Title

Synthetic Design of Vehicles for Macromolecule Delivery

Author

Roeise, Joachim Justad

Publication Date

2021

Peer reviewed|Thesis/dissertation

Synthetic Design of Vehicles for Macromolecule Delivery

by

Joachim Justad Røise

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Niren Murthy, Co-chair
Professor Matthew B. Francis, Co-chair
Professor Phillip Messersmith
Professor Michelle C. Chang

Summer 2021

Abstract

The following dissertation presents work done towards the development of delivery strategies for proteins and nucleic acids to mammalian cells, as well as targeted imaging materials to bacteria. Delivery of large molecules such as proteins, nucleic acids, and large synthetic molecules remains a major bottleneck in the field of drug delivery, and despite an abundance of research into the topic, the delivery of intracellular proteins has still not seen clinical significance. To address this, we sought to develop new delivery vehicles in an endeavor to address current challenges in drug delivery.

Chapter 1 discusses the development of acid-sensitive surfactants for the delivery of nucleic acids. This class of molecules, termed caged surfactants, consists of a membrane disruptive surfactant that has been masked with two PEG chains through an acetal linker. This prevents it from entering the membrane in its intact state; however, after acetal hydrolysis in the acidic endosomal compartment, release of free surfactant causes endosomal release of any co-delivered cargo. The simultaneous delivery of caged surfactant with nucleic acid was achieved using two strategies: (i) inclusion of two primary amines on the caged surfactant to allow complexation through electrostatic interactions, and (ii) inclusion of the RNA-binding dye thiazole orange (TO) to allow tight binding through intercalation. In this study we show that amine-containing caged surfactant (PCS) and TO-containing caged surfactant (TCS) increase the delivery of mRNA and siRNA, respectively.

Chapter 2 discusses the development of delivery vehicles for the delivery of proteins to mammalian cells. The chapter consists of two separate studies. The first study utilized a fluorescent acid-sensitive membrane disruptor (FEDS) as an adjuvant to Lipofectamine-based delivery of CRISPR-Cas9. FEDS successfully increased the delivery efficacy of Cas9 to HEK cells. In addition, due to its fluorescent properties the cells with Cas9-triggered gene editing could be sorted through a fluorescence cell sorter. The second study describes the development of a polyethylene glycol-Eosin Y block copolymer (PEG-pEosin) for protein delivery. The Eosin Y moieties bind a wide range of proteins, and PEG-pEosin can therefore be used as a general protein delivery tool. Using PEG-pEosin we successfully delivered Cre recombinase to Ai9 cells through a PEG-pEosin/Cre/Listeriolysin O complex, wherein the Lysteriolysin O serves as an endosomal disruptive component.

Chapter 3 presents a maltohexaose-indocyanine green conjugate for the detection of infective endocarditis. Maltohexaose (MH) is a sugar that is specifically taken up by bacteria, and it is relatively tolerant of modifications to its reducing end. The FDA-approved dye indocyanine green (ICG) was conjugated to MH to yield a low toxicity, high wavelength probe that was shown to accumulate in bacterial vegetations in an endocarditis rat model.

Table of contents

Abbreviations

Bn	Benzyl
Bz	Benzoyl
Cas9 RNP	Cas9 ribonucleoprotein
CDC	Cholesterol-dependent cytolysin
DEPC	Diethylpyrocarbonate
DIPEA	N,N-Diisopropylethylamine
DLS	Dynamic light scattering
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
eGFP	Enhanced Green Fluorescent Protein
FACS	Fluorescence cell sorting
HEK cells	Human embryonic kidney cells
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectroscopy
HUVEC	Human umbilical vein endothelial cell
ICG	Indocyanine green
IE	Infective endocarditis
LLO	Listeriolysin O
MD	Molecular dynamics
MeCN	Acetonitrile
MH	Maltohexaose
mRNA	Messenger RNA
pAsp	Polyaspartic acid
PBS	Phosphate-buffered saline
PCS	Positively charged (amine containing) caged surfactant
pDNA	Plasmid DNA
PEG	Polyethylene Glycol

RBC	Red blood cells
siRNA	Small interfering RNA
TCS	Thiazole orange containing caged surfactant
TEA	Triethylamine
TFA	Trifluoroacetic acid
TLC	Thin layer chromatography
TO	Thiazole Orange
V-ATPases	Vacuolar-type ATPases
ZFN	Zinc finger nuclease

Acknowledgements

I would like to thank my family, in particular my parents, Kjell Åge and Hilde, and my brother and sister, Alexander and Andrea, for always being supportive and there for me. I would also like to thank Jenna for being an important part of my life and staying supportive throughout my time in Berkeley. I can not imagine how I would have gotten through my time here without you.

I would also like to thank the graduate students and postdocs in the Murthy lab who were welcoming and always willing to teach me new things, without you I would not have been able to grow as a researcher. Specifically I would like to thank Corinne Sadlowski, Santanu Maity, Xiaojian Wang, Jingtuo Zhang, Bora Park, Tara deBoer, Vanessa Mackley, Kunwoo Lee, Giri Vegesna, Subhamoy Das, Yumiao Zhang, Jeremy Adams, Jie Li, Hesong Han, Maomao He, I-Che Li, Eli Espinoza, Sören Reinhard, Teena Bajaj, Michael Stentzel and Dake Chen.

My undergraduate students, especially Aaron, Lucas, and Emily, for trusting me enough to play a part in their education, and always providing me with new perpectives on all of our projects. Hopefully the experience you gained will be as helpful to you as it has been for me.

And lastly, I would like to thank Professor Murthy for letting me join the lab with minimal research experience, and no experience in bioengineering. I have learned a lot the last six years through our discussions, and you have always stayed patient and willing to explain concepts that I was not familiar with.

Chapter 1 – Acid-Sensitive Surfactants Designed to Enhance the Delivery of Nucleic Acids

The development of endosomal disruptive agents is a major challenge in the field of drug delivery and pharmaceutical chemistry. Current endosomal disruptive agents are composed of polymers, peptides, and nanoparticles, and have been difficult to optimize and translate because of their complicated chemistry. In this report we demonstrate that low molecular weight surfactants can be converted into endosomal disruptive agents via reversible PEGylation. This new class of endosomal disruptive agents are termed caged surfactants and are the first example of low molecular weight molecules that are pH sensitive membrane disruptive agents. The caged surfactants have the potential to address several of the limitations hindering the development of current endosomal disruptive agents, and will be amenable to traditional medicinal chemistry approaches for optimization. In this report we synthesized three generations of caged surfactants, and demonstrate that they can enhance the ability of cationic lipids to deliver mRNA into primary cells. We also show that caged surfactants are able to deliver siRNA into cells when modified with the RNA-binding dye thiazole orange (TO). We anticipate that the caged surfactants will have numerous applications in pharmaceutical chemistry and drug delivery given their small size and well-defined chemistry.

Chapter 1 is based on the following research study:

Røise JJ ‡, Han H ‡, Li J, Kerr DL, Taing C, Behrouzi K, He M, Ruan E, Chan LY, Espinoza EM, Reinhard [S], Thakker K, Kwon J, Mofrad MRK & Murthy N. *Acid-Sensitive Surfactants Enhance the Delivery of Nucleic Acids,* In Review

1.1 Introduction

The degradation of endocytosed nucleic acid therapeutics in the lysosome is a central problem in drug delivery and limits the development of a wide range of experimental therapeutics[1,2]. For example, the percentage of endocytosed nucleic acid/cationic lipid complexes that reach the cytoplasm is only between 1-5%[3]. There is consequently an urgent need to develop strategies for enhancing the endosomal release of nucleic acid therapeutics. Upon endocytosis, the endosomal compartment sees a significant change in pH as it progresses down the endosomal ladder, with early endosomes having a pH of 6.5, to lysosomes which have pH values down to 4.0.[4] This acidification is driven by vacuolar-type ATPases (V-ATPases), which is a signature of the endosomal pathway. This difference in pH has been widely exploited to develop agents that can selectively disrupt endosomes, termed endosomal disruptive agents.[5–7]

Mainly two classes of pH sensitive triggers have been developed to engineer pH-sensitive membrane disruption, based upon either the protonation of a carboxylic acid group or the hydrolysis of an acid degradable linkage. Both strategies function by increasing the hydrophobicity of the endosomal disruptive agent at pH 5.0 versus 7.4, enabling membrane insertion at pH 5.0 but not at pH 7.4.

The earliest endosomal disruptive agents were based upon the protonation of carboxylic acid groups. Carboxylic acids are found in a variety of peptides, proteins, and polymers and undergo a significant change in hydrophobicity upon protonation, hence they are ideal for designing endosomal disruptive agents and have found numerous applications[5–8]. For example, glutamic acid-containing peptides, such as the GALA peptide[9], disrupt lipid bilayers at micromolar concentrations at pH 5.0, but have no membrane disruptive ability at pH 7.4. These endosomal disruptive peptides were capable of enhancing DNA, mRNA, and siRNA delivery into cells[10–12], and demonstrated that endosomal disruptive agents could have a significant impact on nucleic acid delivery. Endosomal disruptive agents based upon synthetic polymers, such as poly(propyl acrylic acid) have also been developed[13], which undergo large changes in hydrophobicity in acidic environments and are capable of dramatically enhancing the delivery of nucleic acids to cells and *in vivo*.

The second class of endosomal disruptive agents utilize reversible PEGylation to achieve an increase in hydrophobicity[14,15]. In this strategy, PEG is generally conjugated to a hydrophobic membrane disruptive molecule, typically a polymer, through an acid degradable acetal linkage. The PEG chains prevent the hydrophobic backbone from partitioning into the lipid bilayer; however, after endocytosis, the linkage between the PEG and polymer hydrolyzes and the polymer can then insert into the membrane and disrupt it. This strategy has tremendous versatility because PEG can be conjugated to a variety of endosomal disruptive agents, such as cationic polymers, liposomes, and nanoparticles[14–17]. In addition, because the size of the PEG can be easily varied, this strategy allows for the robust chemical optimization of endosomal disruptive agents.

A key limitation of existing endosomal disruptive agents is their large size and chemical complexity. Low molecular weight endosomal disruptive agents have the potential to address several of the challenges that have limited the translational development of endosomal disruptive agents. In particular, they will be well-defined and straightforward to synthesize and characterize, in contrast to polymers and nanoparticles which often do not achieve the same level of uniformity. In addition, low molecular weight endosomal disruptive agents are excretable, lack antigenicity,

and are amenable to accessible chemical modifications and optimization. However, despite their potential, the development of low molecular weight endosomal disruptive agents has been challenging. At present the only strategy available for disrupting endosomes with small molecules is via the proton sponge effect based on molecules such as chloroquine. However, chloroquine requires high concentrations and incubation times of 10-12 hours to efficiently promote nucleic acid delivery, which in addition to its toxicity severely limits its use as an endosomal disruptive agent.[18,19]

Due to the current lack of small molecule endosomal disruptors that are not based on the proton sponge effect, alternative designs for enabling pH-dependent membrane disruptive features are sorely needed. We therefore sought to develop a class of small molecule endosomal disruptors that act by direct insertion and pore-formation on the endosomal membrane. In this report, we utilize the acid-triggered release of a low molecular weight surfactant to achieve endosomal disruption, as seen in Figure 1A. In addition, we utilized two nucleic acid-binding moieties, amines and thiazole orange (Figure 1B and 1C), to enable complexation with nucleic acids, allowing for their cytosolic delivery.

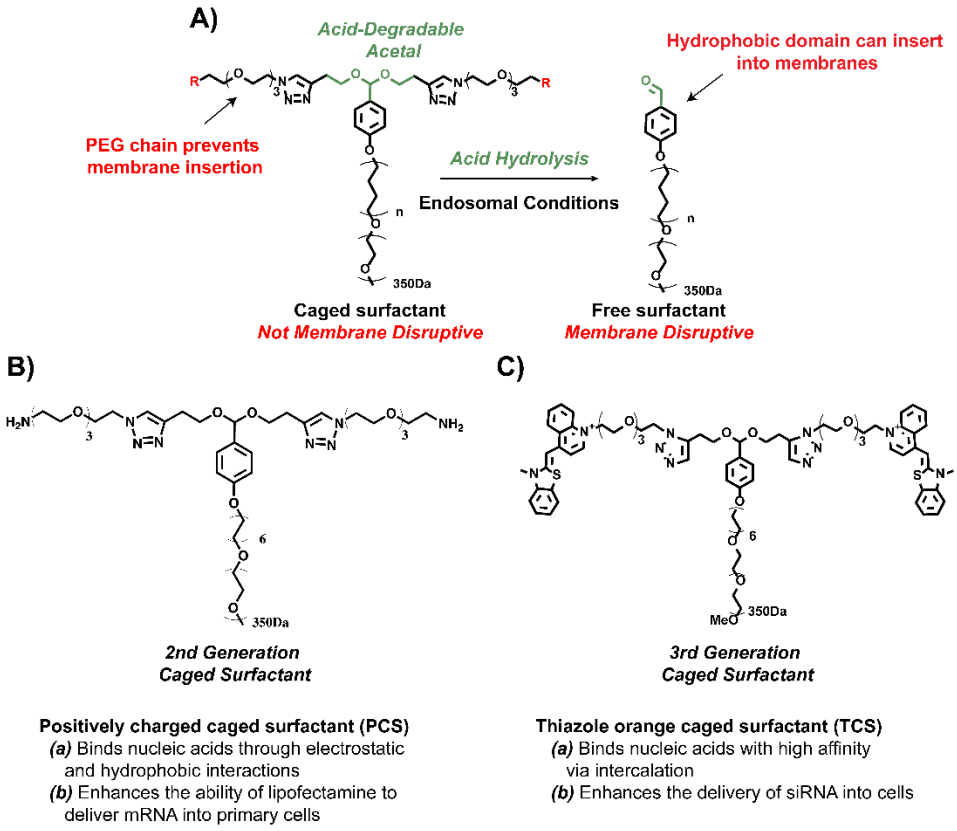


Figure 1: ***The caged surfactants: Low molecular weight endosomal disruptive agents that can enhance the delivery of mRNA and siRNA***

(A) The caged surfactants are composed of a surfactant that has been reversibly masked with PEG chains. The caged surfactants are not membrane disruptive at pH 7.4. However, after endocytosis, the acidic pH

of the endosome hydrolyzes their acetal linkage, unmasking their hydrophobic domains, which can then disrupt endosomes. To our knowledge, caged surfactants achieve pH-sensitive membrane disruption with the most efficient molecular weight economy available. In this report we synthesized 2nd generation and 3rd generation caged surfactants, which were able to enhance the delivery of mRNA and siRNA into cells.

(B) ***2nd generation caged surfactants****: A positively charged caged surfactant (PCS) that complexes nucleic acids via electrostatic and hydrophobic forces. PCS was able to enhance the ability of lipofectamine to transfect mRNA into hard-to-transfect primary cell cultures.*

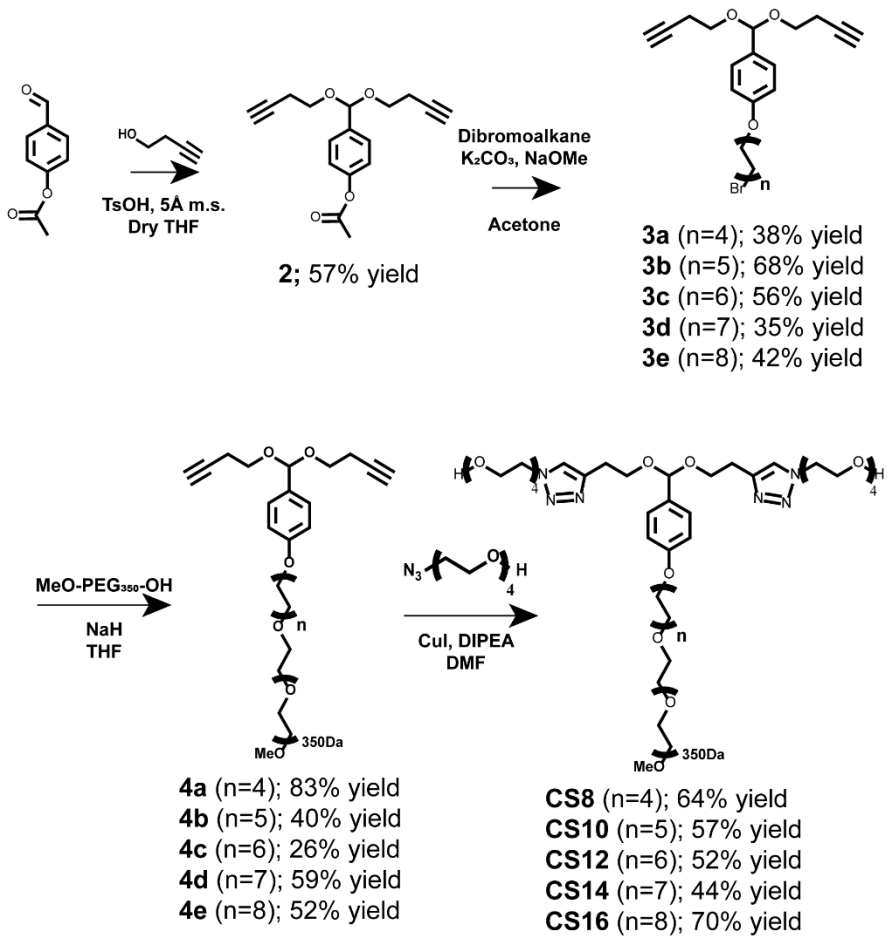
(C) ***3rd generation caged surfactants****: A caged surfactant that contains the RNA binding dye thiazole orange (TCS). TCS was able to enhance the delivery of siRNA into cells.*

1.2 Alcohol Caged Surfactants

In order to synthesize efficient caged surfactants, it is important to optimize two parameters: (1) their membrane disruptive ability, and (2) their hydrolysis kinetics. The 1st generation of caged surfactants were synthesized to investigate the effect of varied hydrophobic chain lengths to achieve optimal pH-dependency by optimizing for these two parameters. Although membrane partitioning, which leads to membrane disruption, of Triton X-like surfactants scales with the length of their hydrophobic domains, it is unclear if caged surfactants may themselves cause membrane disruption if their hydrophobic segments are made too prominent. We consequently synthesized a variety of caged surfactants, which had hydrophobic domains composed of eight to sixteen-carbon chain followed by a benzaldehyde moiety, to determine the optimal hydrophobic chain length. We then measured their acid-dependent hydrolysis rates to choose the hydrophobic length with optimal hydrolysis kinetics.

Synthesis

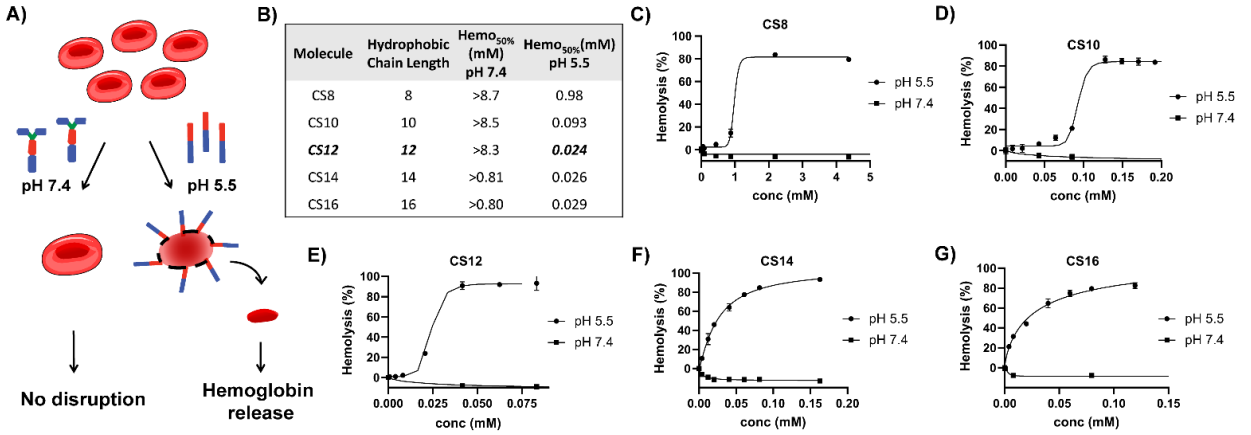
We synthesized five caged surfactants with alkyl chains ranging from eight to sixteen carbons to determine the optimal hydrophobicity. The 1st generation caged surfactants contain a $(HO\text{-}PEG)_2$-caging group. The synthetic strategy for the five caged surfactants (**CS8**-**CS16**) is shown in Scheme 1. In short, treatment of aldehyde **1** with 1-butyn-4-ol in acidic conditions gave the acetal **2** with a yield of 57%. The acetal **2** has an ester in its para position, with a σ_{para} value of 0.31, and consequently generates a relatively stable acetal that forms in high yields. In contrast, the hydroxy and methoxy groups have σ_{para} values of -0.38 and -0.27, respectively[20], and making acetals with these fragments is therefore much more challenging. Compound **2** was then treated with NaOMe in the presence of a dibromo-alkyl compound with 8, 10, 12, 14 or 16 carbons, yielding compounds **3a-e,** respectively. These compounds were then added to a solution of $MeO\text{-}PEG_{350}\text{-}OH$ pretreated with NaH, followed by stirring overnight to yield compounds **4a-e**. Finally, the caged surfactants **CS8-CS16** were synthesized through a copper-click cycloaddition with tetraethyleneglycol monoazide.



Scheme 1: Synthesis of the 1st generation caged surfactants CS8-CS16.

