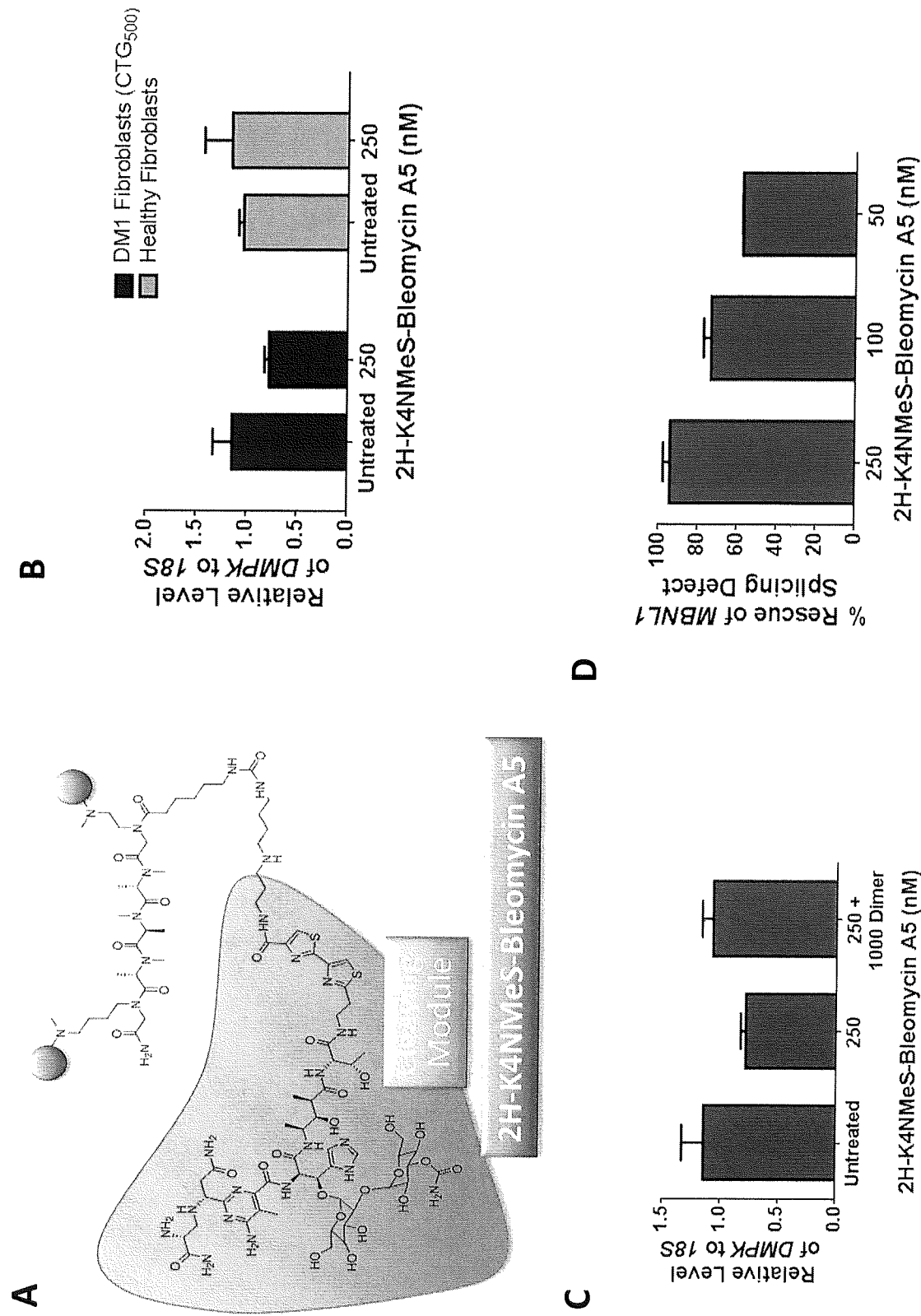


Figure 32

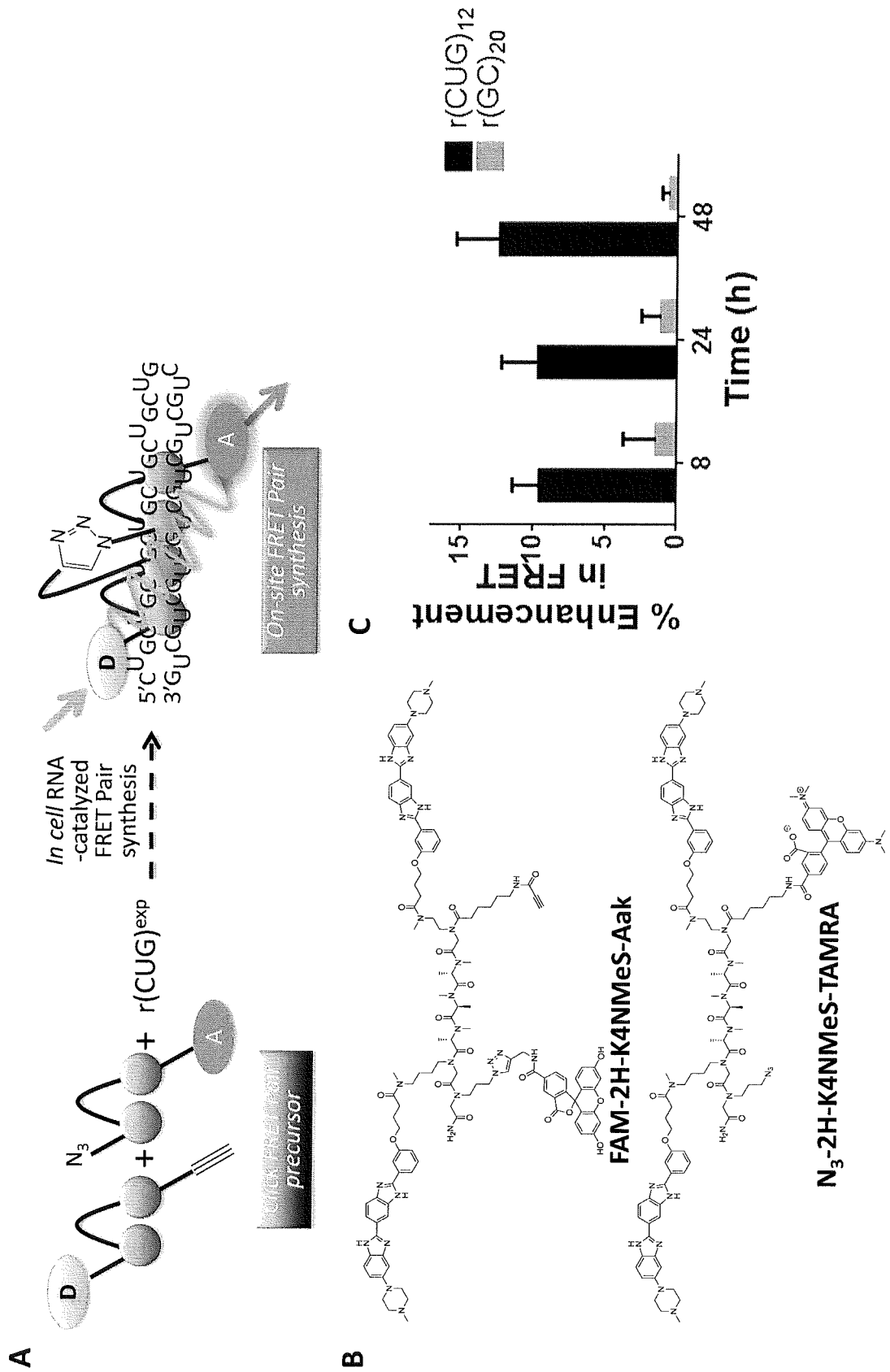