Membrane Disruptive Abilities

The caged surfactants are designed to function as pH-sensitive membrane disruptive agents. As a model system, we measured the ability of the caged surfactants **CS8-CS16** to disrupt red blood cell membranes at pH 7.4 and 5.5 (Figure 2A) [5]. Using this hemolysis assay, we found that all the caged surfactants are pH-sensitive membrane disruptive agents (Figure 2C-G). For example, the caged surfactant **CS12** has a $Hemo_{50\%}$ of 24 µM at pH 5.5 and generated no hemolysis at pH 7.4 at concentrations up to 8.3 mM, the highest concentration measured. Interestingly, **CS16**, which bears the longest hydrocarbon chain also shows no hemolysis at pH 7.4 despite having the largest hydrophobic surface. The hydrolyzed aldehyde corresponding to **CS8** was independently synthesized and showed no pH dependency (Figure 4). This confirms that the pH-sensitive membrane disruptive property of the caged surfactants is due to the masking PEG chains. These results demonstrate that the branched $(HO\text{-}PEG)_2$- structure is remarkably efficient at preventing membrane insertion of long hydrophobic domains even though it consists of only short PEG chains. The dePEGylation strategy can therefore be used to generate low molecular weight pH-sensitive membrane disruptive agents.



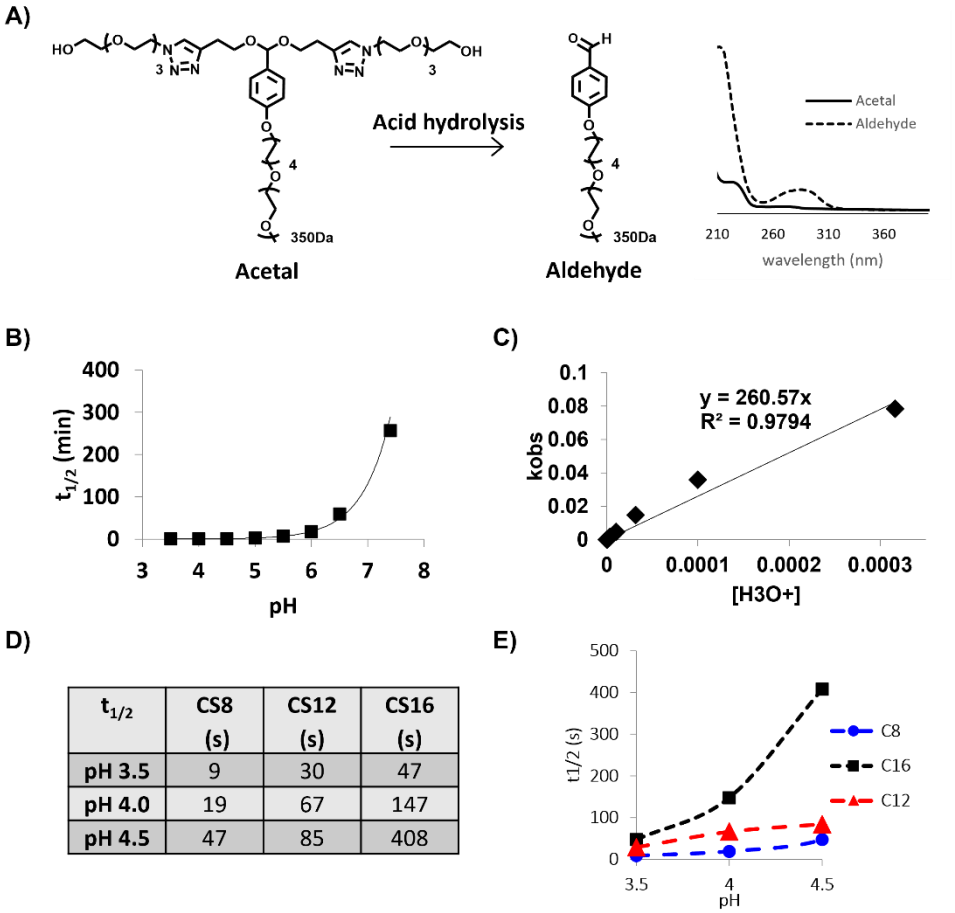
Molecule	Hydrophobic Chain Length	$Hemo_{50\%}$ (mM) pH 7.4	$Hemo_{50\%}$(mM) pH 5.5
CS8	8	>8.7	0.98
CS10	10	>8.5	0.093
CS12	***12***	>8.3	***0.024***
CS14	14	>0.81	0.026
CS16	16	>0.80	0.029

Figure 2: ***The caged surfactants are pH sensitive membrane disruptive agents****. The caged surfactants disrupt red blood cell membranes in a pH sensitive manner. A) The caged surfactants were mixed with red blood cells and their ability to disrupt them was measured at pH 5.5 and 7.4. B) Shows the concentration of CS at which 50% hemolysis is achieved at either pH 7.4 or 5.5. All of the caged surfactants CS8-CS16 are pH sensitive membrane disruptive agents (see Scheme 1 for chemical structures). The values at pH 7.4 represent the highest concentration tested, and several of the caged surfactants, in particular CS12-CS16 are remarkably efficient at performing membrane disruption and have Hemo50% in the micromolar range. 50% hemolysis was not reached at pH 7.4 for any of the caged surfactants. C-F) Shows the hemolysis curves at pH 5.5 and pH 7.4 for compounds CS8-CS16 respectively.*

The membrane disruptive efficiency of the caged surfactants also increases with the hydrocarbon chain length (Figure 2B). For example, **CS12** was approximately 37 times more efficient at disrupting membranes than **CS8**, with a $Hemo_{50\%}$ value decreasing from 900 µM to 24 µM at pH 5.5. This activity is similar to endosomal disruptive polymers such as polyethyl-acrylic acid, despite being 1/10 the molecular weight[8]. The increased membrane disruptive activity of **CS12** over **CS8** is not surprising due to its larger hydrophobic surface. However, it is unclear why the membrane disruptive efficacy of the caged surfactants does not improve for caged surfactants with chains longer than for **CS12**.

Hydrolysis Kinetics

The hydrolysis kinetics of the caged surfactants is a key parameter that will have a significant impact on their efficacy. Upon endosomal uptake, biomolecules are trafficked to lysosomes within 30 minutes[21], and consequently, the caged surfactants need to hydrolyze with a half-life of <30 minutes in endosomal conditions to ensure that endosomal disruption occurs before lysosomal degradation. We therefore measured the hydrolysis kinetics of the caged surfactants **CS8, CS12** and **CS16** via absorption spectroscopy (λ=286 nm), at a 50 µg/mL concentration at 37°C. Figure 3 demonstrates that the caged surfactants hydrolyze rapidly at pH 5.0. For example, **CS8** had a $k_{2nd} = 261\ s^{-1}M^{-1}$, which corresponds to a hydrolysis half-life of 2.5 minutes at pH 5.0 and over 4 hours at pH 7.4. The hydrolysis kinetics of **CS12** and **CS16** show similar pH-sensitive hydrolysis, but overall, longer hydrocarbon chains lead to slower hydrolysis rates. For example, hydrolysis rates at pH 4.5 for **CS8**, **CS12**, and **CS16** were found to be 0.8 min, 1.4 min, and 6.8 min respectively. Hydrolysis graphs are shown in Figure 13-26 of Appendix A.

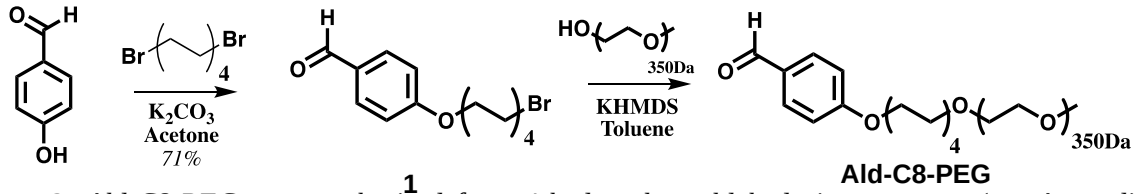


$t_{1/2}$	CS8 (s)	CS12 (s)	CS16 (s)
pH 3.5	9	30	47
pH 4.0	19	67	147
pH 4.5	47	85	408

Figure 3: Hydrolysis kinetics for the caged surfactants. A) CS8 hydrolyzes in acid, giving the free aldehyde. The aldehyde has an absorption peak with λ=286nm. B) Plot of hydrolysis half-life of CS8 at a range of pH conditions. C) Plotting k_{obs} over $[H_3O^+]$ gives the k_{2nd} of CS8 as the slope. D) and E) Hydrolysis half-lives for CS8, CS12 and CS16 at pH 3.5, 4.0 and 4.5 at 37°C, here shown in seconds.

Synthesis and Hemolytic Efficacy of Aldehyde Surfactant

An important feature of the caged surfactants is their pH sensitivity. In order to confirm that the hemolytic efficacy could be attributed to the release of free surfactant, the free surfactant **Ald-C8-PEG** (Scheme 2) was synthesized and tested for its hemolytic efficacy. Briefly, 4-hydroxybenzaldehyde was reacted with 1,8-dibromooctane in the presence of K_2CO_3 in acetone to yield compound **1**. This was further reacted with HO-PEG350Da-OMe which was pretreated with KHMDS in toluene to yield the final product **Ald-C8-PEG**.



Scheme 2: Ald-C8-PEG was synthesized from 4-hydroxybenzaldehyde in two steps (see Appendix for protocol.

The hemolytic efficacy of Ald-C8-PEG was hypothesized to be independent of solution pH. In order to confirm this, Ald-C8-PEG in increasing amounts was added to a 2% RBC solution (pH 5.5 or pH 7.4). The mixtures were allowed to incubate at 37°C for 60 minutes, followed by centrifugation to separate intact blood cells with supernatant. The supernatant was then measured for absorbance at 541nm, which corresponds with the absorbance of hemoglobin. The hemolytic efficacies of Ald-C8-PEG at pH 5.5 and 7.4 were compared. As shown in Figure 4 below, there is no observed pH dependence for the free surfactant Ald-C8-PEG. This confirms that the hemolytic efficacies for C8-C16 are due to their acid-dependent release of free surfactant.

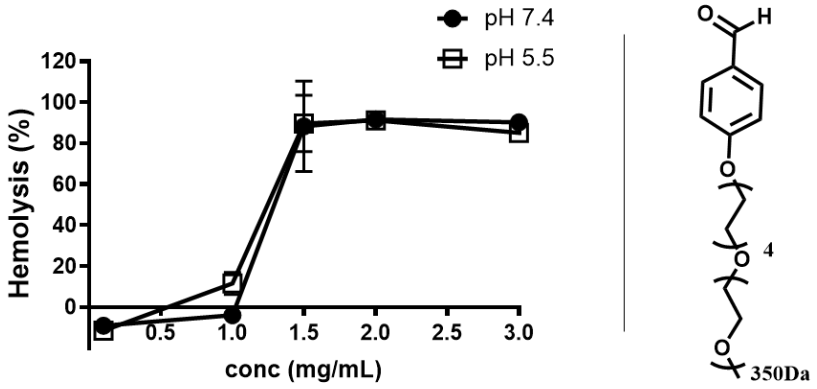


Figure 4: Ald-C8-PEG shows pH-independent hemolytic properties.

Conclusion

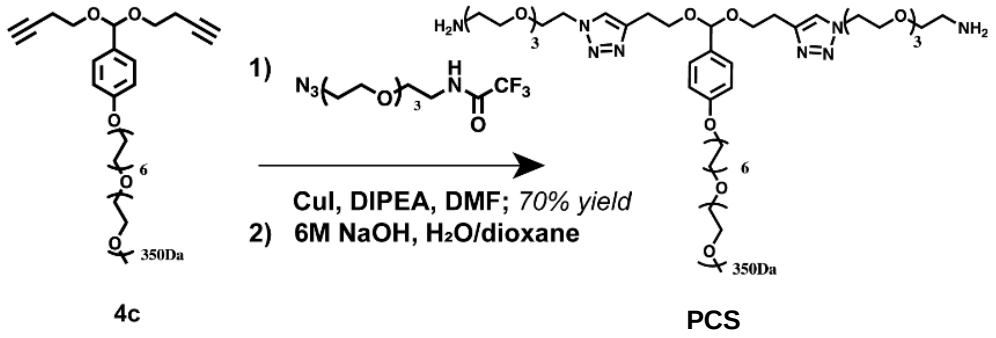
The results of the structure-activity relationship suggest that CS12 is the optimal structure for endosomal disruption. The hemolytic efficiency reaches a maximum at CS12, with similar efficiencies for CS14 and CS16. However, the hydrolysis half-life increases steadily as the hydrophobic core is extended. This factor makes CS12 preferable over CS14 and CS16, as rapid hydrolysis at endosomal conditions is necessary to lyse the endosomal compartment before the cargo is degraded in lysosomes.

1.3 Amine Caged Surfactants

The caged surfactants need to complex nucleic acids in order to enhance their cytoplasmic delivery. We consequently synthesized a 2nd generation caged surfactant, which contained two primary amines conjugated to the distal ends of the (PEG)2- chains, termed positively charged caged surfactants (PCS). The chemical structure of the PCS is shown in Scheme 3. PCS is designed to bind nucleic acids via a combination of electrostatic forces and hydrophobic forces, similar to cationic surfactants. However, after endocytosis the PCS will hydrolyze, generating individual positive charges and hydrophobic domains, which independently will not bind nucleic acids tightly, leading to efficient release of nucleic acids

Synthesis

The diamine caged surfactant (PCS) was synthesized from compound **4c** as shown in Scheme 3. Briefly, compound **4c** was reacted with N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2,2,2-trifluoroacetamide (See Appendix A) in the presence of CuI, DIPEA and DMF, to yield CS12-DITFAA in 70% yield. **PCS** was then formed *in situ* before use by addition of 6M NaOH (aq) in 1,4-dioxane. After 2 hours, the solution was added phosphate buffer (233.7µL, 0.2M at pH 8.3) to give a 10mg/mL stock solution. Deprotection was confirmed by TLC and HRMS, and stored at -80°C.



Scheme 3: Synthetic scheme for the preparation of PCS.

Nucleic acid binding efficacy

The caged surfactants need to bind nucleic acids to function as efficient delivery vehicles. Two positive charges were introduced into the caged surfactants at the distal ends of their (PEG)2-chains to promote nucleic acid binding. The ability of PCS to bind nucleic acids was tested by mixing with mRNA and plasmid DNA (pDNA) at various ratios and measuring retention on an agarose gel. Figure 5A and Figure 27-28 (Appendix A) show that PCS does indeed bind both mRNA and pDNA, and retained 50% at N/P=11 and N/P=10 respectively. Subsequent release is also vital, as nucleic acids which are bound tightly to positively charged polymers have been shown to be less transcriptionally and translationally active than their free analogs. Therefore, we

investigated whether PCS releases nucleic acids at low pH due to separation of the two positive charges from the surfactant domain. Figure 5B shows that acid hydrolysis of PCS leads to pDNA release even at an N/P ratio of 111. These results suggest that PCS can release complexed nucleic acids in endosomes following acid hydrolysis, allowing for efficient transcription and translation.

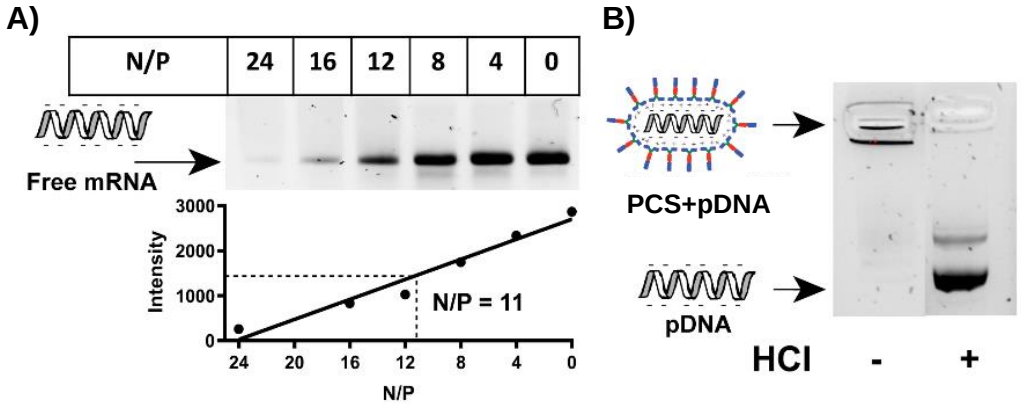


Figure 5: A) PCS binds mRNA: PCS was mixed with mRNA at various ratios and run on an agarose gel. 50% retention is reached at roughly N/P=11. B) Analysis of the pH dependent binding of PCS with plasmid DNA. Plasmid DNA was used for these studies because mRNA was unstable at acidic pHs. PCS + plasmid DNA (N/P=111) does migrate into the gel because of complexation. Addition of HCl causes hydrolysis of the acetal linker in PCS, separating the two positive charges from the hydrophobic core, which causes release of plasmid DNA.

Comparison of transfection agents

Three transfection agents were investigated for their ability to transfect Hela cells. Transfections were performed by mixing together mRNA (0.5µg/mL final concentration), PCS (2µg/mL final concentration) and X-tremeGene™ 360 (low: 0.2µL/mL), TurboFect™ (low: 0.3µL/mL), or Lipofectamine 2000 (low: 0.2 µL/mL). The complexes were added to Hela cells and incubated overnight in OptiMEM. This was compared to mRNA delivered at optimal delivery concentrations for X-tremeGene™ 360 (high: 2µL/mL), TurboFect™ (high: 3µL/mL), and Lipofectamine 2000 (high: 2 µL/mL). Figure 6 shows that neither X-tremeGene™ 360 or TurboFect™ achieved transfection levels comparable to that of Lipofectamine 2000, which saw roughly 80% transfection at optimal conditions. X-tremeGene™ 360 was the only control transfection reagent with high levels of transfection, which gave roughly 30% transfection, while low levels of transfection agent with PCS only gave around 10% transfection. For TurboFect™ these numbers were 40% for mRNA/transfection agent and 30% for mRNA/TurboFect™/PCS.

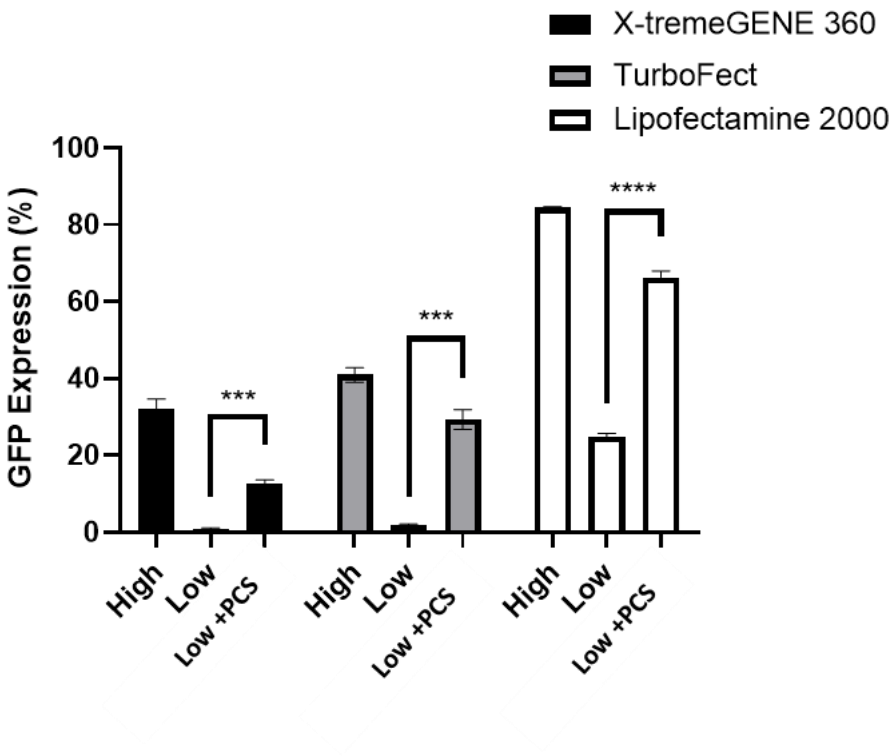


*Figure 6: PCS increases the transfection rate of low levels of both X-tremeGene™ 360 (0.2µL/mL), TurboFect™ (0.3µL/mL) and Lipofectamine 2000 (0.2µL/mL) respectively. High levels of X-tremeGene™ 360, TurboFect™, and Lipofectamine 2000 are 2µL/mL 3µL/mL and 2µL/mL respectively. ***P<0.001, ****P<0.0001 (unpaired two-tailed t-test).*

Size Determination (DLS)

DLS measurements were done by mixing 50 µL (50 µg/mL in PBS) of nucleic acid (mRNA or pDNA) with 7.5µL of PCS (75 µg in PBS). The mixture was incubated for 30 minutes before adding 1 µL of Lipofectamine 2000. For the measurements of mRNA and pDNA alone, as well as their complexes with PCS without Lipofectamine, the missing components were replaced with PBS. The measurements were done in triplicate.

The data in Figure 7 shows that Lipofectamine is vital to the tight packing of mRNA and pDNA in the presence of PCS, with average sizes of 77nm (mRNA) and 92nm (pDNA) for Nucleic acid/PCS/Lipo complexes, compared to 2023nm (mRNA) and 2157nm (pDNA) for Nucleic acid/PCS complexes. In addition, PCS also provides tighter packing when added to Nucleic acid/Lipo complexes, with number average sizes of 163 nm for mRNA/Lipo, and 77 nm for mRNA/PCS/Lipo.

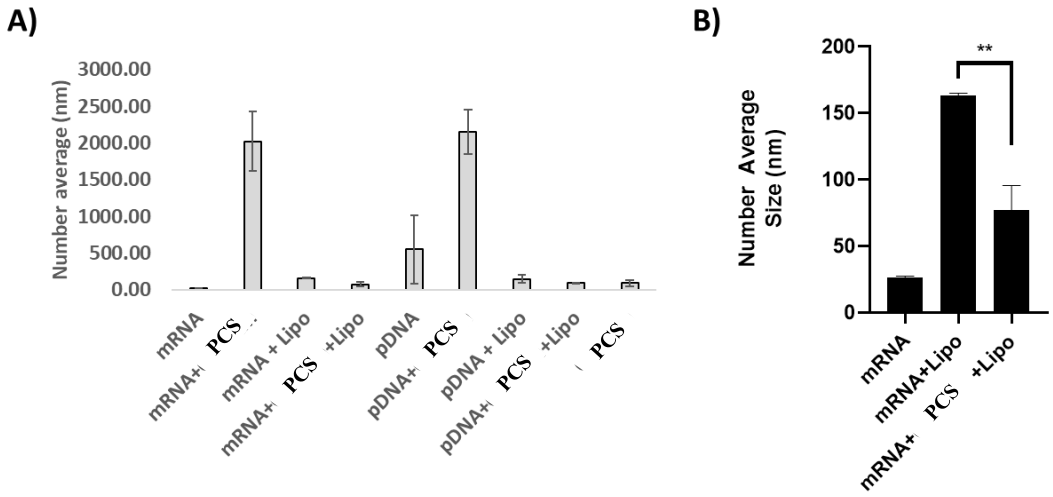


*Figure 7: DLS measurements of mRNA, mRNA+PCS, and mRNA+PCS in the presence of Lipofectamine. A) The data shows a clear dependence of Lipofectamine to achieve efficient packing of nucleic acids, with average sizes of 77nm (mRNA) and 92nm (pDNA) for Nucleic acid/PCS/Lipo complexes, compared to 2023nm (mRNA) and 2157nm (pDNA) for Nucleic acid/PCS complexes. B) Same data as A, but highlighting the reduction of size upon addition of PCS compared to mRNA+Lipofectamine alone. **P<0.01 (unpaired two-tailed t-test).*

mRNA delivery

PCS was then investigated for its ability to enhance the ability of Lipofectamine to deliver eGFP-expressing mRNA into HUVECs (Figure 8A). Lipofectamine 2000 was selected as a cationic lipid for transfection studies because it is one of the most widely used transfection agents, and in our studies greatly outperformed the commercially available transfection agents X-tremeGENE™ 360 and TurboFect™ (Figure 6). In addition, transfection reagents such as Lipofectamine were found to be vital to PCS-mediated delivery, as no eGFP expression was observed upon treatment of cells with mRNA/PCS alone (data not shown). However, despite its widespread use Lipofectamine has issues with toxicity[22] as well as low *in vivo* efficiency[23], and consequently, adjuvants that enhance its transfection efficiency are needed. Lipofectamine consists of a mixture of two positively charged lipids that disrupts endosomes through membrane fusion or the proton sponge effect, following endosomal entrapment[24,25]. The caged surfactants, however, disrupt endosomes through membrane insertion and destabilization, and should complement Lipofectamine's ability to disrupt endosomes.

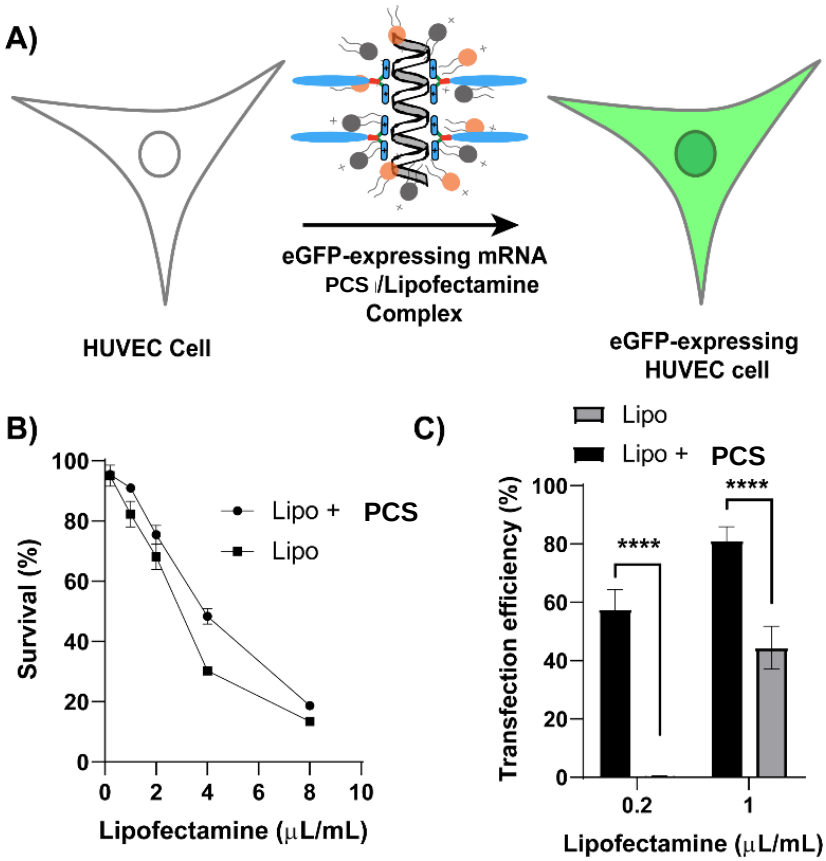


Figure 8: PCS increases the mRNA transfection efficiency of Lipofectamine in HUVECs. A) Summary of experiment. eGFP-expressing mRNA (0.5 µg/mL) was mixed with PCS (2.0 µg/mL), and incubated for 30 minutes followed by addition of Lipofectamine 2000. The resulting complex was added to HUVEC cells, and the expression of eGFP was measured by flow cytometry. B) Toxicity data for Lipofectamine alone as well as Lipofectamine/PCS. PCS dose not increase the toxicity of Lipofectamine. C) PCS increases the transfection efficiency of Lipofectamine in HUVEC cells. Lipofectamine/PCS had an 81% transfection efficiency at 1µg/mL of Lipofectamine, whereas the maximum transfection efficiency of Lipofectamine alone was 51%. In all cases the concentrations of PCS (2 µg/mL) and mRNA (500 ng/mL) were kept constant.

We selected primary cells as a testbed for these studies because they are challenging to transfect due to their sensitivity to cationic reagents, and because of their biomedical importance. Primary cells are the basis of a wide number of cell therapies and are also being intensely investigated as platforms for screening drugs. There is consequently great interest in being able to transfect primary cells with high efficiencies. HUVECs, like most primary cells are challenging to transfect, in part due to their sensitivity to cationic reagents and the sequestration of endocytosed material in lysosomes. For example, HUVEC transfection with cationic reagents rarely exceeds 45%, and frequently causes high levels of toxicity[26]. Consequently, an endosomal disruptive agent that can increase the efficiency of lipofectamine has the potential to impact multiple areas of biotechnology.

We therefore performed experiments to determine if PCS could enhance the ability of lipofectamine to deliver mRNA to HUVECs. PCS was added to eGFP-expressing mRNA and allowed to complex for 30 minutes. Lipofectamine was then added to the mixture, which was subsequently added to HUVEC cells and allowed to incubate at 37°C for 24 hrs. Finally, the

percentage of eGFP-positive cells was determined by flow cytometry (see Appendix A). In addition, we also determined the toxicity of Lipofectamine and Lipofectamine/PCS (Figure 8B).

Figures 8C demonstrates that PCS can increase the transfection efficiency of Lipofectamine, while also lowering its toxicity by minimizing the amount of Lipofectamine that needs to be added to cells. Lipofectamine's CC50 is between 2-4µL/mL (Figure 8B), which is problematic because at concentrations below 2µL/mL it is inefficient at transfecting cells (Figure 8C). At 2 µL/mL Lipofectamine does have a 51% transfection rate, but also causes significant cell death, killing approximately 32% of HUVECs. In contrast, 2 µg/mL of PCS combined with 1 µL/mL of Lipofectamine generated a transfection efficiency of 81%. In addition, PCS/Lipofectamine caused only 9% toxicity at this dose, as opposed to 32% toxicity for Lipofectamine at 2 uL/mL. These results were further verified by a separate transfection study, using fluorescent microscopy as the read-out (see Figure 37 in Appendix A), which showed a clear increase in the number of transfected cells in the PCS/Lipofectamine case.

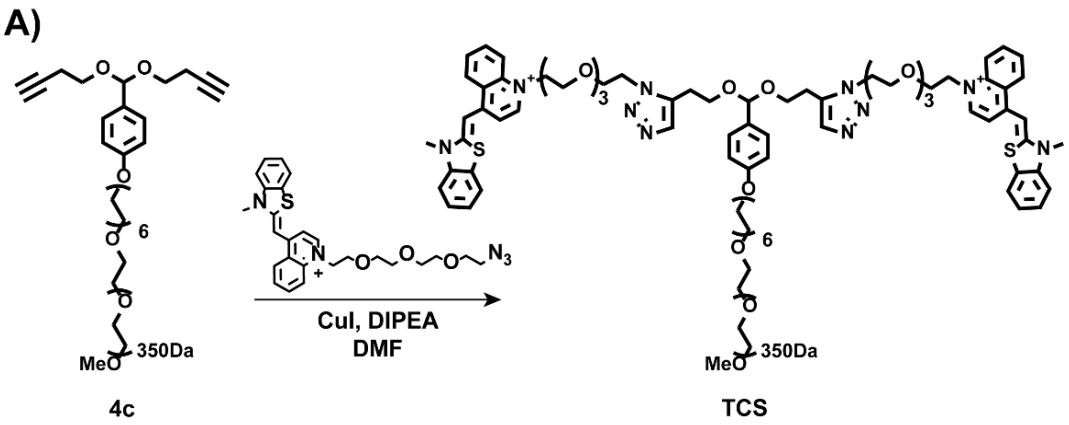
1.4 Thiazole Orange Caged Surfactants

Introduction

A key limitation of PCS is its relatively weak affinity for nucleic acids, which will potentially limit their applications. We therefore synthesized a 3rd generation caged surfactant, which had the RNA binding dye thiazole orange conjugated to the distal ends of its PEG chains, termed the Thiazole Orange Caged Surfactant (TCS). TCS should bind nucleic acids with a higher affinity than PSC because thiazole orange binds nucleic acids with a K_d in the micromolar range[27]. Using nucleic acid-binding dyes for delivery has previously been developed by attaching acridine on a PEG-peptide backbone[28–30]. This creates tight binding through multivalency; however, due to the lack of an acid-sensitive linker, the absence of efficient nucleic acid release could pose a problem as bound nucleic acid has been shown to be less transcriptionally and translationally active.[3] We therefore foresee that a dimeric TO with an acid-sensitive linker can prove a viable alternative.

Synthesis

The Thiazole Orange caged surfactant (TCS) was synthesized from compound 4c as shown in Scheme 4. Briefly, compound **4c** was reacted with TO-PEG4-N3 (See Appendix A for synthesis) in the presence of CuI, DIPEA in DMF to yield TCS in 38% yield after purification by neutral alumina flash chromatography. The compound was stored in the presence of TEA. In the case of preparation of DMSO stocks, final TCS concentration was standardized to TO concentration by absorbance at 514nm, using TO as a standard.



Scheme 4: Synthetic scheme for the preparation of TCS.

siRNA Binding Characterization

TCS was then characterized for its ability to bind siRNA. This was assessed using a TCS/siRNA titration curve, as well as DLS for inspecting particle size. The siRNA binding assay was performed by gradually adding increasing amounts of TCS (25µM in DMSO) to 100µL siRNA (1µM). The mixture was then allowed to equilibrate for 60 minutes before measuring fluorescent intensity (490/535nm), which increases upon TO/RNA intercalation. The results shown in Figure 9A reveals that TCS does in fact intercalate with siRNA, and reaches saturation at around two TCS per siRNA molecule. This corresponds to four TO molecules per siRNA, which in this case has 19 base pairs (with a two-nucleotide overhang). This roughly corresponds to the saturation of TO, which was found to reach saturation at four TO per siRNA (results not shown).

Dynamic light scattering (DLS) was performed by mixing siRNA (50µL, 10µM in DEPC water) with TCS (1µL, 0.5mg/mL in DMSO). The number average sizes were measured in triplicate. As shown in Figure 9B, siRNA/TCS has an average number size of 55nm in diameter, whereas siRNA by itself has an average number size of 2.5nm. Interestingly, TCS by itself has an average size of 270nm, indicating that it aggregates in some form when not in the presence of siRNA.

These results combined shows that TCS does bind siRNA with an efficiency comparable to that of TO, and forms particles of small size, appropriate for those needed for efficient cell uptake. This is in direct contrast for the complexes formed by PCS, which formed particles of <1000nm in size (see Figure 7).

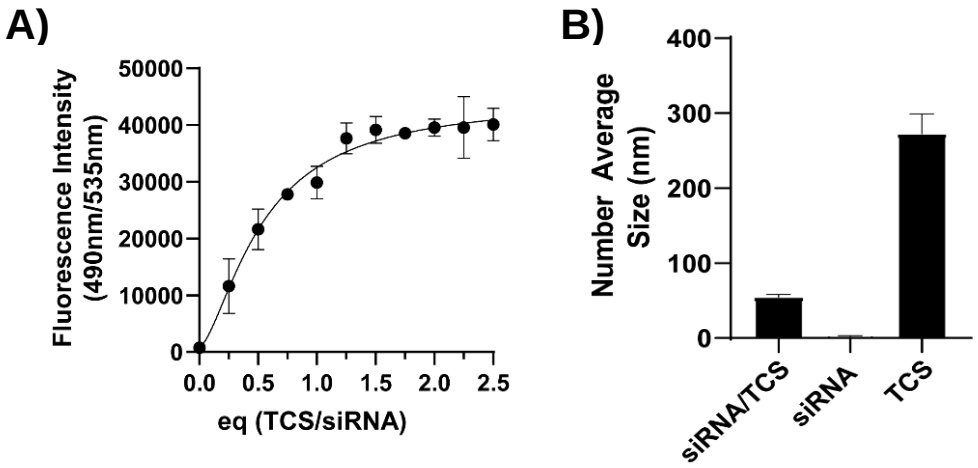


*Figure 9: **A)** Binding of TCS to siRNA causes an increase in fluorescence. TCS was mixed with siRNA at various mole ratios and the increase in fluorescence was measured. A 23 base siRNA can bind 2 TCS molecules. **B)** TCS forms 55 nm sized particles with siRNA. The particle size was determined via dynamic light scattering.*

Thiazole Orange Caged Surfactant-Mediated siRNA Delivery

The ability of TCS to enhance the delivery of siRNA targeting the luciferase gene was investigated in luciferase expressing Hela cells (Figure 10A). Luciferase siRNA (sense strand; GAU UAU GUC CGG UUA UGU AUU) was complexed with TCS and was then incubated with Luciferase-expressing Hela cells in serum-containing media for 24 hours followed by addition of luciferin and measurement of chemiluminescence. Figure 10B demonstrates that TCS can enhance the delivery of siRNA into cells. For example, at a 25 nM concentration of siRNA, siRNA by itself caused only a 18% knockdown of luciferase, whereas siRNA complexed to TCS caused a 78% inhibition of luciferase. Similar enhancements in siRNA delivery with TCS were also seen at a 100 nM concentration of siRNA.

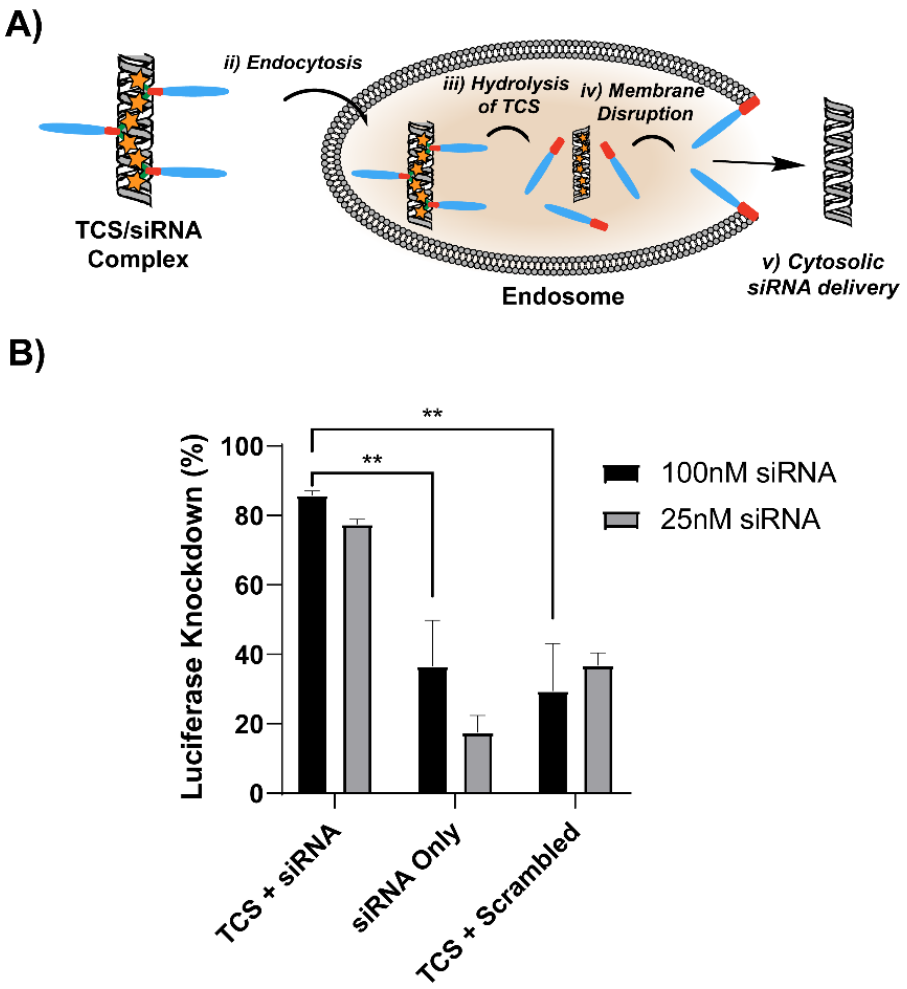


*Figure 10: **TCS** can enhance the delivery of siRNA into cells. **A)** Schematic showing intracellular delivery of siRNA by TCS. TCS binds siRNA and hydrolyses in the endosome, triggering endosomal release. **B)** TCS can enhance the delivery of siRNA into cells. TCS+siRNA (100 nM) causes 86% luciferase inhibition, whereas siRNA by itself caused only 37% inhibition. At 20nM siRNA, these values were 78% and 18%, respectively. **P<0.01 (unpaired two-tailed t-test).*

1.5 Conclusion

In this chapter we demonstrated that Triton X-like surfactants can be converted into pH sensitive membrane disruptive agents via reversible PEGylation. Interestingly, short PEG chains, 4 units in length were able to prevent large hydrophobic domains, as large as 23 carbons in length (hexadecane and a benzaldehyde) from disrupting membranes. This result suggests that reversible PEGylation will be able to enable the design and synthesis of a wide variety of new small molecule endosomal disruptive agents. A key benefit of low molecular weight endosomal disruptive agents is their well-defined chemical structure, which makes them easily amenable to chemical optimization. In this report we synthesized 2nd and 3rd generation caged surfactants, which enhanced the ability of lipofectamine to deliver mRNA and enhanced the delivery of siRNA, respectively.

Appendix A – Chapter 1

A.1 Methods and Materials

All compounds were purchased from Sigma Aldrich (St. Louis, MO, USA), Fisher Scientific (Hampton, NH, USA), or VWR (Radnor, PA, USA). THF was purchased from Fisher Scientific (Hampton, NH, USA) and distilled from solid sodium and benzophenone before using in acetal formation reactions. siSTABLE Luc-knockdown siRNA (sense strand; GAU UAU GUC CGG UUA UGU AUU) was purchased from Horizon Discovery Biosciences Limited (Cambridge, UK). Compound purity was assessed by ^{1}H-NMR spectroscopy. HPLC was not utilized due to the instability of acetal-containing compounds in standard 0.1% TFA HPLC mobile phases. NMR spectra were collected on either a Bruker Avance 400 console with Oxford Instruments 9.4 T magnet, a Bruker Avance 500 console with Bruker 11.7 T magnet, or a Bruker Avance III 600 console with Bruker 14 T magnet instrument at the College of Chemistry NMR Facility at UC Berkeley. Additionally, low-mass ^{13}C NMR spectra were collected on a Bruker 900 MHz NMR spectrometer. Aldehydes and acetals were visualized on TLC with a 2,4-Dinitrophenylhydrazine solution. Mass spectra were collected at the QB3/Chemistry Mass Spectrometry Facility at UC Berkeley. Absorbance was measured on a Tecan infinite M200 plate reader (Männedorf, Switzerland). Fluorescent Microscopy images were taken on an Invitrogen Evos M5000 instrument (Carlsbad, CA, USA). UV Vis spectra were collected on a Shimadzu UV-2600 Spectrophotometer with a Shimadzu TCC-100 Temperature Controlled Cell Holder. Flow cytometry was collected on an Invitrogen Attune NxT flow cytometer. Gels were imaged using a Bio-Rad ChemiDoc MP Imaging System using 1:1000 SYBR^{TM} Safe (Thermo Fischer Scientific) as the intercalating imaging agent. Dynamic light scattering (DLS) measurements were done on a Zetasizer Nano ZS with backscatter detection (Malvern Instruments). CleanCap® eGFP mRNA was purchased from Trilink Biotechnologies, (San Diego, CA).

A.2 Hemolysis Protocol

Preparation of 2% RBC solution: Defibrinated red blood cells (Hemostat Laboratories, Dixon CA, USA) were diluted in PBS at the desired pH (1:10), and the suspension was centrifuged at 4°C (750 x g, 10 minutes), followed by removal of the supernatant. This was repeated three additional times, diluting with PBS (1:50) each time. After the last centrifugation, 200 µL of the collected red blood cells (RBC) were diluted with 9.8mL of PBS at the desired pH.

Hemolysis assay: Caged surfactants were dissolved in PBS at various concentrations and 100 µL was added to 200 µL centrifuge tubes, which were capped and allowed to stand for 1 hour at RT for hydrolysis to occur. Then, each tube was added 50 µL of the prepared 2% RBC solution for a total of 150 µL. The tubes were incubated for 30 minutes at 37°C, followed by centrifugation at 2,000 x g for 10 minutes. A spectrophotometer was used to measure the hemoglobin content of the supernatant at 546 nm. All hemolysis experiments were performed in technical triplicate.