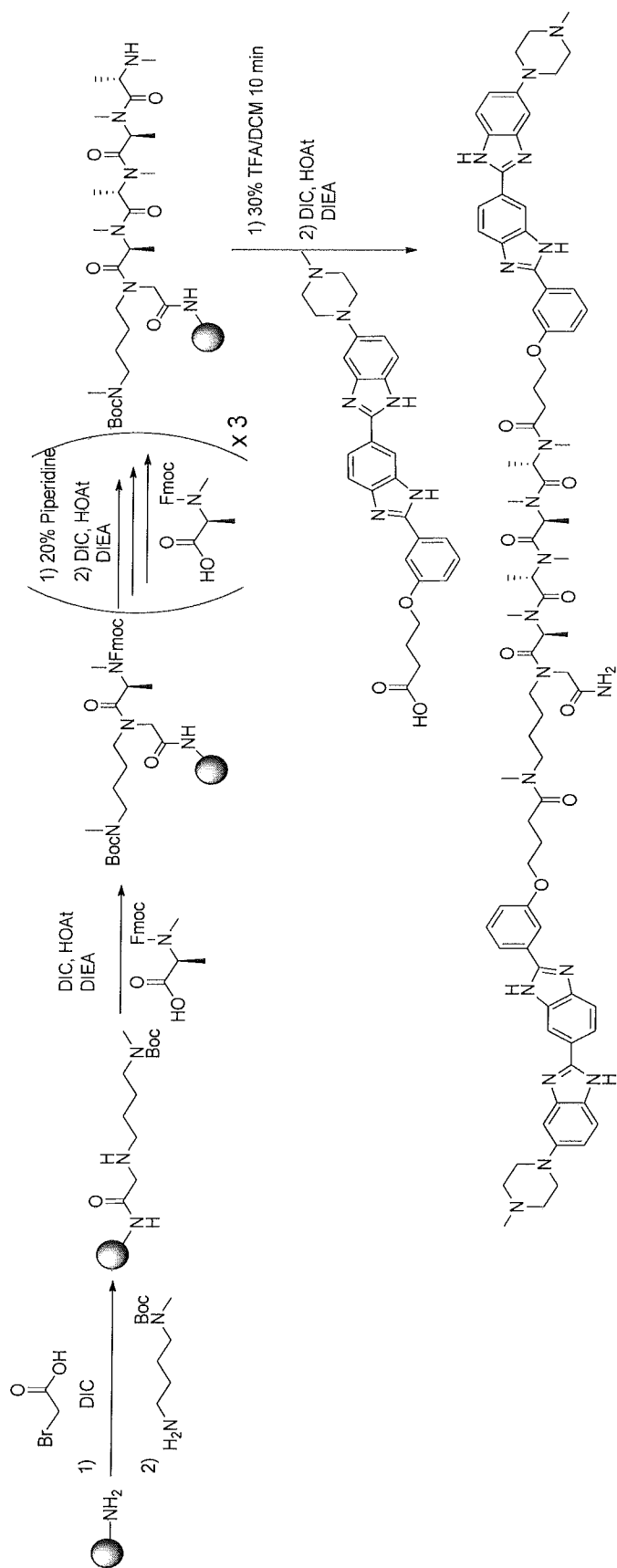