Preparation of positive control: Positive control (100% hemolysis) was prepared by adding 50µL of the prepared RBC solution to 100µL of DI H_2O. The solution was then frozen and thawed three times, leading to complete disruption of the blood cells. This solution was used as the 100% positive control.

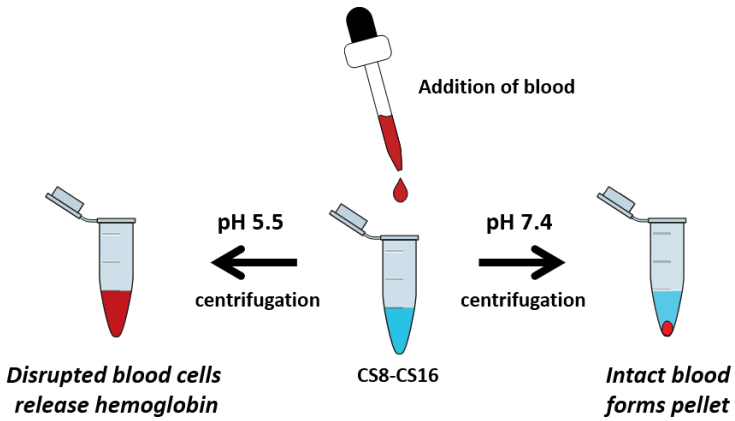


Figure 11: The hemolysis assays were done by adding a 2% RBC solution (pH 5.5 or 7.4) to a mixture of caged surfactant (CS8-CS16) pre-incubated in PBS for 60 minutes (pH 5.5 or 7.4). This mixture was then incubated for 30 min at 37°C, before centrifugation. The supernatant was measured for hemoglobin content by absorbance (546nm).

A.3 Kinetics of Hydrolysis

Time dependency of hemolysis

In order to achieve efficient endosomal disruption, the acetal linker in the caged surfactants needs to hydrolyze rapidly at endosomal conditions. However, the caged surfactants of longer carbon chains show slow hydrolysis kinetics. We therefore tested whether hydrolysis kinetics will limit the degree of hemolysis, since the hemolytic efficiency of the caged surfactants is dependent of the degree of acetal hydrolysis. **CS14** was incubated in PBS at pH 7.4 and 5.5 for 30 min and 60 min before adding to RBCs as explained in Section A.2. The hemolytic efficiency was measured after co-incubation with the RBCs at 37°C for 30 minutes. Figure 12 shows that the hemolytic efficiency is indeed time dependent, and suggests that the hydrolysis half-life of the caged surfactants will have a significant influence on their ability to disrupt endosomes.

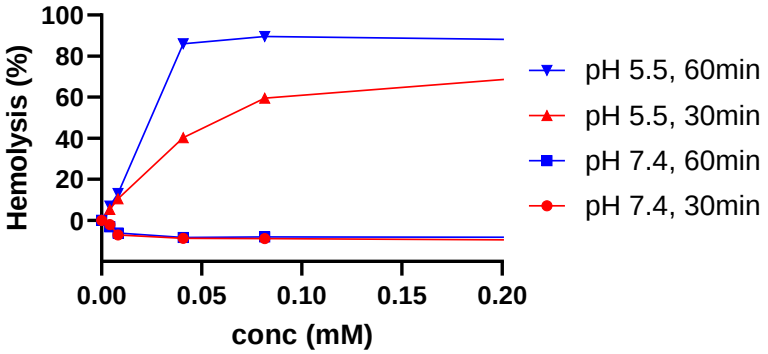


Figure 12: The hemolytic efficacy of CS14 is time dependent. Shown here is the hemolytic efficiency of CS14 after incubation in PBS (pH 7.4 and 5.5) for 30 and 60 minutes before measuring hemolysis.

CS8 Kinetics Data

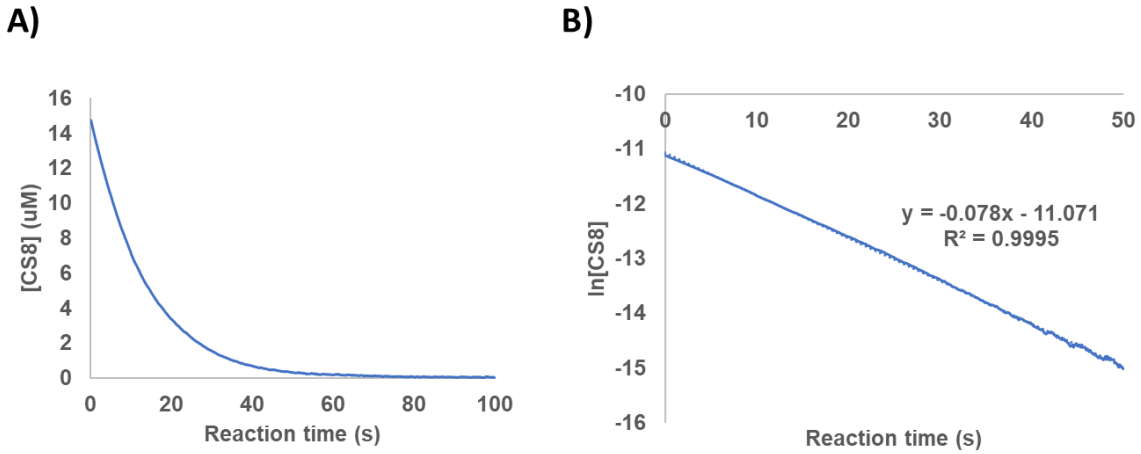


Figure 13: Hydrolysis kinetics of CS8 at pH 3.5.

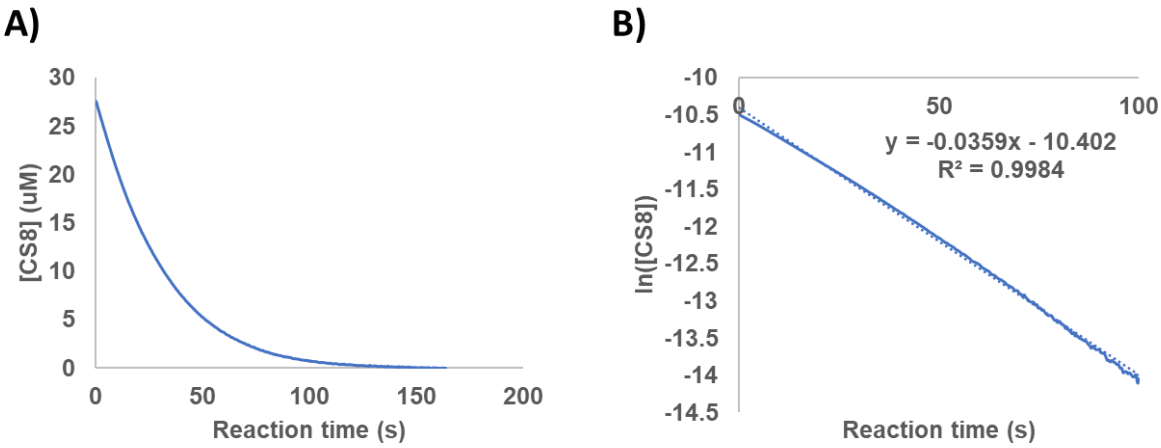


Figure 14: Hydrolysis kinetics of CS8 at pH 4.0.

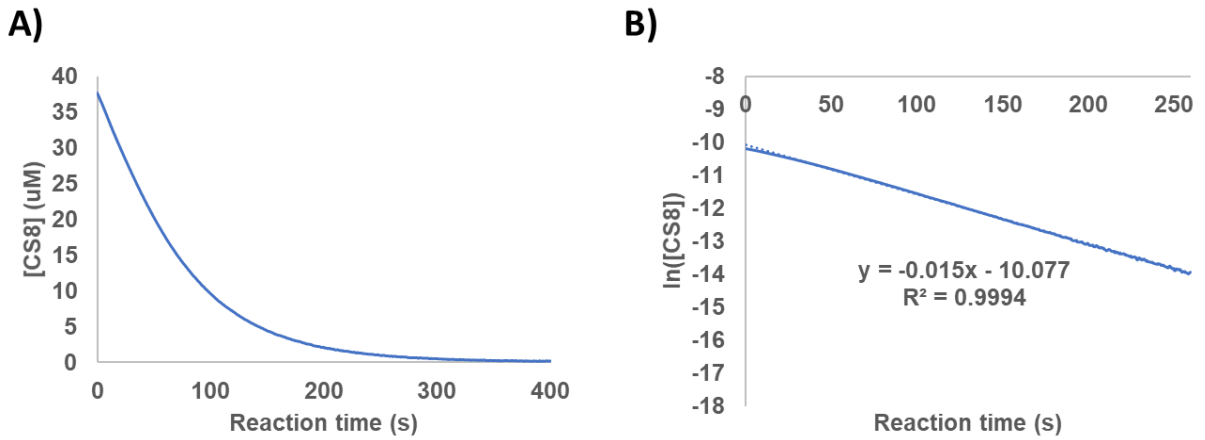


Figure 15: Hydrolysis kinetics of CS8 at pH 4.5.

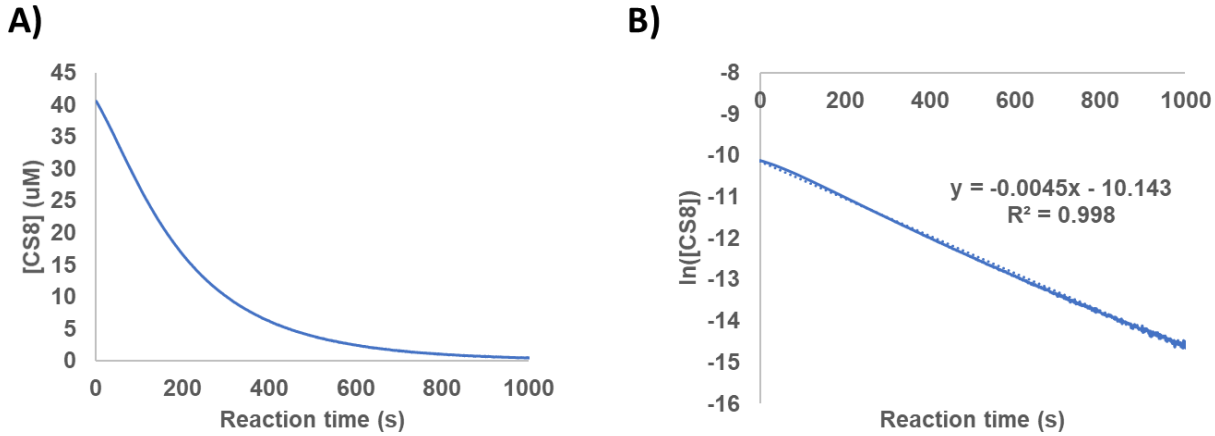


Figure 16: Hydrolysis kinetics of CS8 at pH 5.0.

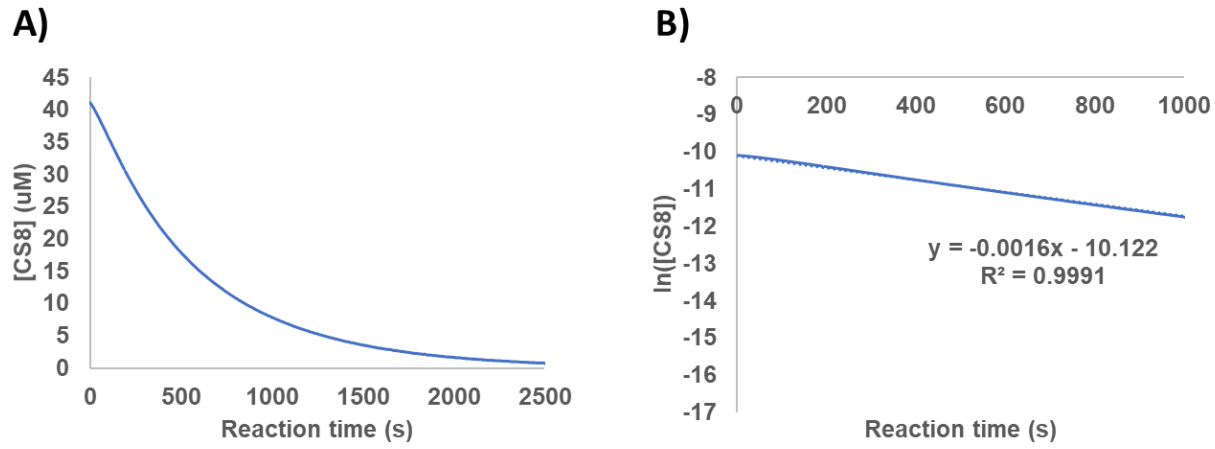


Figure 17: Hydrolysis kinetics of CS8 at pH 5.5.

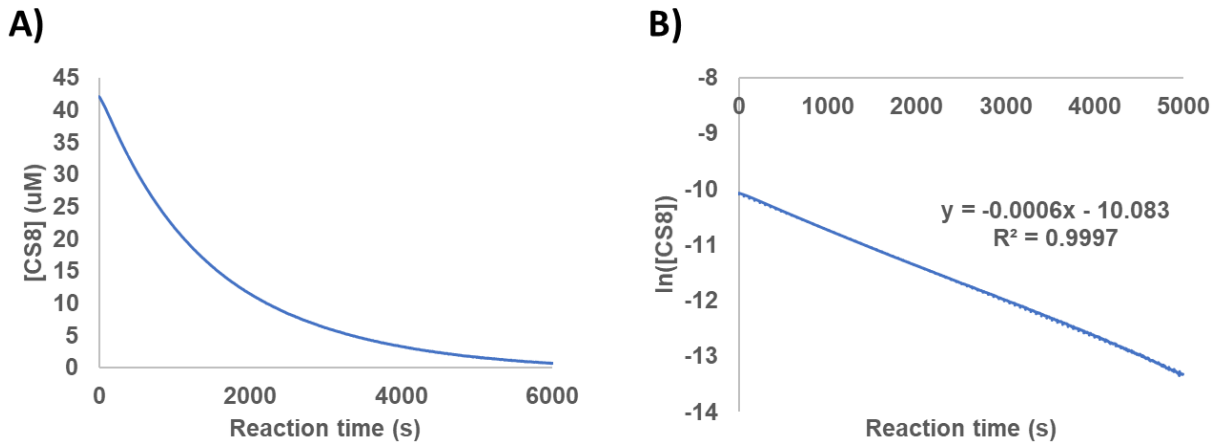


Figure 18: Hydrolysis kinetics of CS8 at pH 6.0.

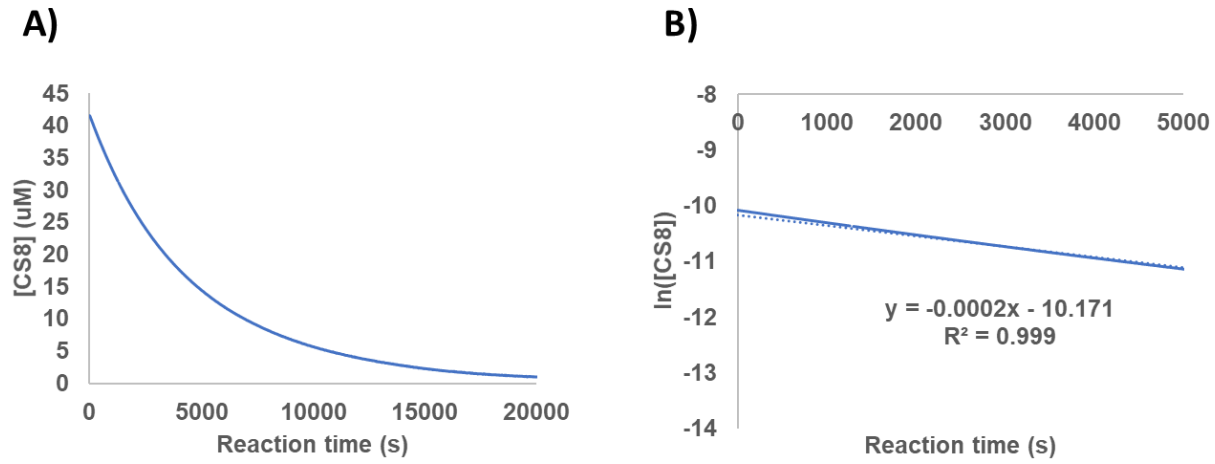


Figure 19: Hydrolysis kinetics of CS8 at pH 6.5.

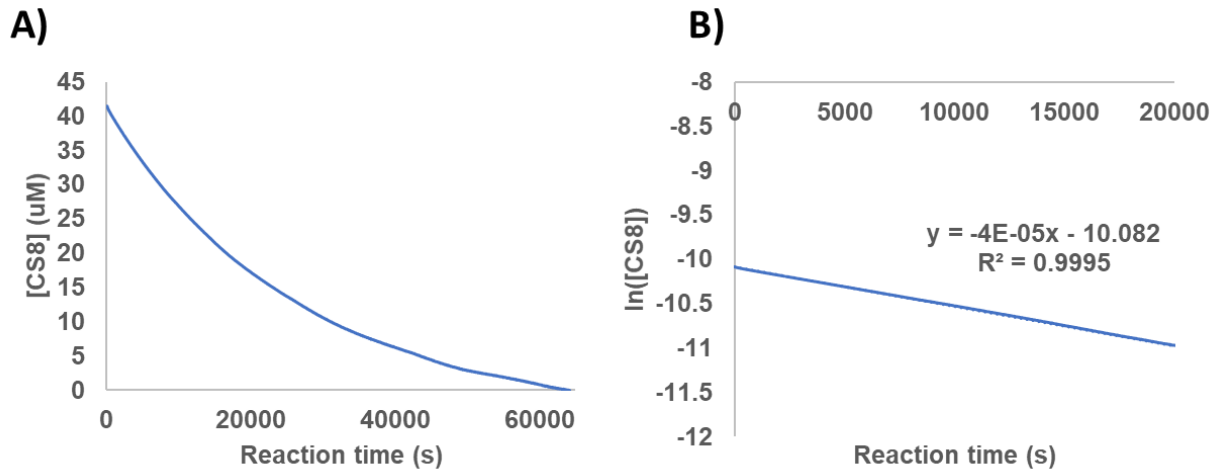


Figure 20: Hydrolysis kinetics of CS8 at pH 7.4.

CS12 Kinetics Data

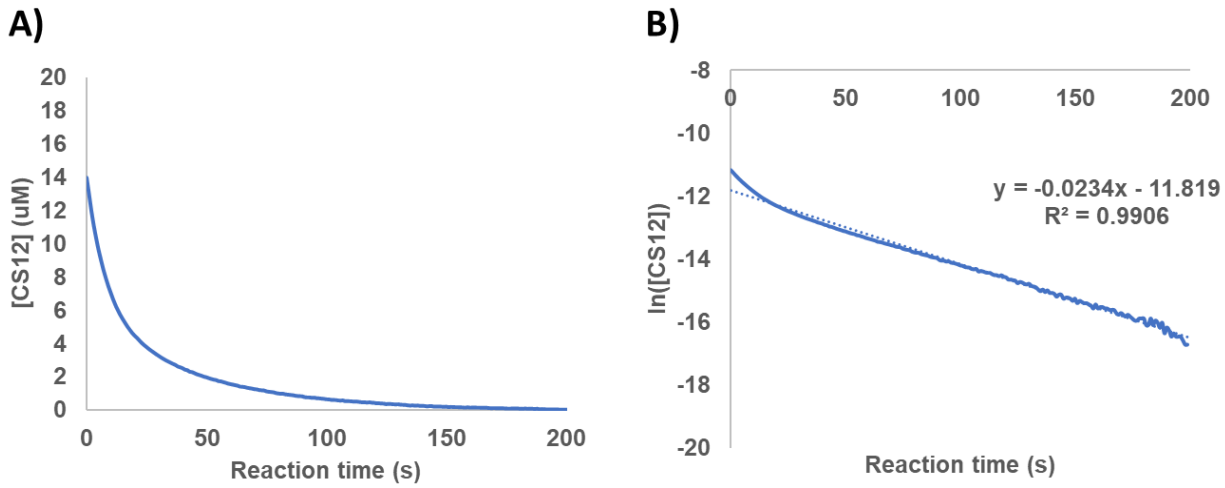


Figure 21: Hydrolysis kinetics of CS12 at pH 3.5

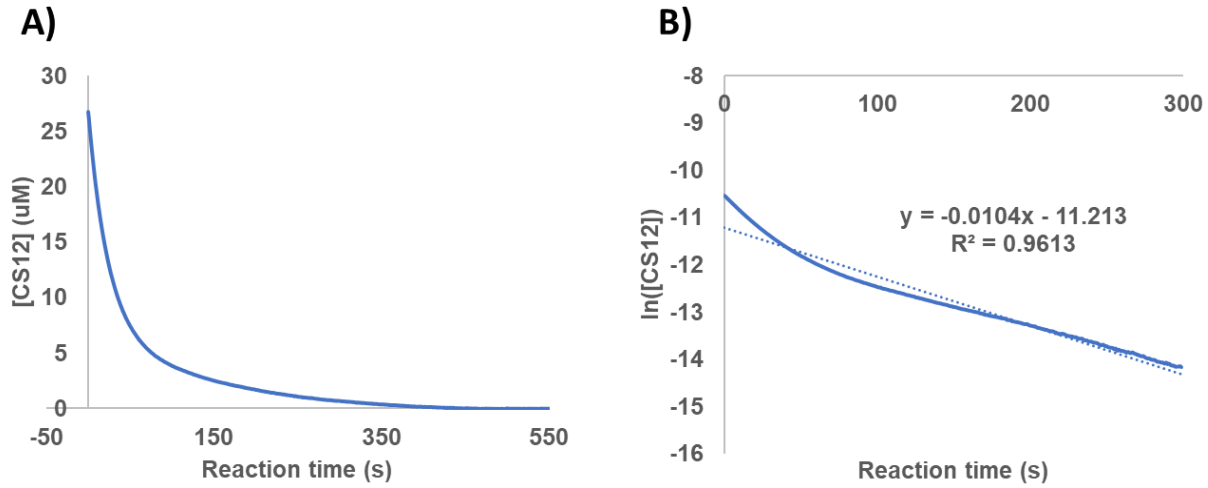


Figure 22: Hydrolysis kinetics of CS12 at pH 4.0.

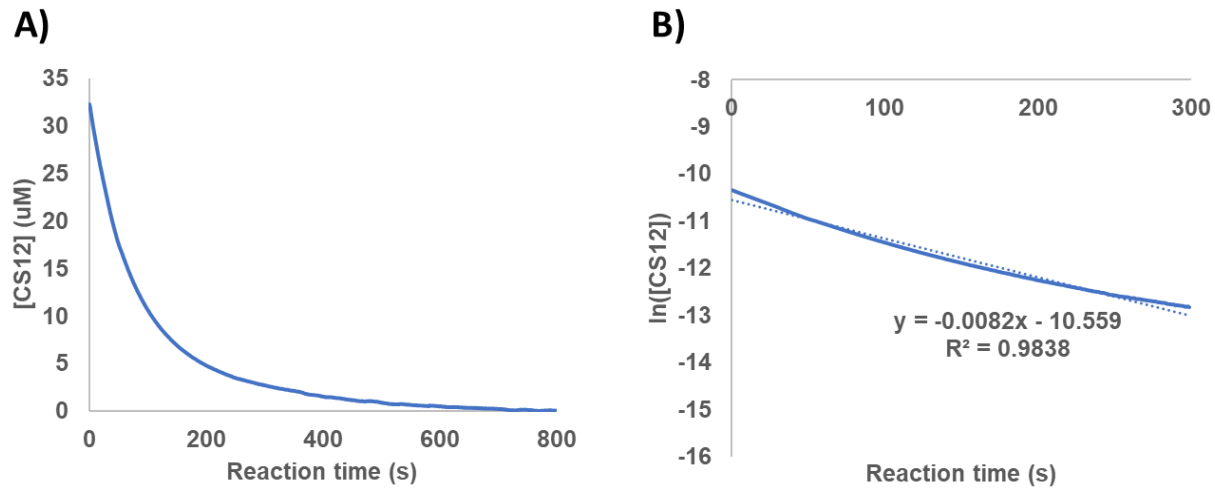


Figure 23: Hydrolysis kinetics of CS12 at pH 4.5.

CS16 Kinetics Data

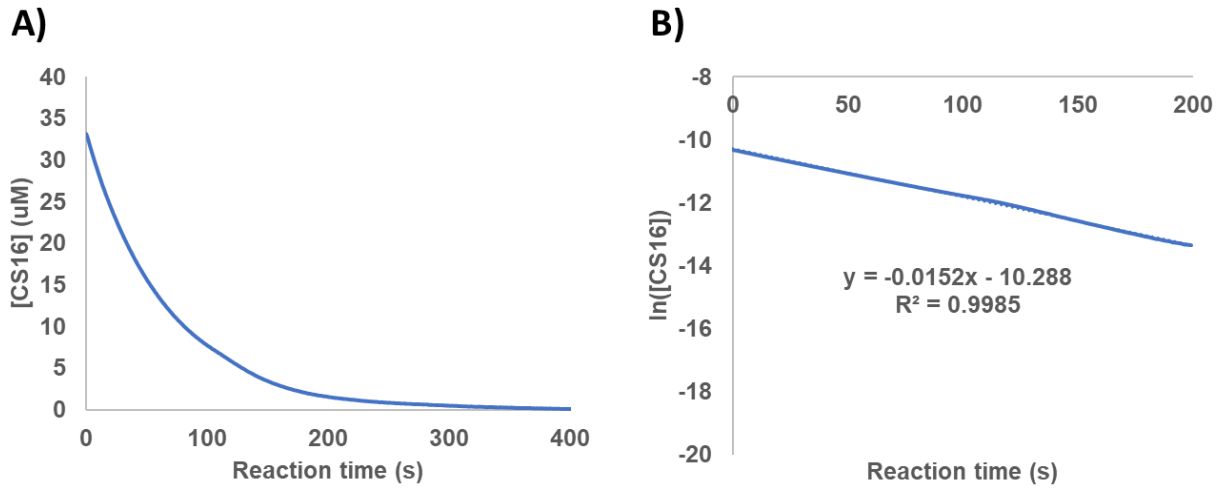


Figure 24: Hydrolysis kinetics of CS16 at pH 3.5.

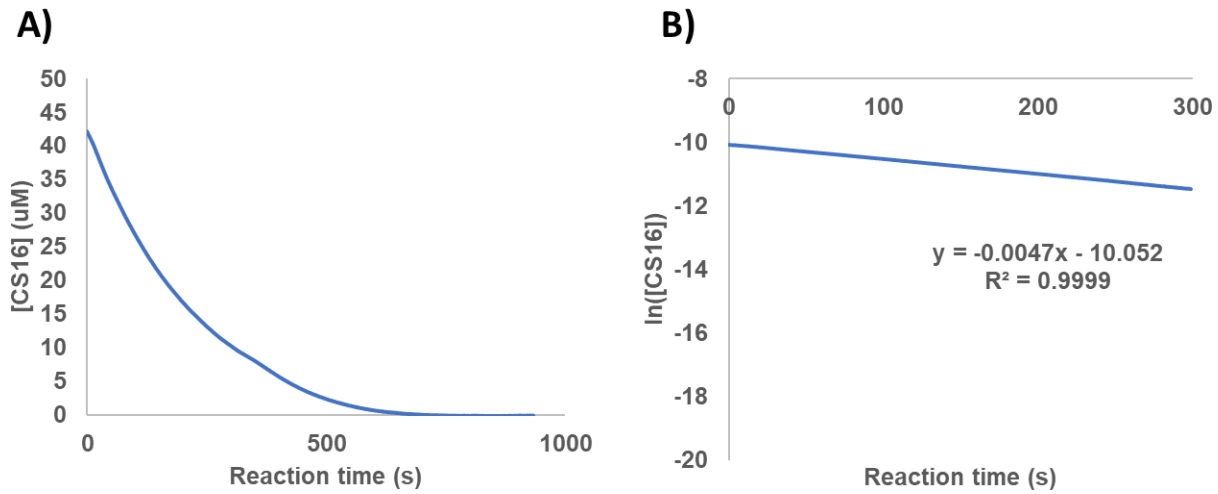


Figure 25: Hydrolysis kinetics of CS16 at pH 4.0.

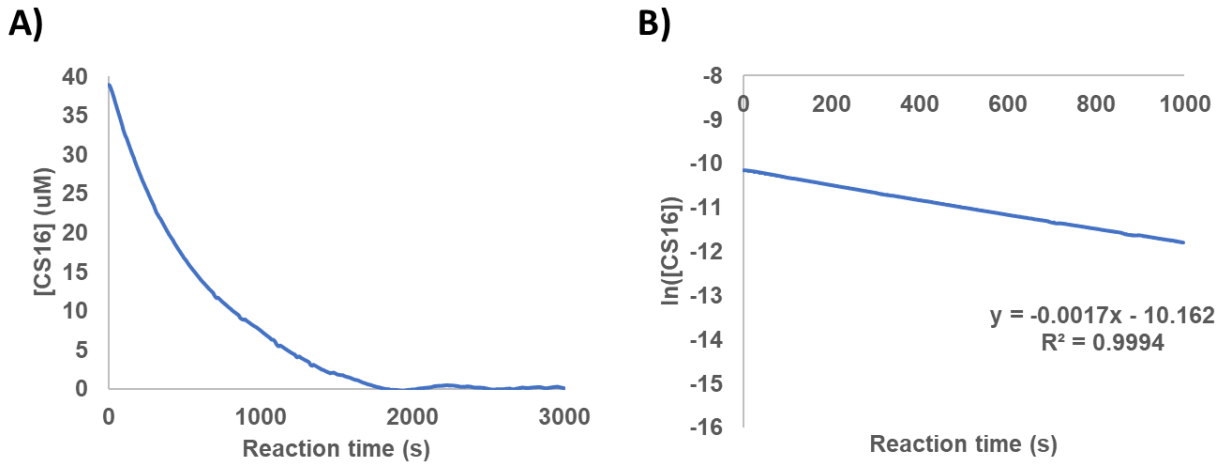


Figure 26: Hydrolysis kinetics of CS16 at pH 4.5.

A.4 Nucleic acid Retention Gels

mRNA retention by PCS was measured by adding PCS (6µL, various concentrations) to mRNA (eGFP encoding, 600ng in 2µL) followed by incubation for 1 hour before loading onto a 1% agarose gel. After running at 120V for 30 minutes, the gel was imaged using a Bio-Rad ChemiDoc MP Imaging System using 1:1000 SYBR™ Safe (Thermo Fischer Scientific) as the intercalating imaging agent.

pDNA retention was measured similarly to mRNA retention on an agarose gel (1%). eGFP-pDNA (250ng in 5µL) was added to 5uL of PCS at various concentrations. After incubating for 1 hour at room temperature, the gel was run at 120V, and retention was quantified based on the dominant band of free pDNA (Figure 28). The efficiency was shown to be 50% retention at roughly N/P=10, which is similar to that of mRNA with a 50% retention efficiency at N/P = 11. Full gels are shown in Figure 27 (mRNA) and Figure 28 (pDNA).

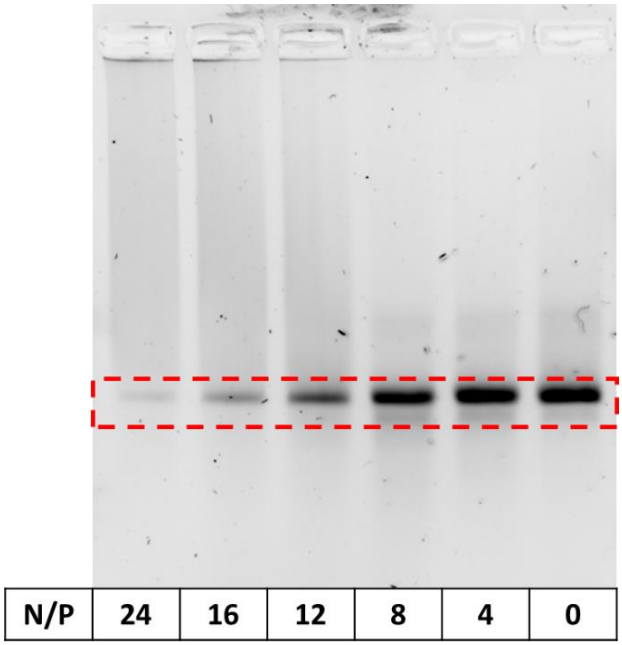


Figure 27: PCS retains mRNA on a gel. A) Agarose gel showing retention of mRNA upon addition of PCS. The ratio is given as N/P, referencing the amount of amine groups per phosphate group. 50% retention is achieved at roughly N/P = 11.

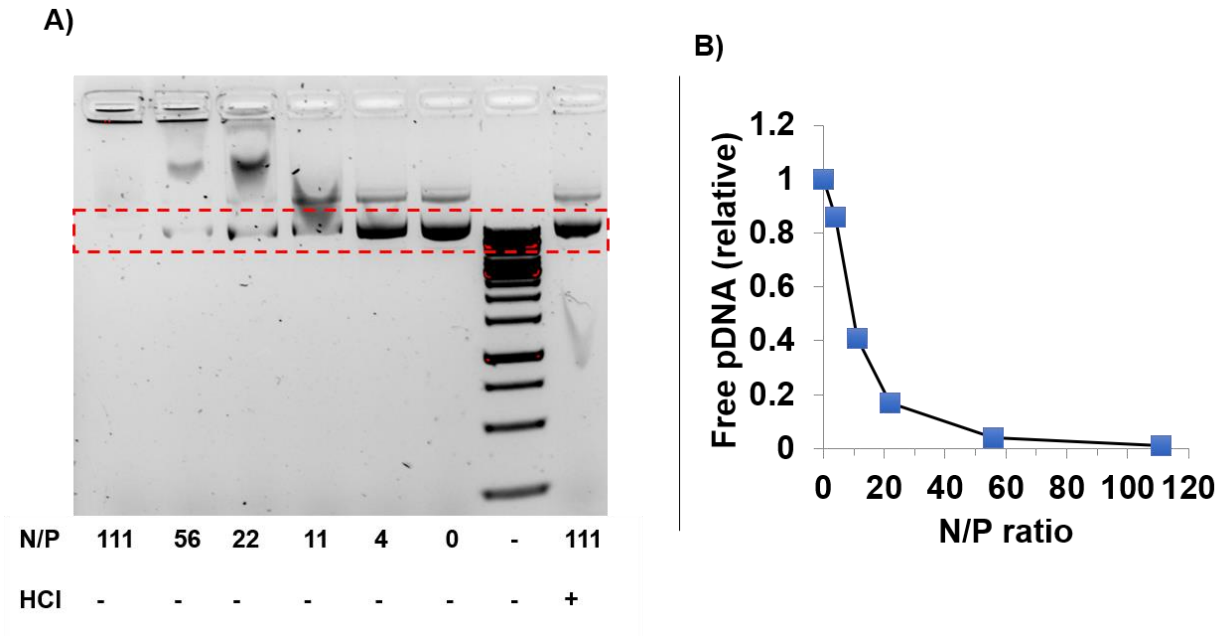


Figure 28: PCS retains pDNA on a gel. A) Agarose gel showing retention of pDNA. B) A plot showing degree of retention of pDNA. A retention efficiency of 50% is achieved around N/P=10. This is similar to the value found for mRNA of N/P=11.

A.5 PCS Transfection and Toxicity Protocols

HUVEC Transfection and Cell Toxicity Protocols

HUVEC cells were transfected by first adding PCS (2µg/mL) to CleanCap® eGFP mRNA (0.5µg/mL) (Trilink Biotechnologies, San Diego, CA), followed by incubation for 30 minutes at room temperature. Then, Lipofectamine 2000 was added at different amounts followed by an additional 30-minute incubation. The resulting mixture was then added to HUVEC cells in a 96-well plate in OptiMEM media. The cells were incubated for 24 hours before eGFP expression was measured by flow cytometry. The transfection was expressed as the percentage of positive cells/total cells. This also takes into consideration the toxicity that is caused by Lipofectamine, and is a better measurement of the total efficiency of the system.

Toxicity of mRNA/Lipo/PCS was measured on HUVEC cells by keeping the PCS concentration constant at 2µg/mL, while varying the Lipofectamine 2000 concentration. Toxicity was then determined as the number of dead cells as seen by flow cytometry. The protocol is the same as for transfection. Briefly, eGFP-expressing mRNA (0.5µg/mL) was mixed with PCS (2µg/mL) and incubated for 30 minutes, followed by the addition of Lipofectamine. After an additional 30-minute incubation, the complexes were added to HUVEC cells in a 96-well plate (100uL/well) in OptiMEM media. After a 24-hour incubation, flow cytometry was performed to determine cell toxicity.

Hela PCS Cell Toxicity

Hela cells were transfected by first adding PCS (2μg/mL) to CleanCap® eGFP mRNA (0.5μg/mL) (Trilink Biotechnologies, San Diego, CA), followed by incubation for 30 minutes at room temperature. Then, Lipofectamine 2000 was added at different amounts followed by an additional 30-minute incubation. The resulting mixture was then added to Hela cells in a 96-well plate in OptiMEM media. The cells were incubated for 24 hours before eGFP expression was measured by flow cytometry. The toxicity of PCS alone was measured on Hela cells using a simple Resazurin assay. As seen in Figure 29, this showed a CC50% of roughly 50 μg/mL, which was significantly higher than the concentrations used for mRNA transfections, which was 2 μg/mL.

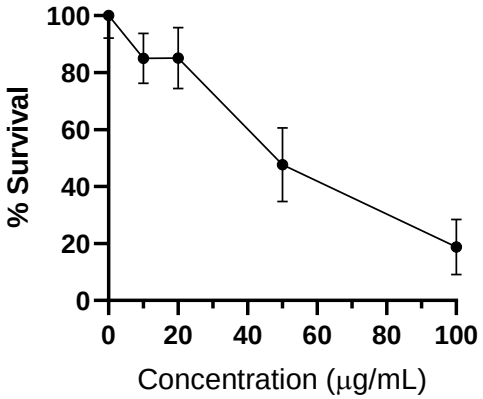


Figure 29: PCS shows toxicity on Hela cells with a CC50% of roughly 50ug/mL. This is above the concentration used for transfection, which was 2μg/mL for Hela cells.

Endocytosis Inhibitor Assay

In order to investigate the mechanism of delivery for the mRNA/PCS/Lipo complexes, Hela cells were transfected in the presence of Wortmannin (0.15μg/mL), Chlorpromazine (1.5μg/mL), Genistein (5μg/mL) and methyl-β-cyclodextrin (7.5mg/mL). The inhibitors were incubated with Hela cells for 1hr before replacing the media. The mRNA/PCS/Lipo complexes were prepared by making a mixture of 2μg/mL PCS, 500ng/mL mRNA, and 2μL/mL Lipofectamine 2000. The complex was incubated for 30 minutes before being added to cells and incubated overnight in OptiMEM. A control was done by using only Lipofectamine 2000 for delivery to see if there were any significant changes in delivery mechanism. The inhibitors used and their inhibiting pathways are as follows[31]: (A) Wortmannin - Clathrin-dependent endocytosis (CDE), (B) Chlorpromazine – CDE, (C) Genistein – Clathrin-independent endocytosis (CIE), (D) Methyl-β-cyclodextrin (MβCD) – Cholesterol-dependent uptake mechanisms (CDE and CIE at used concentrations[32]). The results below show that the uptake of the mRNA/PCS/Lipo-complexes happens by a similar mechanism to that of mRNA/Lipo complexes. The uptake is dependent on both Clathrin-dependent and Clathrin-independent endocytosis, as seen in the knock-down in the presence of chlorpromazine and genistein, respectively. This is in agreement with previously reported data.[33] In addition, MβCD caused a total knockdown in transfection efficiency, while maintaining cell

viability (data not shown). At high concentrations, MβCD can inhibit both Clathrin-dependent and independent pathways that are dependent on cholesterol, specifically those that involve lipid raft domains.[34] These results confirm that endocytosis plays an important part in the delivery of mRNA complexed with PCS and Lipofectamine.

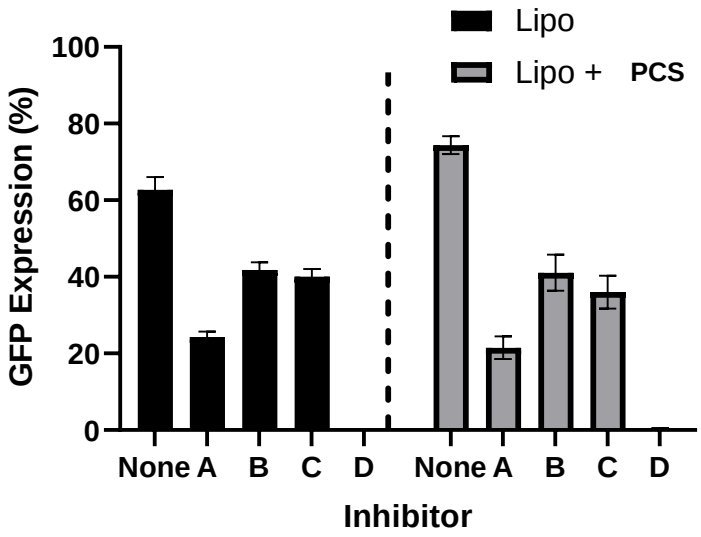


Figure 30: mRNA/PCS/Lipofectamine 2000 in endocytosed by a similar mechanism to Lipofectamine 2000-mediated delivery. The transfection efficiency is dependent on both Clathrin-dependent endocytosis (CDE) and Clathrin-independent endocytosis (CIE). The following inhibitors were incubated with Hela cells before transfection: (A) Wortmannin - (CDE), (B) Chlorpromazine – CDE, (C) Genistein – (CIE), (D) Methyl-β-cyclodextrin (MβCD) – Cholesterol-dependent uptake mechanisms (CDE and CIE at the concentrations used[32]).

A.6 TCS Characterization and Cell Work Protocols

siRNA Binding Constants

To 100uL of siRNA (1µM in DEPC-treated water) was added increasing amounts of TCS (25µM in DMSO) in 1µL intervals until saturation was observed. From the resulting graph binding constants were measured and calculated according to published procedures[35].

DLS measurements of siRNA/TCS

Average sizes were measured using dynamic light scattering (DLS). Luc-siRNA stocks were prepared in DEPC-treated water to a concentration of 10µM, and TCS was dissolved in DMSO to

a concentration of 0.5mg/mL. In a cuvette, 50µL of siRNA and 1uL of TCS was added and left to complex for 15 minutes before size measurement (n=3). siRNA alone was added 1µL of DMSO instead of TCS solution, and TCS alone was added 50µL of DEPC-treated water instead of siRNA mixture.

siRNA Transfection with TCS

Luc-siRNA (100uM in DEPC-treated water) was mixed with TCS (1 mg/mL in DMSO) and allowed to complex for 1hr at room temperature. The siRNA/TCS mixture was then added to Luciferase-expressing Hela cells, with a final concentration of 10 ug/mL of TCS and 25nM/100 nM of Luc-siRNA. The final DMSO concentration was 1%. The cells were allowed to incubate for 24 hours before cell lysis and luciferase addition according to standard procedures. Chemiluminescence was then measured and the results were given in % inhibition. Cells with 1% DMSO were used as a reference for 0% inhibition.