Figure 33



[illegible]

Figure 35

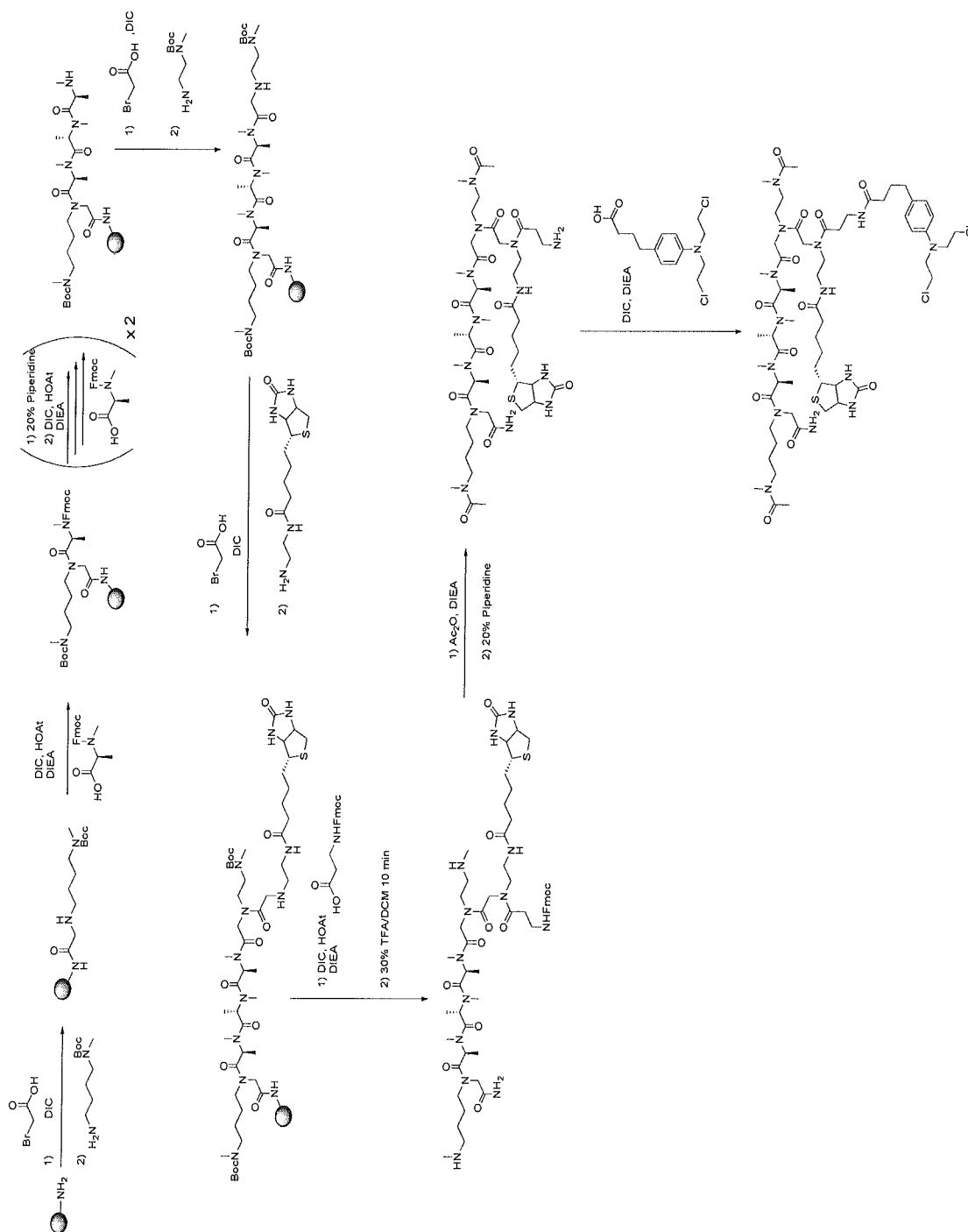