Cell Toxicity of TCS

TCS (20 mg/mL in DMSO) was added to Hela cells at increasing amounts and allowed to incubate for 24 hours. Cell toxicity was then investigated using a standard resazurin assay, and results were reported in % cell survival.

A.7 Flow Cytometry Histograms

Flow Cytometry Data Histograms for Hela cells

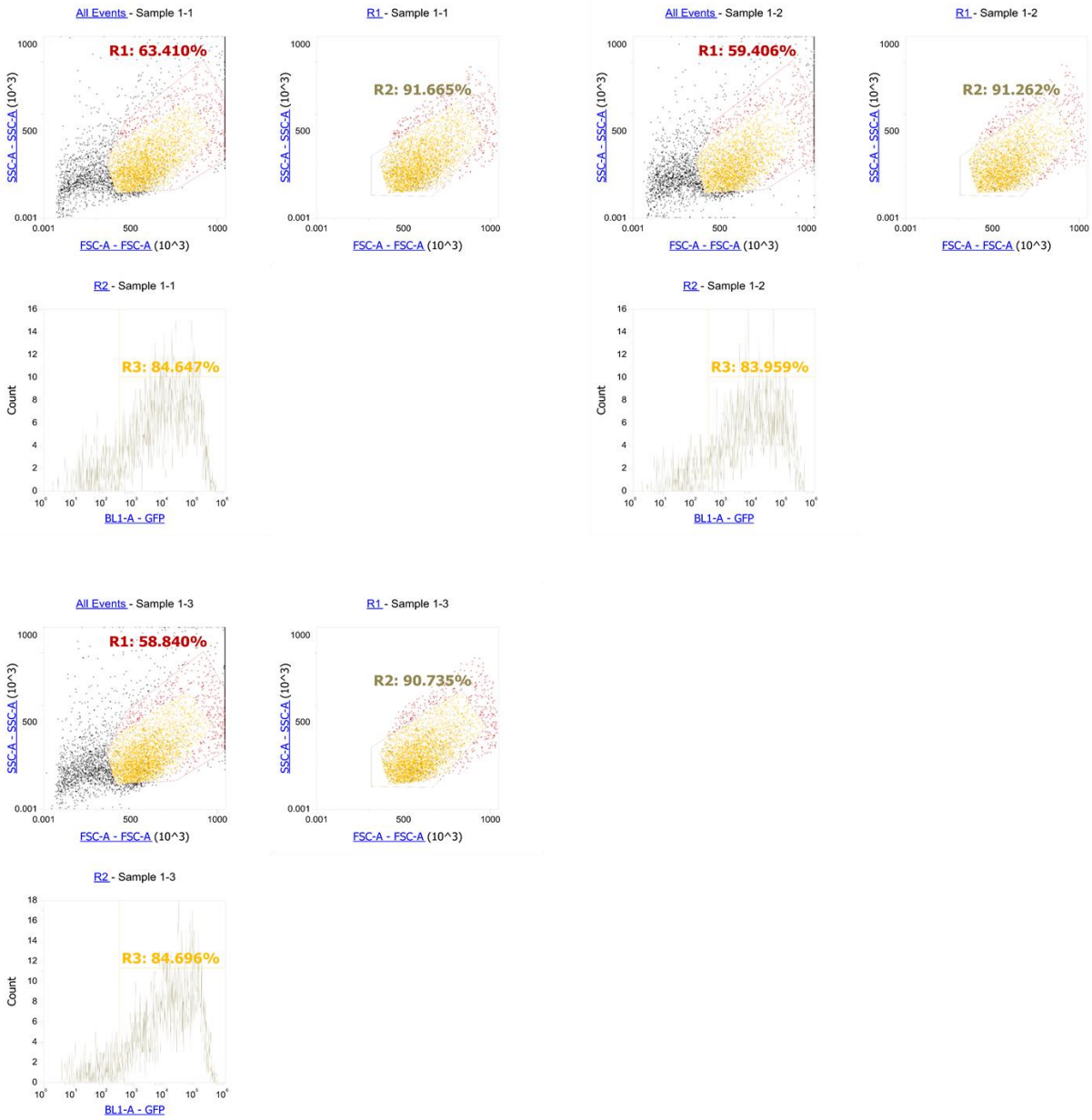


Figure 31: Flow Histograms for Hela cells transfected with GFP-expressing mRNA with Lipofectamine (2µL/mL).

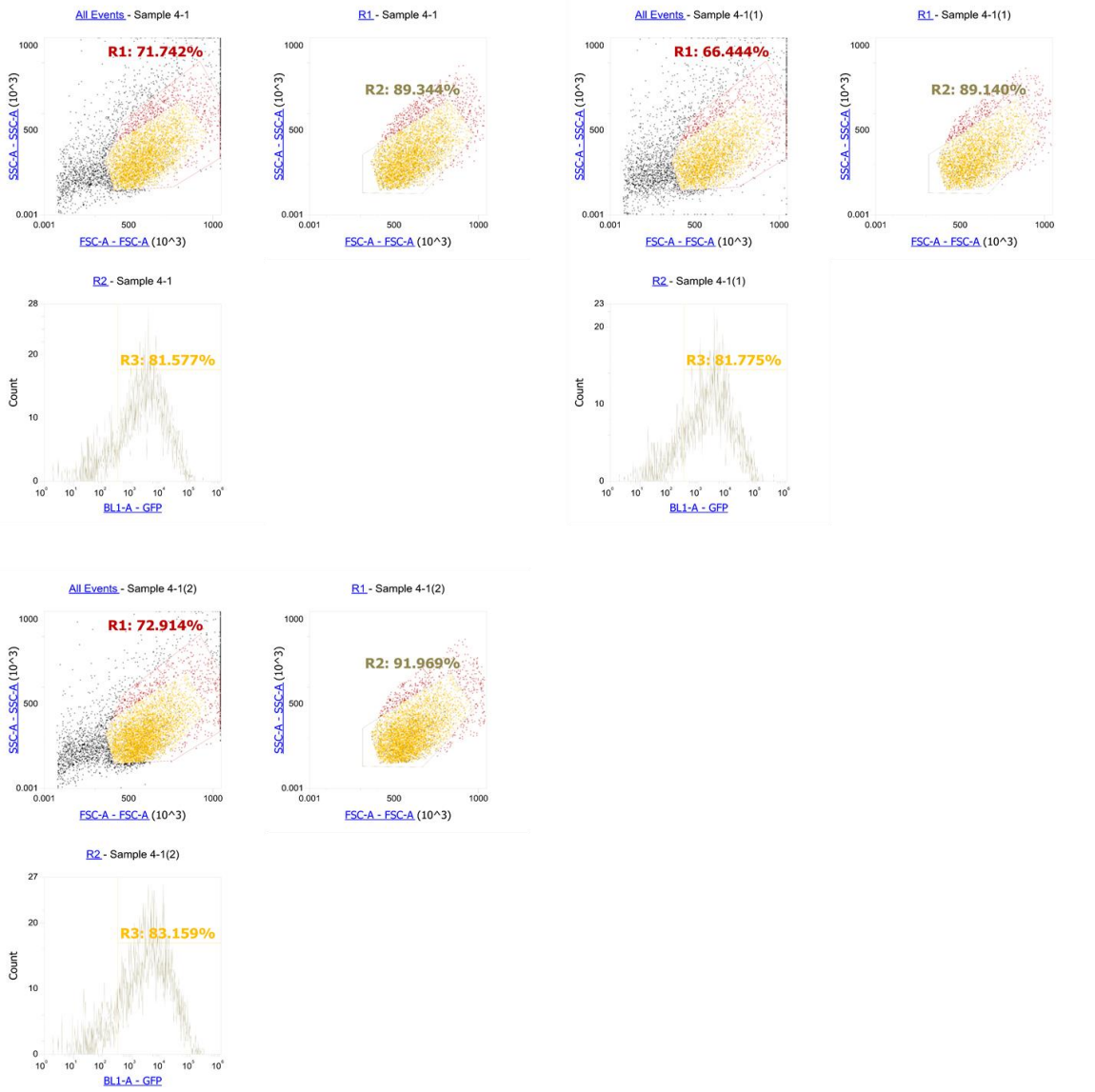


Figure 32: Flow Histograms for Hela cells transfected with GFP-expressing mRNA with Lipofectamine (2µL/mL) and PCS (2µg/mL).

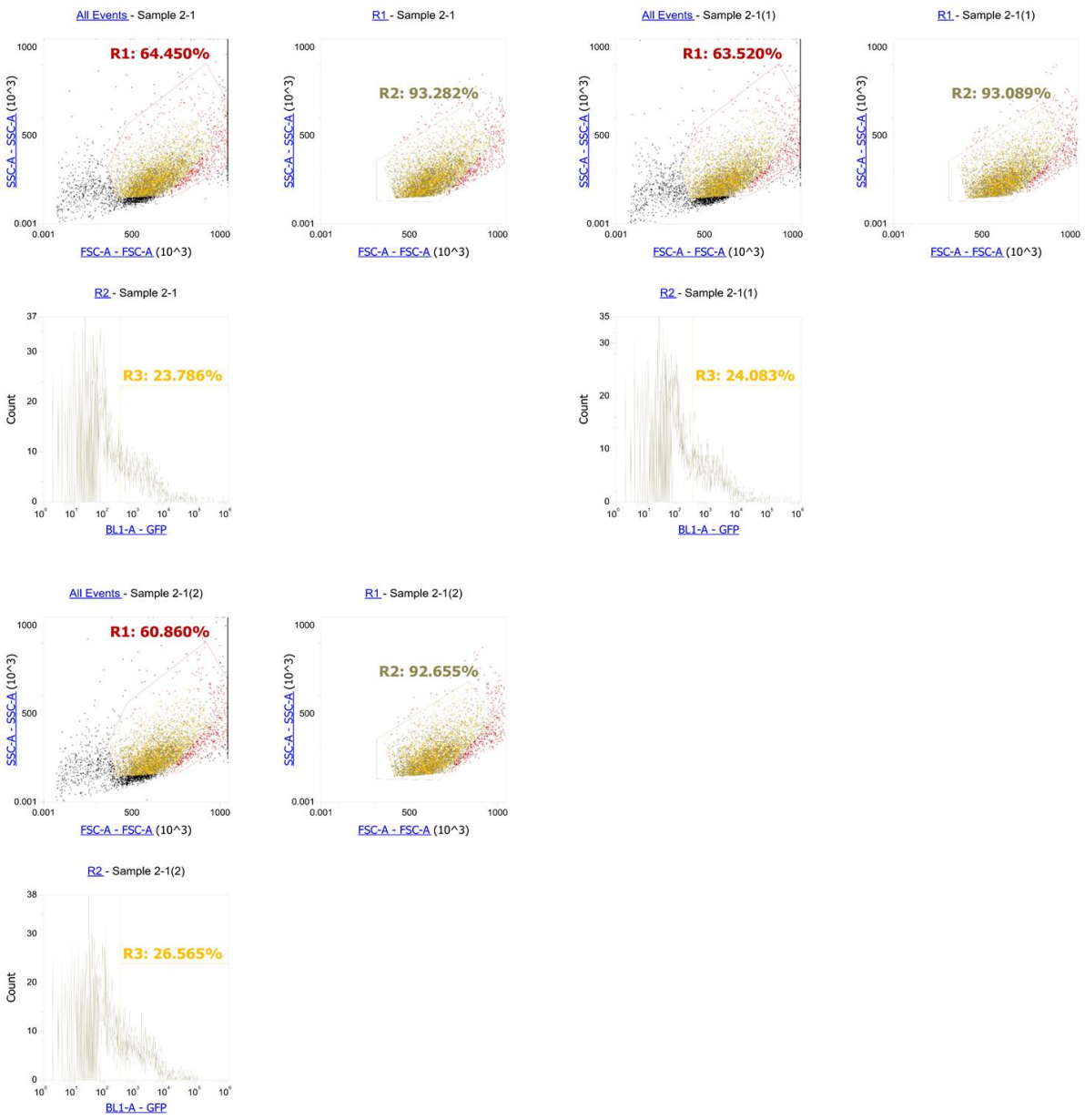


Figure 33: Histograms for Hela cells transfected with GFP-expressing mRNA with Lipofectamine (0.2μL/mL).

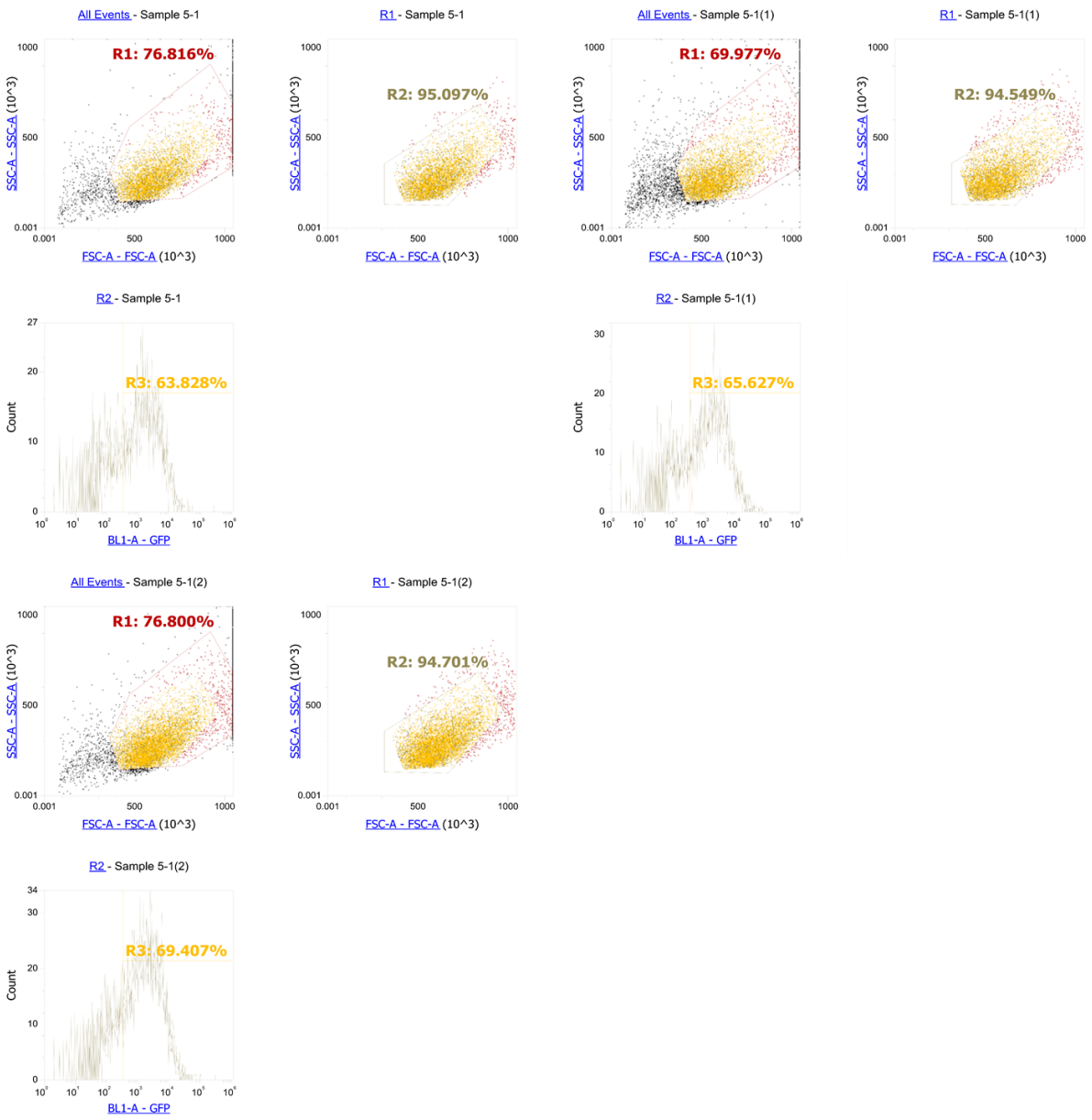


Figure 34: Histograms for Hela cells transfected with GFP-expressing mRNA with Lipofectamine (0.2µL/mL) and PCS (2µg/mL).

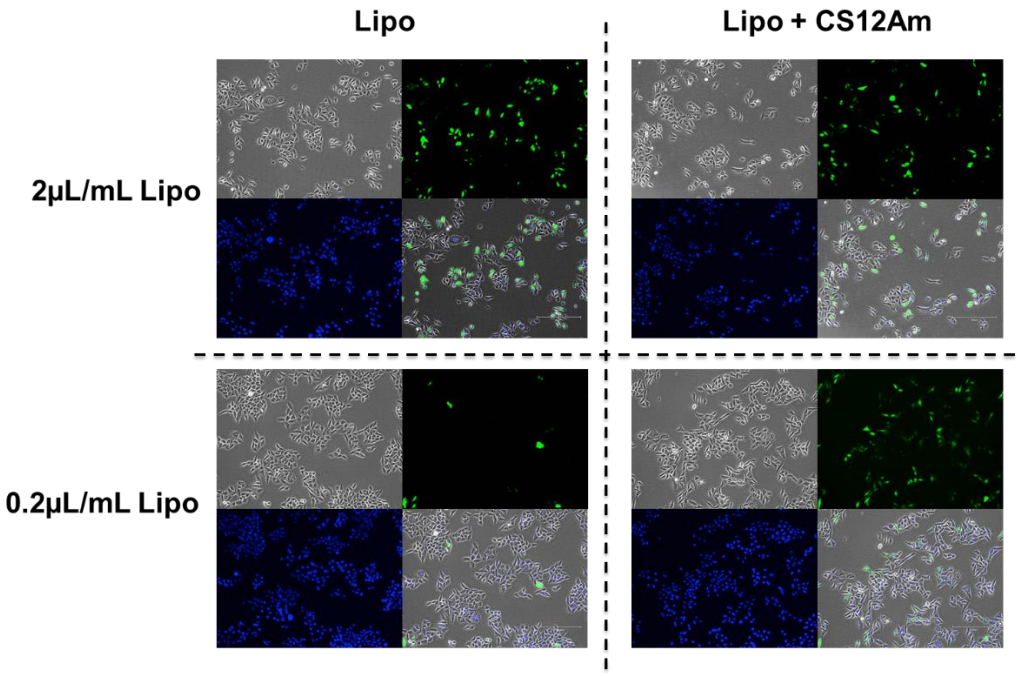


Figure 35: ***PCS increases the efficiency of low Lipofectamine 2000 doses in Hela cells*** *Microscopy images for the transfection of Hela cells using Lipofectamine (2µL/mL and 0.2µL/mL) with and without the addition of PCS (2µg/mL) The cells were incubated with the Lipo/PCS mixture for 4 hours before removal, and incubated further for another 24 hours in fresh media.*

Flow Cytometry Data Histograms for HUVEC Cells

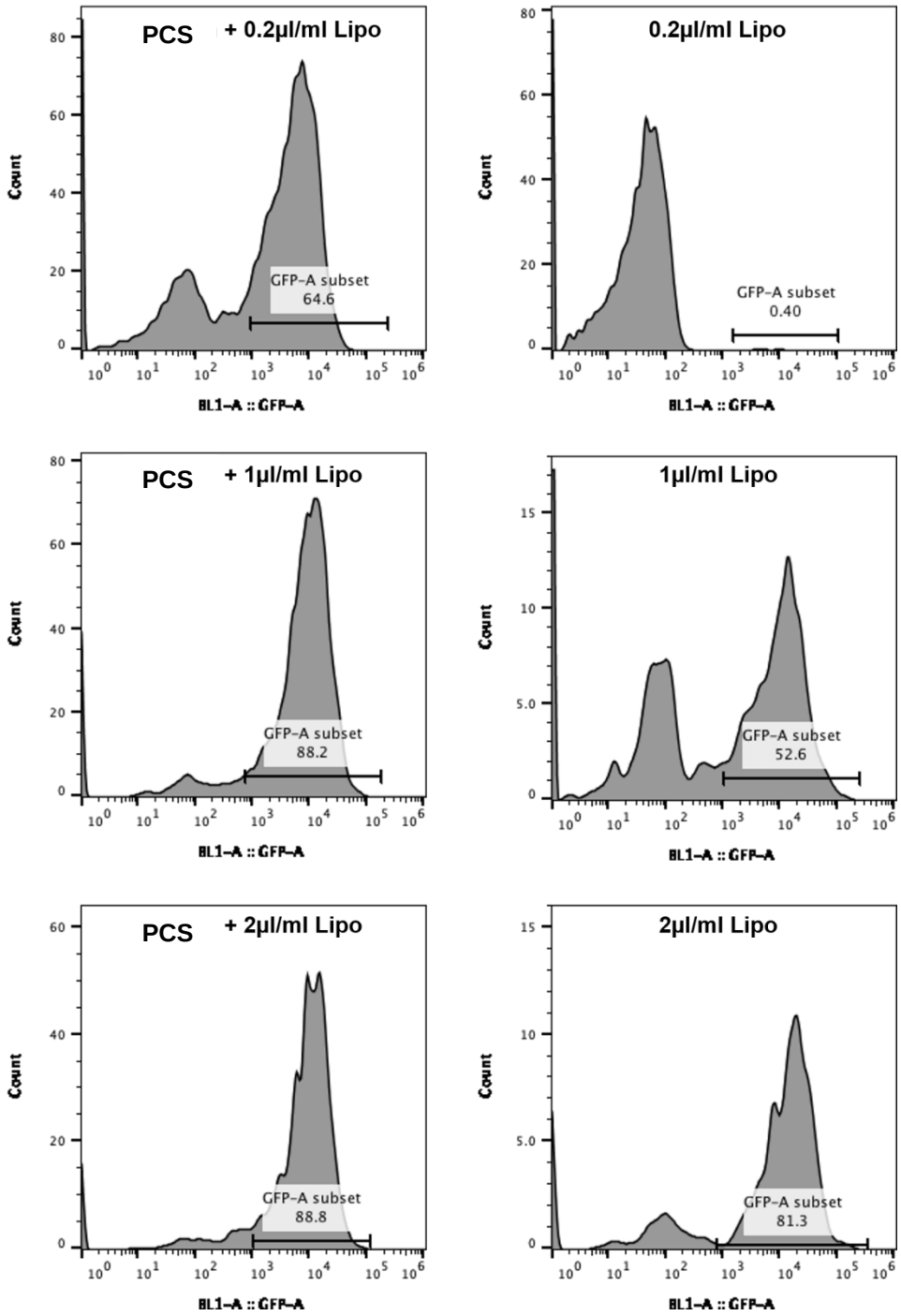


Figure 36: Representative flow cytometry histograms for transfection of HUVECs with PCS/Lipo and Lipofectamine 2000 alone at 0.2, 1, and 2 µL/mL of Lipofectamine. PCS concentrations are kept constant at 2µg/mL. The transfection was expressed as the percentage of positive cells/total cells. This also takes into consideration the toxicity that is caused by Lipofectamine 2000, and is a better measurement of the total efficiency of the system.

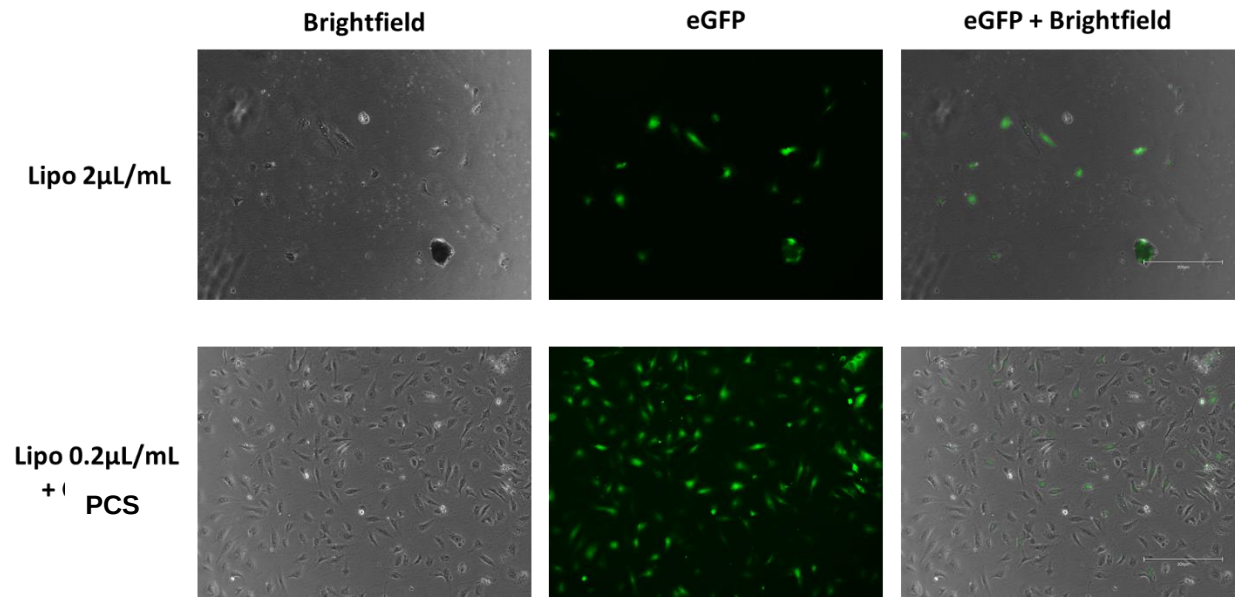


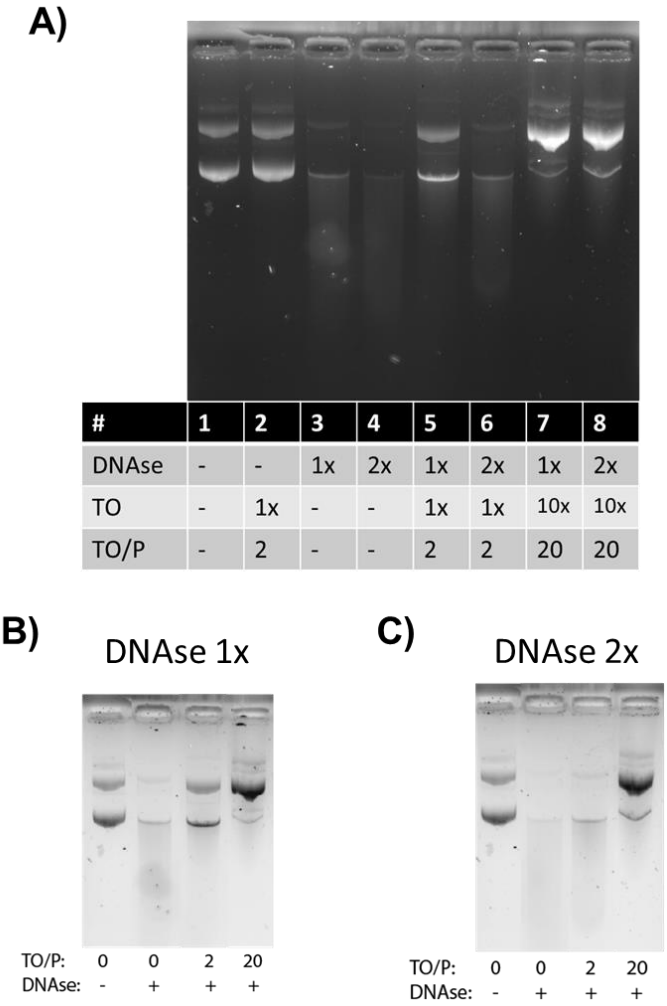
Figure 37: Comparison of cell counts of HUVEC cells treated with high amounts of Lipofectamine (2µL/mL), compared to treatment with low levels of Lipofectamine (0.2µL/mL) and PCS (5µg/mL). There is a marked increase in cell viability with cells treated with low levels of Lipofectamine.

A.8 TCS DNAse protection

In addition to increasing the delivery efficacy of nucleic acid through endosomal disruption, thiazole orange-based caged surfactants have the added benefit of potentially inhibiting nuclease activity due to the tight binding of TO with nucleic acids. In order to probe for this, we tested if C8-TCS/pDNA complex would be cleaved at a slower rate by DNAse than free pDNA. ***C8-TCS*** was utilized for this assay (see Section A.10).

Protocol: To 10µL of C8-TCS at various concentrations (in PBS at pH 7.4) was added pDNA (500ng in 10uL of PBS at pH 7.4). DNAse was then added (2.5µL at either 1mg/mL – 1X or 2mg/mL – 2X), and the mixture was incubated for 20 minutes at 37°C before loading onto a 1% agarose gel which was run at 120V until sufficient resolution was achieved. The gel was stained with SYBR Safe.

As shown in Figure 38A-C, binding of C8-TCS to pDNA significantly inhibited DNAse activity at both TO/P ratios of 2 and 20 at 110µg/mL of DNAse, while at 220µg/mL of DNAse significant protection is observed only at TO/P of 20. This confirms that TCS inhibits nuclease-catalyzed hydrolysis of nucleases.

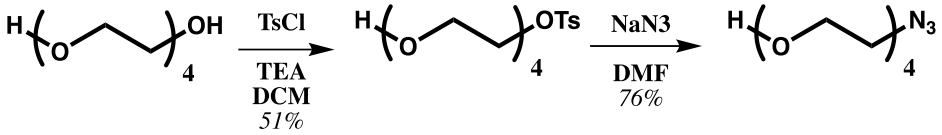


#	1	2	3	4	5	6	7	8
DNAse	-	-	1x	2x	1x	2x	1x	2x
TO	-	1x	-	-	1x	1x	10x	10x
TO/P	-	2	-	-	2	2	20	20

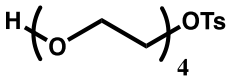
Figure 38: C8-TCS protects against nuclease-catalyzed hydrolysis. A) Gel showing DNAse protection by C8-TCS. All lanes contain pDNA (500ng), TO is added to either TO/P of 2 or 20 to pDNA, and DNAse is added to final concentrations of 110 and 220 µg/mL. B) Same gel as in A, but showing only lanes with 110µg/mL of DNAse. C) Same gel as in A, but showing only lanes with 220µg/mL of DNAse. Gel conditions: 1% agarose gel run at 120V

A.9 Synthetic Protocols

Synthesis and Characterization of HO-PEG$_4$-N$_3$

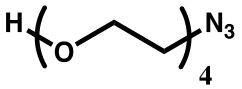


Scheme 5: Synthetic scheme of HO-PEG4-N$_3$.



5

2-(2-(2-(2-hydroxyethoxy)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (**5**): To a solution of tetraethylene glycol (26.7 mL, 154 mmol, 2 eq) in DCM (500 mL) was added TsCl (14.7 g, 77 mmol, 1 eq) and TEA (41.8 mL, 300 mmol, 3.9 eq). The mixture was stirred at room temperature for 22 hours. The reaction mixture was then washed with saturated ammonium chloride and extracted with additional DCM. The combined organic layers were dried and concentrated. The crude product was purified by flash chromatography (100% DCM) to afford a colorless oil (13.58 g, 62.8 mmol, 51% yield). ^{1}H NMR (600 MHz, Chloroform-d) δ 7.80 (d, J = 8.1 Hz, 2H), 7.34 (d, J = 8.0 Hz, 2H), 4.16 (t, J = 4.8 Hz, 2H), 3.86 – 3.57 (m, 14H), 2.45 (s, 3H), 2.27 (b, 1H). ^{13}C NMR (151 MHz, Chloroform-d) δ 144.94, 133.15, 129.97, 128.14, 72.58, 70.90, 70.82, 70.64, 70.50, 69.39, 68.87, 61.90, 21.80. HRMS (ESI+): Found: 371.1135m/z, Calc: 371.1135m/z for $[C_{15}H_{24}O_7SNa]^+$

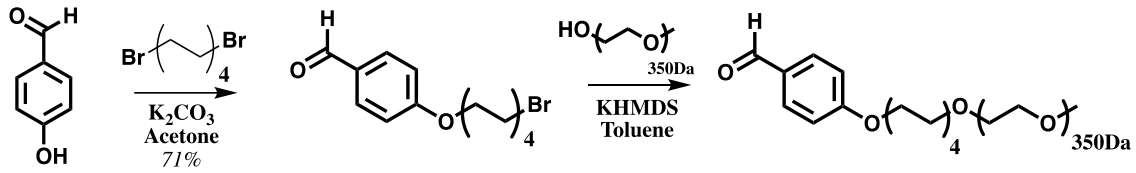


6

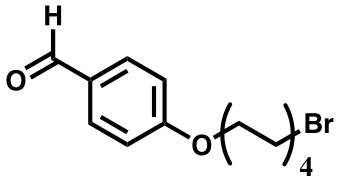
2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethan-1-ol (**6**): To a solution of **5** (994.5 mg, 2.85 mmol, 1 eq) in DMF (10 mL) was added sodium azide (986 mg, 15.17 mmol, 5 eq). The mixture was set to reflux at 60ºC for 5 hours, after which it was concentrated and resuspended in water. The mixture was extracted with DCM and washed with water and brine. The organic phase was dried with sodium sulfate, filtered, and concentrated. The resulting crude product was purified by flash chromatography (100% EtOAc to 1:19 MeOH:EtOAc) to yield a colorless oil (477.7 mg, 2.18 mmol, 76% yield). ^{1}H NMR (600 MHz, Chloroform-d) δ 3.73 (t, J = 4.5 Hz, 2H), 3.70 – 3.66 (m, 10H), 3.62 (t, J = 4.5 Hz, 2H), 3.40 (t, J = 5.1 Hz, 2H), 2.13 (b, 1H). ^{13}C NMR (151 MHz,

Chloroform-d) δ 72.60, 70.87, 70.84, 70.77, 70.53, 70.21, 61.92, 50.83. HRMS (ESI+): Found: 242.1110 m/z, Calc: 242.1111 m/z for $[C_8H_{17}O_4N_3Na]^+$

Synthesis of Ald-C8-PEG

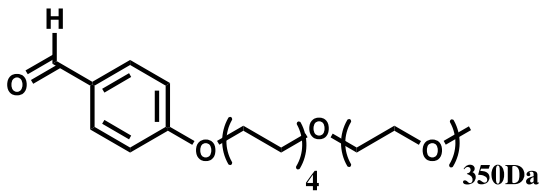


Scheme 6: *Synthetic scheme for Ald-C8-PEG.*



1

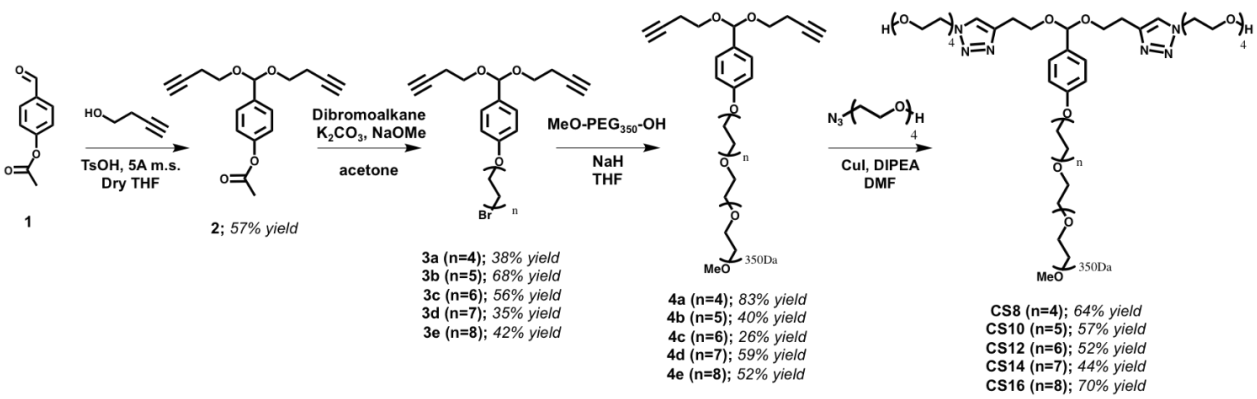
4-((8-bromooctyl)oxy)benzaldehyde (**1**): To a solution of 4-hydroxybenzaldehyde (2.4025 g, 19.7 mmol, 1 eq) in acetone (100 mL) was added 1,8-dibromooctane (11 mL, 16 g, 59.4 mmol, 3 eq) and anhydrous potassium carbonate (5.4064 g, 39.1 mmol, 2 eq) . The mixture was set to reflux at 70ºC for 13.5 hours, after which it was concentrated and resuspended in DCM (100 mL). The organic phase was washed with brine (100mL), dried with sodium sulfate, filtered and concentrated. The resulting crude was purified by flash chromatography (100% hexanes to 1:9 EtOAc:hexanes) to yield a colorless solid (4.3620 g, 13.93 mmol, 71 % yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 9.87 (s, 1H), 7.82 (d, J = 8.8 Hz, 2H), 6.98 (d, J = 8.8 Hz, 2H), 4.03 (t, J = 6.5 Hz, 2H), 3.41 (t, J = 6.8 Hz, 2H), 2.04 – 1.74 (m, 4H), 1.56 – 1.30 (m, 8H). ^{13}C NMR (101 MHz, Chloroform-d) δ 190.96, 164.34, 132.13, 129.88, 114.86, 68.46, 34.12, 32.88, 29.27, 29.14, 28.79, 28.20, 26.01. HRMS (EI+): Found: 312.0722m/z, Calc: 312.0725m/z for $[C_{15}H_{21}O_2Br]^+$



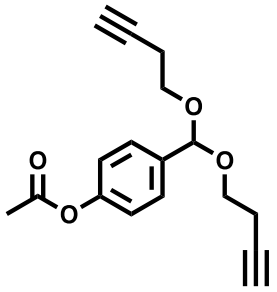
Ald-C8-PEG

*4-((8-(PEG$_{350Da}$)octyl))benzaldehyde (***Ald-C8-PEG)**: To a mixture of HO-PEG$_{350Da}$-OMe (2.124g, 6.07mmol, 0.95eq) in toluene (16mL) was added 4Å molecular sieves. The mixture was cooled to 0^{o}C under a nitrogen flow, and KHMDS (15% in toluene, 9.31mL, 1.21g, 6.07mmol, 0.95eq) was added dropwise. After stirring for 10 minutes, a mixture of aldehyde (2.003g, 6.36mmol, 1eq) in toluene (12mL) was added dropwise to the mixture. The reaction was then allowed to heat to room temperature overnight while stirring. The reaction was then quenched with sat. NH_4Cl, and extracted into DCM (3x100mL). The organic phases were dried with brine and Na_2SO_4, before purifying by silica column chromatography (DCM:MeOH; 1:0 to 95:5) to yield an orange-tinted oil (205.6mg, 0.35mmol, 6% yield). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 9.88 (s, 1H), 7.82 (d, *J* = 8.8 Hz, 2H), 6.98 (d, *J* = 8.7 Hz, 2H), 4.03 (t, *J* = 6.5 Hz, 2H), 3.67-3.62 (m, 31H), 3.59 – 3.56 (m, 2H), 3.56-3.53 (m, 2H), 3.45 (t, *J* = 6.8 Hz, 2H), 3.38 (s, 3H), 1.84-1.76 (m, 2H), 1.63 – 1.54 (m, 2H), 1.50-1.41 (m, 2H), 1.38 – 1.30 (m, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 190.98, 164.40, 132.14, 114.89, 72.09, 71.62, 70.73, 70.22, 68.53, 59.19, 29.75, 29.52, 29.41, 29.18, 26.17, 26.05. HRMS (ESI+): Found: 639.3711m/z, Calc: 639.3715m/z for $[C_{32}H_{56}O_{11}Na]^+$

General syntheses of CS8-CS16

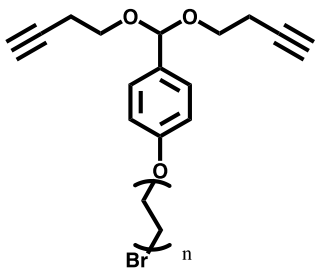


*Scheme 7: Synthetic scheme for the caged surfactants **CS8-CS16**.*



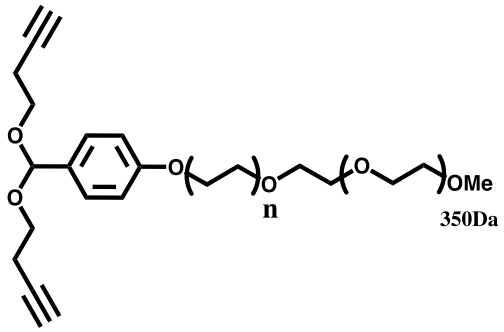
2

4-(bis(but-3-yn-1-yloxy)methyl)phenyl acetate (**2**): To a mixture of 4-formylphenyl acetate (5.0033g, 30.5mmol, 1eq) in THF (30mL, distilled from sodium using benzophenone as an indicator) was added but-3-yn-1-ol (13.9mL, 12.8g, 0.183mol, 6eq) and 5Å molecular sieves (roughly ½ of THF volume). Then, toluenesulfonic acid (0.834g, 4.87mmol, 0.16eq, anhydrous, dried in a Dean Stark trap) was added, and the reaction was allowed to stir overnight. The reaction was then quenched with TEA (6mL), and the mixture was concentrated by rotary evaporation. The crude product was purified by silica column chromatography (EtOAc:Hex:TEA; 10:90:0.1) to yield a white solid (4.982g, 57% yield). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.51 (d, *J* = 8.6 Hz, 2H), 7.09 (d, *J* = 8.6 Hz, 2H), 5.65 (s, 1H), 3.73 – 3.57 (m, 4H), 2.50 (td, *J* = 6.8, 2.7 Hz, 4H), 2.30 (s, 3H), 1.99 (t, *J* = 2.7 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.49, 150.86, 135.60, 128.04, 121.49, 100.81, 81.42, 69.53, 63.34, 21.28, 20.02. HRMS (EI+): Found: 286.1203m/z, Calc: 286.1205m/z for $[C_{17}H_{18}O_4]^+$



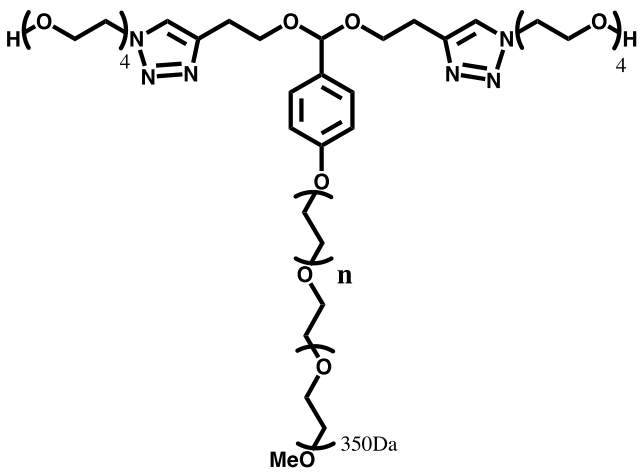
3a-3e

4-(bis(but-3-yn-1-yloxy)methyl)phenyl alkylbromides (**3a-3e**): To a suspension of K_2CO_3 (2eq) in acetone (0.1M to **2**) was added **2** (1eq) and dibromoalkane (3eq) under a flow of nitrogen. A solution of NaOMe (30% in MeOH, 1eq) was added to the solution, and it was allowed to stir at room temperature overnight, after which the reaction was quenched with TEA and concentrated by rotary evaporation. The crude mixtures were then purified by silica flash chromatography (EtOAc:HEX:TEA; 5:95:0.1 to 10:90:0.1).