Figure 36

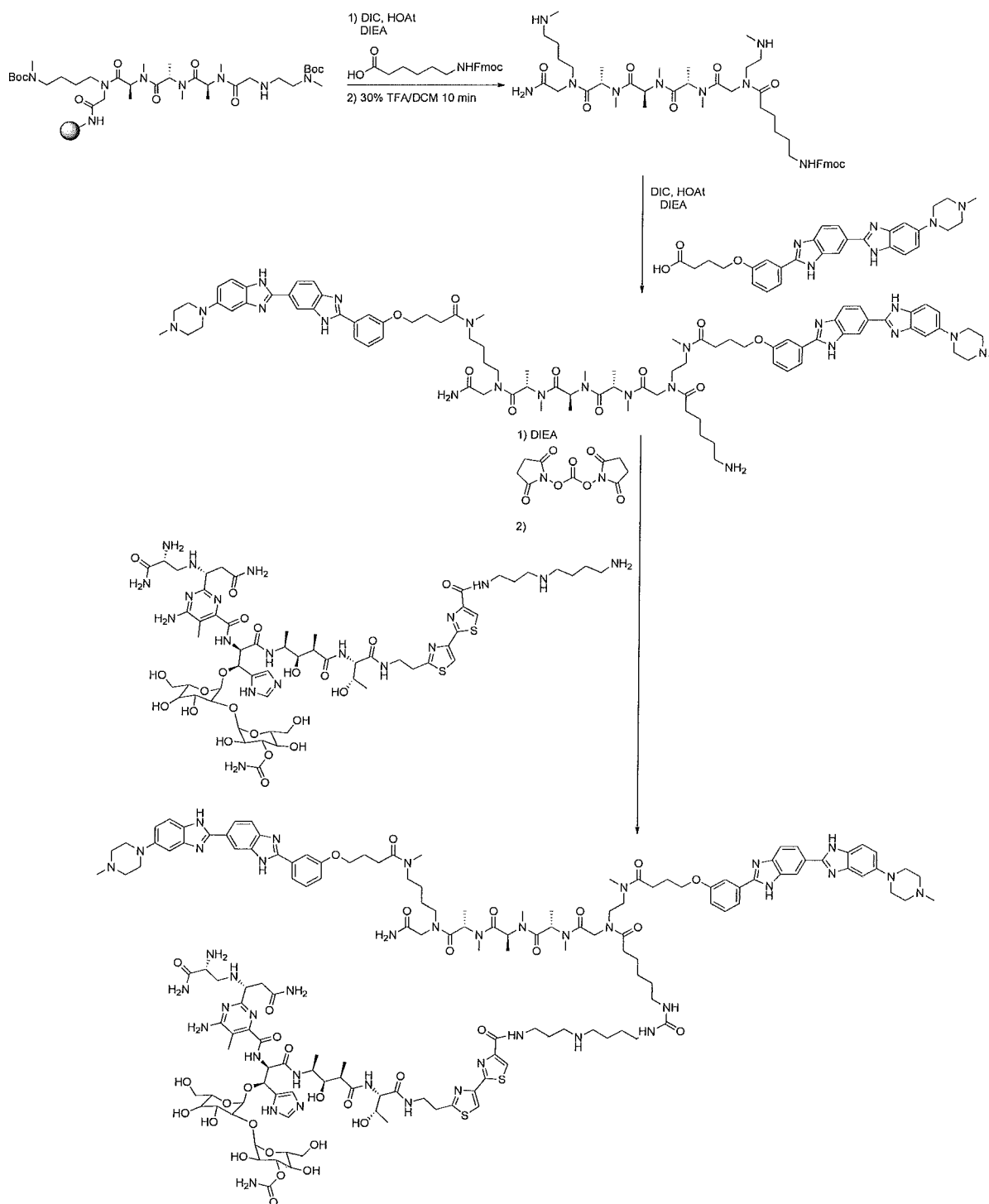
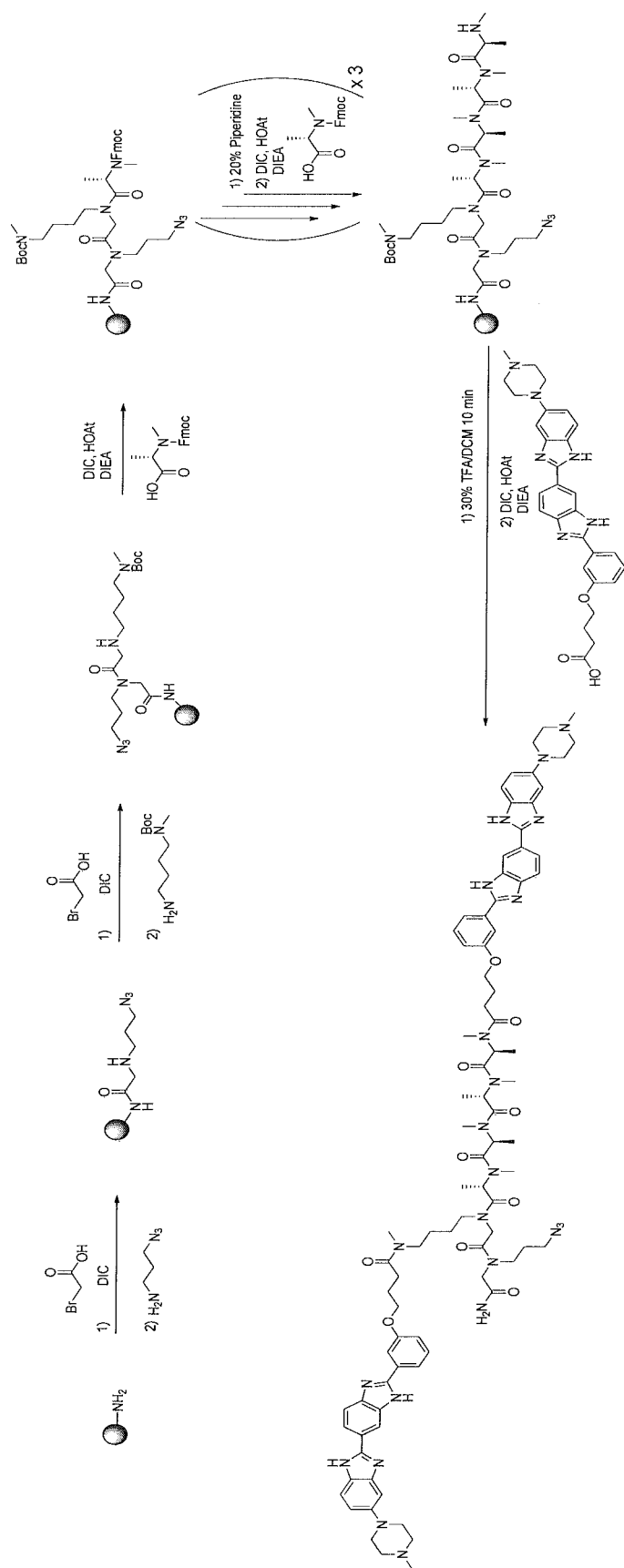
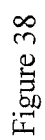


Figure 37







[illegible]

Figure 40

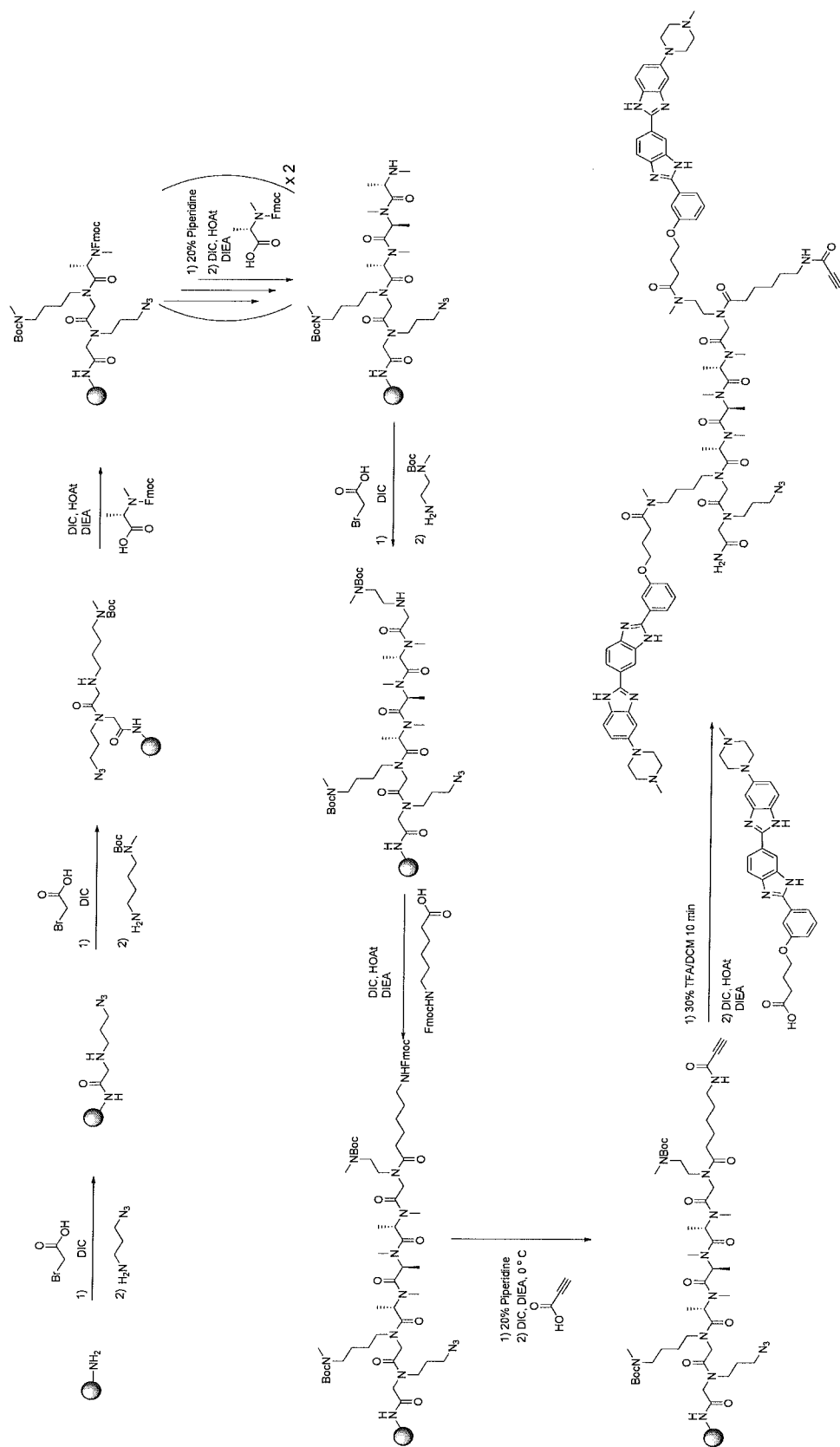


Figure 41

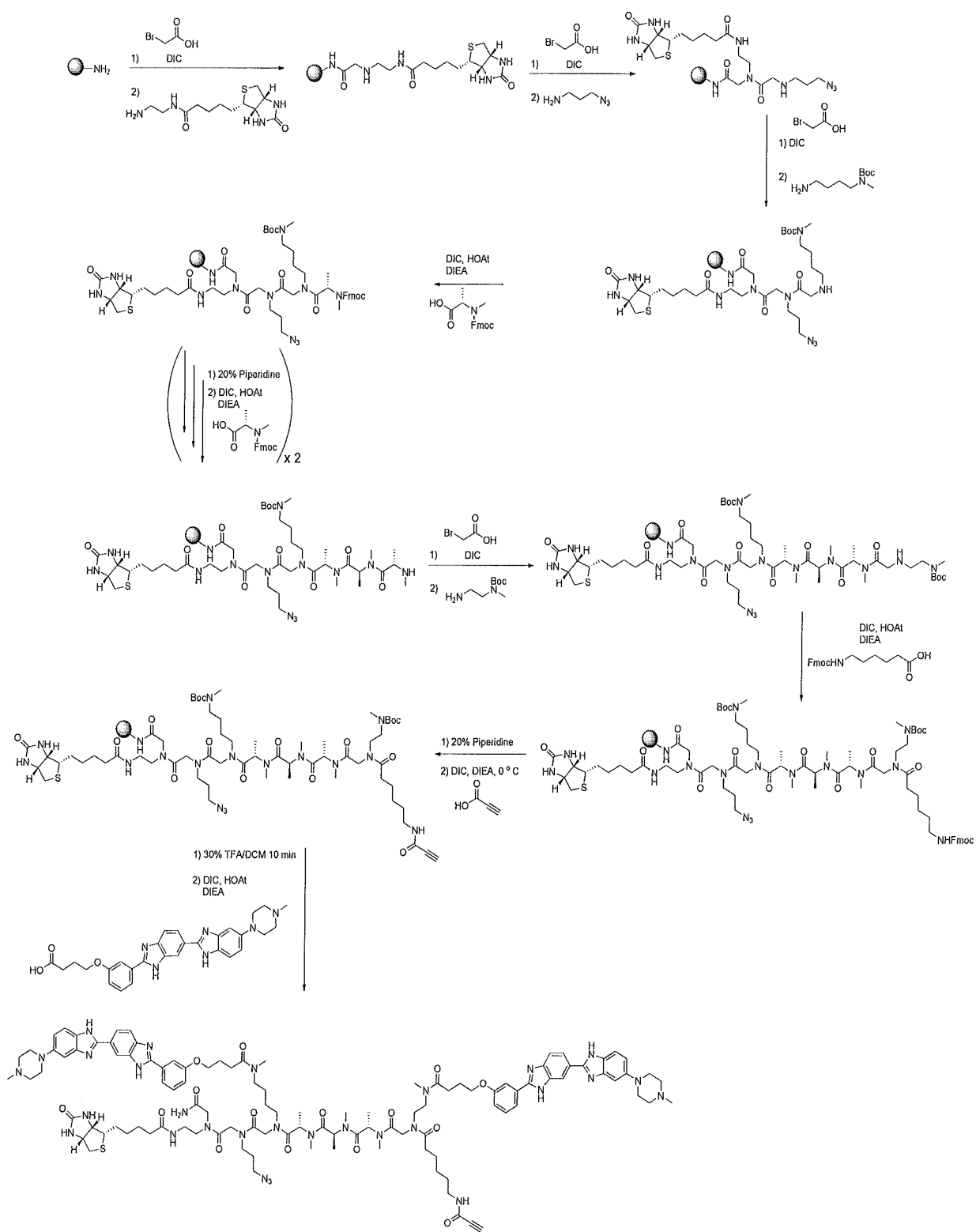


Figure 42

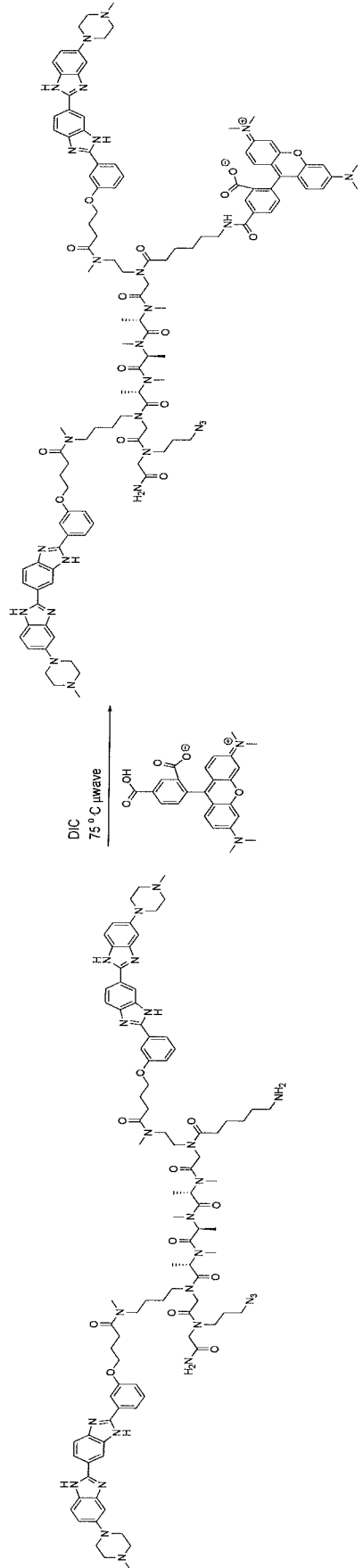


Figure 43