4a-4e

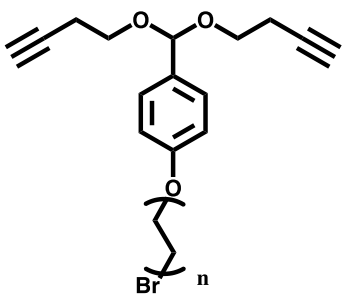
4-(bis(but-3-yn-1-yloxy)methyl)phenyl alkyl-PEG350Da (**4a-4e**): To a mixture of HO-PEG_{350Da}-OMe (2eq) in THF (0.7M) at 0^{o}Cwas added NaH (60% in mineral oil, 2eq). The suspension was allowed to stir for 30 minutes before **3n** (1eq) in THF (0.5M) was added dropwise. The mixture was then left to stir overnight, before some TEA was added for stability, and the solvent was removed by rotary evaporation. The crude was then purified by silica flash chromatography (MeOH:DCM:TEA; 1:99:0.1 to 4:96:0.1), and the PEG-containing fractions were collected to yield a mixture of **4a-4e** and excess HO-PEG350-OMe. These fractions were used directly in the next step.



CS8-CS16

Caged surfactants (**CS8-CS16**): To dry DMF (6 mL) was added DIPEA (1 eq), **4a-e** (1 eq), tetraethylene glycol azide (4 eq) and lastly CuI (0.2 eq). The reaction was allowed to stir at RT for 48 hours. TEA was added to the crude reaction to stabilize the acetal, and the majority of the DMF was removed by co-evaporation with toluene. The resulting crude product was purified by flash chromatography (90:10:0.1 EtOAc:MeOH:TEA to 20:80:0.1 MeOH:DCM:TEA) to yield **CS8-CS16**.

Characterization of CS8-CS16 and intermediates



3a-3e

1-(bis(but-3-yn-1-yloxy)methyl)-4-((8-bromooctyl)oxy)benzene (**3a**): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.39 (d, *J* = 8.5 Hz, 2H), 6.87 (d, *J* = 8.8 Hz, 2H), 5.61 (s, 1H), 3.95 (t, *J* = 6.5 Hz, 2H), 3.64 (qt, *J* = 9.5, 6.9 Hz, 4H), 3.41 (t, *J* = 6.8 Hz, 2H), 2.49 (td, *J* = 6.9, 2.7 Hz, 4H), 1.98 (t, *J* = 2.7 Hz, 2H), 1.90-1.81 (m, 2H), 1.81-1.74 (m, 2H), 1.51 – 1.31 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.42, 130.09, 128.05, 114.25, 101.35, 81.53, 69.44, 68.05, 63.30, 34.13, 32.91, 29.34, 29.32, 28.82, 28.23, 26.09, 20.07. HRMS (EI+): Found: 434.1453m/z Calc: 434.1457m/z for $[C_{23}H_{31}O_3Br]^+$

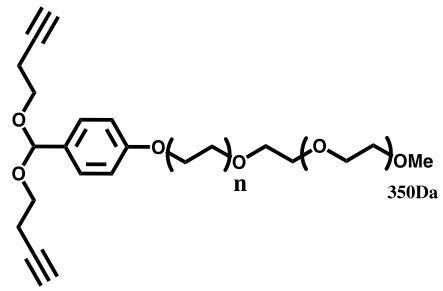
1-(bis(but-3-yn-1-yloxy)methyl)-4-((10-bromodecyl)oxy)benzene (**3b**): ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.39 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.7 Hz, 2H), 5.61 (s, 1H), 3.95 (t, *J* = 6.5 Hz, 2H), 3.64 (qt, *J* = 9.5, 6.9 Hz, 4H), 3.41 (t, *J* = 6.9 Hz, 2H), 2.49 (td, *J* = 6.9, 2.7 Hz, 4H), 1.98 (t, *J* = 2.7 Hz, 2H), 1.93 – 1.69 (m, 4H), 1.49 – 1.28 (m, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.42, 130.04, 128.03, 114.24, 101.33, 81.53, 69.44, 68.09, 63.28, 34.19, 32.94, 29.56, 29.48, 29.37, 28.87, 28.29, 26.15, 20.05. HRMS (EI+): Found: 462.1758m/z Calc: 462.1770m/z for $[C_{25}H_{35}O_3Br]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((12-bromododecyl)oxy)benzene **(3c):** ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.39 (d, *J* = 8.7 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 5.61 (s, 1H), 3.95 (t, *J* = 6.5 Hz, 2H), 3.64 (qt, *J* = 9.3, 6.8 Hz, 4H), 3.40 (t, *J* = 6.8 Hz, 2H), 2.49 (td, *J* = 6.9, 2.7 Hz, 4H), 1.98 (t, *J* = 2.7 Hz, 2H), 1.93 – 1.71 (m, 4H), 1.49 – 1.26 (m, 16H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.42, 130.01, 128.02, 114.22, 101.32, 81.52, 69.44, 68.11, 63.26, 34.22, 32.95, 29.65, 29.55, 29.51, 29.37, 28.88, 28.30, 26.16, 20.04. HRMS (EI+): Found: 490.2074m/z Calc: 490.2083m/z for $[C_{27}H_{39}O_3Br]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((14-bromotetradecyl)oxy)benzene **(3d):** ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.39 (d, *J* = 8.6 Hz, 2H), 6.88 (d, *J* = 8.8 Hz, 2H), 5.61 (s, 1H), 3.95 (t, *J* = 6.4 Hz, 2H), 3.64 (tdd, *J* = 9.5, 7.0, 2.7 Hz, 4H), 3.41 (t, *J* = 6.8 Hz, 2H), 2.49 (td, *J* = 6.9, 2.7 Hz, 4H), 1.98 (t, *J* = 2.7 Hz, 2H), 1.92 – 1.72 (m, 4H), 1.44 (dt, *J* = 10.7, 6.0 Hz, 4H), 1.37 – 1.23 (m,

16H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.48, 130.07, 128.05, 114.28, 101.39, 81.54, 69.44, 68.17, 63.33, 34.20, 32.99, 29.74, 29.68, 29.54, 29.41, 28.92, 28.33, 26.20, 20.08. HRMS (EI+): Found: 518.2399m/z Calc: 518.2396m/z for $[C_{29}H_{43}O_3Br]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((16-bromohexadecyl)oxy)benzene **(3e):** ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.39 (d, *J* = 8.7 Hz, 2H), 6.87 (d, *J* = 8.7 Hz, 2H), 5.61 (s, 1H), 3.95 (t, *J* = 6.6 Hz, 2H), 3.64 (qt, *J* = 9.4, 6.9 Hz, 4H), 3.40 (t, *J* = 6.9 Hz, 2H), 2.49 (td, *J* = 6.9, 2.7 Hz, 4H), 1.98 (t, *J* = 2.6 Hz, 2H), 1.85 (dt, *J* = 14.5, 7.0 Hz, 2H), 1.81 – 1.73 (m, 2H), 1.47 – 1.38 (m, 4H), 1.35-1.24 (m, 20H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.46, 130.04, 128.03, 114.26, 101.36, 81.53, 77.37, 69.44, 68.16, 63.30, 34.21, 32.98, 29.79, 29.73, 29.68, 29.58, 29.40, 28.91, 28.32, 26.19, 20.06. HRMS (EI+): Found: 546.2701m/z Calc: 546.2709m/z for $[C_{31}H_{47}O_3Br]^+$



4a-4e

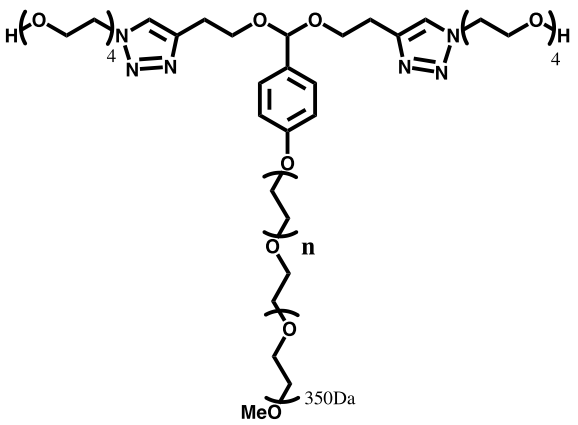
1-(bis(but-3-yn-1-yloxy)methyl)-4-((8-PEG$_{350Da}$octyl)oxy)benzene (**4a**): HRMS (ESI+): Found: 761.4448m/z Calc: 761.4446m/z for $[C_{40}H_{66}O_{12}Na]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((10-PEG$_{350Da}$decyl)oxy)benzene (**4b**): HRMS (ESI+): Found: 789.4760m/z Calc: 789.4759m/z for $[C_{42}H_{70}O_{12}Na]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((12-PEG$_{350Da}$dodecyl)oxy)benzene (**4c**): HRMS (ESI+): Found: 817.5068m/z Calc: 817.5072m/z for for $[C_{44}H_{74}O_{12}Na]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((14-PEG$_{350Da}$tetradecyl)oxy)benzene (**4d**): HRMS (ESI+): Found: 845.5384m/z Calc: 845.5385m/z for $[C_{46}H_{78}O_{12}Na]^+$

1-(bis(but-3-yn-1-yloxy)methyl)-4-((16-PEG$_{350Da}$hexadecyl)oxy)benzene (**4e**): HRMS (ESI+): Found: 873.5695m/z Calc: 873.5699m/z for $[C_{48}H_{82}O_{12}Na]^+$



CS8-CS16

CS8: ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.80 (s, 2H), 7.27 (d, J = 8.4 Hz, 2H), 6.86 (d, J = 8.8 Hz, 2H), 5.49 (s, 1H), 4.53 (t, J = 5.2 Hz, 4H), 3.96 (t, J = 6.5 Hz, 2H), 3.86 (t, J = 5.1 Hz, 4H), 3.77 – 3.68 (m, 2H), 3.68 – 3.51 (m, 69H), 3.47 (t, J = 6.5 Hz, 3H), 3.35 (s, 3H), 2.94 (t, J = 6.6 Hz, 4H), 1.82 – 1.71 (m, 2H), 1.57 (d, J = 7.0 Hz, 2H), 1.48 (d, J = 7.8 Hz, 2H), 1.37 (s, 8H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 145.00, 127.94, 123.10, 114.25, 101.86, 72.65, 71.65, 70.72, 70.65, 70.58, 70.44, 69.74, 68.15, 64.67, 61.79, 59.19, 50.23, 29.78, 29.57, 29.42, 26.67, 26.20. HRMS (ESI+): Found: 1199.6918m/z Calc: 1199.6885m/z for $[C_{56}H_{100}O_{20}N_6Na]^+$

***CS10*:** ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.80 (s, 2H), 7.27 (d, *J* = 8.7 Hz, 2H), 6.87 (d, *J* = 8.7 Hz, 2H), 5.49 (s, 1H), 4.53 (t, *J* = 5.0 Hz, 4H), 3.96 (t, *J* = 6.4 Hz, 3H), 3.85 (t, *J* = 5.1 Hz, 4H), 3.74 – 3.70 (m, 2H), 3.68 – 3.51 (m, 61H), 3.47 (t, *J* = 6.6 Hz, 3H), 3.35 (s, 3H), 2.94 (t, *J* = 6.5 Hz, 4H), 1.77 (t, *J* = 7.3 Hz, 2H), 1.57 (t, *J* = 6.8 Hz, 2H), 1.47 (d, *J* = 7.2 Hz, 2H), 1.34 (s, 12H). ^{13}C NMR (101 MHz, MeOD) δ 160.73, 146.15, 131.87, 129.02, 124.84, 115.10, 103.04, 73.68, 72.98, 72.36, 71.58, 71.49, 71.41, 71.37, 71.19, 70.44, 69.03, 65.55, 62.23, 59.10, 51.32, 47.41, 30.74, 30.67, 30.56, 30.42, 27.22, 10.24. HRMS (ESI+): Found: 1227.7228m/z Calc: 1227.7198m/z for $[C_{58}H_{104}O_{20}N_6Na]^+$.

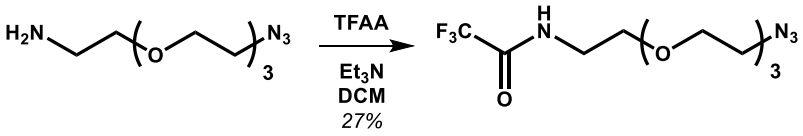
CS12: ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.81 (s, 2H), 7.27 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 5.49 (s, 1H), 4.54 (m, 4H), 3.96 (t, *J* = 6.4 Hz, 2H), 3.87 (m, 4H), 3.83 – 3.77 (m, 2H), 3.76 – 3.50 (m, 71H), 3.47 (t, *J* = 6.6 Hz, 3H), 3.35 (s, 3H), 3.00 – 2.92 (m, 4H), 1.83 – 1.72 (m, 2H), 1.61-1.54 (m, 2H), 1.52-1.45 (m, 2H), 1.32 (b, 16H). ^{13}C NMR (101 MHz, MeOD) δ 160.74, 146.14, 129.02, 124.84, 115.10, 103.04, 73.68, 72.99, 72.37, 71.58, 71.49, 71.42, 71.37, 71.19, 70.44, 69.04, 65.57, 62.23, 59.10, 51.32, 47.43, 30.73, 30.60, 30.58, 30.53, 30.43, 27.23, 10.27. HRMS (ESI+): Found: 1255.7510m/z Calc: 1255.7511m/z for $[C_{60}H_{108}O_{20}N_6Na]^+$.

***CS14*:** ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.85 (s, 2H), 7.32 (d, *J* = 8.6 Hz, 2H), 6.89 (d, *J* = 8.8 Hz, 2H), 5.31 (s, 1H), 4.58 – 4.52 (m, 4H), 3.97 (t, *J* = 6.4 Hz, 2H), 3.88 (dd, *J* = 5.6, 4.6 Hz, 4H), 3.80 (t, *J* = 6.1 Hz, 4H), 3.62 (d, *J* = 7.7 Hz, 53H), 3.46 (t, *J* = 6.6 Hz, 2H), 3.35 (s, 3H), 2.90 (t, *J* = 6.7 Hz, 4H), 1.76 (dq, *J* = 8.2, 6.5 Hz, 2H), 1.56 (q, *J* = 6.9 Hz, 2H), 1.51-1.44 (m, 2H), 1.36-1.28 (m, 20H). ^{13}C NMR (101 MHz, MeOD) δ 160.76, 131.54, 128.99, 124.77, 115.06, 104.56, 73.66, 72.97, 72.36, 71.57, 71.48, 71.41, 71.37, 71.17, 70.41, 68.99, 62.21, 62.13, 59.10, 51.31, 47.94, 30.73, 30.69, 30.58, 30.49, 30.40, 29.93, 27.21, 27.16, 9.22. HRMS (ESI+): Found: 1283.7853m/z Calc: 1283.7824m/z for $[C_{62}H_{112}O_{20}N_6Na]^+$.

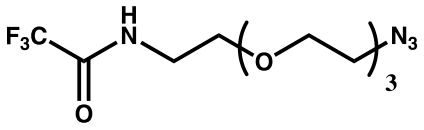
CS16: ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.81 (s, 2H), 7.27 (d, J = 8.7 Hz, 2H), 6.86 (d, J = 8.7 Hz, 2H), 5.49 (s, 1H), 4.53 (t, J = 5.1 Hz, 4H), 3.96 (t, J = 6.5 Hz, 2H), 3.85 (t, J = 5.1 Hz, 4H), 3.69 – 3.49 (m, 59H), 3.46 (t, J = 6.6 Hz, 2H), 3.35 (s, 3H), 1.81 – 1.71 (m, 2H), 1.56 (q, J = 6.8 Hz, 2H), 1.46 (d, J = 7.9 Hz, 2H), 1.31 (d, J = 8.0 Hz, 22H). ^{13}C NMR (101 MHz, MeOD) δ 160.73, 146.14, 131.86, 129.01, 124.85, 115.09, 103.03, 73.67, 72.97, 72.37, 71.56, 71.48, 71.41, 71.17, 70.44, 69.04, 65.55, 62.22, 59.10, 51.32, 47.50, 30.76, 30.59, 30.53, 30.43, 27.23, 10.09. HRMS (ESI+): Found: 1311.8137m/z Calc: 13.11.8137m/z for $[C_{64}H_{116}O_{20}N_6Na]^+$.

Synthesis of N_3-PEG$_4$-NHCOCF$_3$

N3-PEG4-NHCOCF3 (**7**) was synthesized as previously reported[36].



Scheme 8: Synthetic scheme for the synthesis of N_3-PEG4-NHCOCF$_3$.

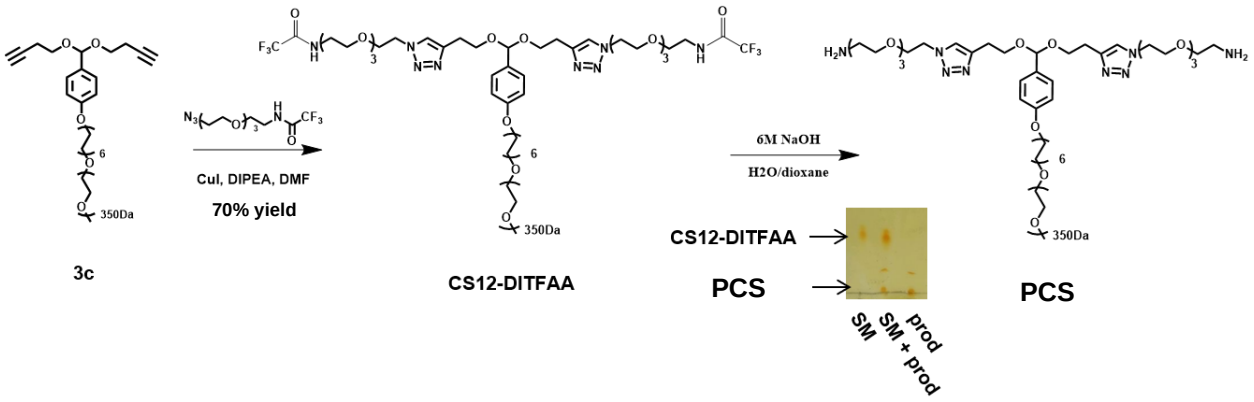


7

N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2,2,2-trifluoroacetamide ***(7):*** A mixture of 2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethan-1-amine (2g, 9.2mmol, 1eq) and TEA (1.55mL, 1.13g, 11.2mmol, 1.2eq) in DCM (20mL) was cooled to 0^{o}C under a flow of nitrogen. Then, TFAA (1.55mL, 2.3g, 11mmol, 1.2eq) was added dropwise, and the reaction was left to slowly cool to room temperature overnight. DCM (150mL) and DI H_2O (100mL) were added, and the organic phase was collected. The organic phase was added to sat. $NaHCO_3$ (100mL) and the organic phase was collected, and dried with Na_2SO_4 and concentrated. The crude product was purified by silica chromatography (MeOH:DCM from 2:98 to 5:95) to yield **7** as a colorless oil (784.8mg, 2.5mmol, 27% yield). The ^{1}H and ^{13}C NMR corresponded with that previously reported. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.11 (s, 1H), 3.73 – 3.58 (m, 12H), 3.58 – 3.49 (m, 2H), 3.42 – 3.33 (m, 2H). 13C NMR (226 MHz, CDCl3) δ 157.36 (d, J = 37.0 Hz), 116.05 (d, J = 287.9 Hz), 70.84,

70.82, 70.72, 70.48, 70.21, 68.84, 50.81, 39.86. HRMS (ESI+): Found 337.1089m/z, Calc 337.1094m/z for $[C_{10}H_{17}O_4N_4F_3Na]^+$.

Synthesis of PCS

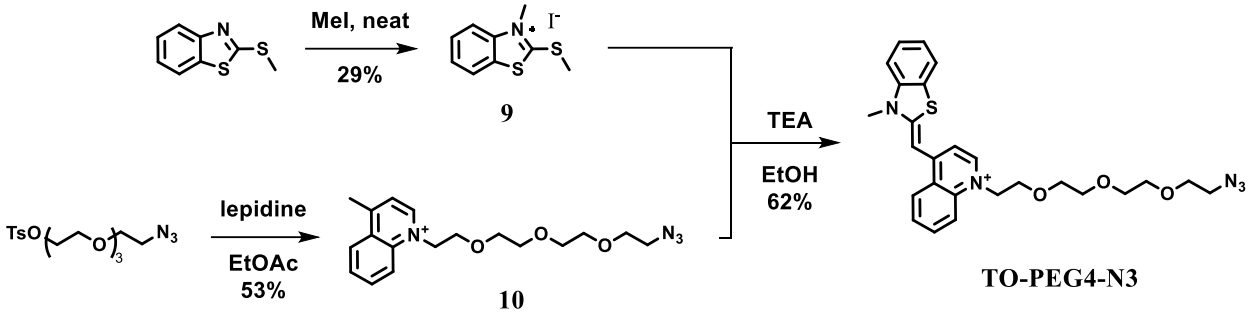


Scheme 9: Synthetic scheme for PCS.

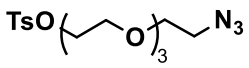
CS12-DITFAA: To a solution of **4c** (301.4mg, 0.4mmol, 1eq) in DMF (2mL) and DIPEA (209µL, 155mg, 1.2mmol, 3eq) was added **7** (499.8mg, 1.6mmol, 4eq) and CuI (15mg, 0.08mmol, 0.2eq). Nitrogen gas was bubbled through the reaction mixture for 30 minutes, and the reaction was set to stir at room temperature for 48 hours before DMF was removed through co-evaporation with toluene (2 x 100mL). The crude product was purified by silica chromatography (MeOH:DCM:TEA from 2:98:0.2 to 5:95:0.2) to yield **CS12-DITFAA** as a clear oil (391mg, 0.28mmol, 70% yield). ^{1}H NMR (400 MHz, MeOD) δ 7.80 (s, 2H), 7.26 (d, J = 8.6 Hz, 2H), 6.86 (d, J = 8.7 Hz, 2H), 5.49 (s, 1H), 4.53 (t, J = 5.0 Hz, 4H), 3.96 (t, J = 6.5 Hz, 2H), 3.85 (t, J = 5.1 Hz, 4H), 3.77 – 3.42 (m, 57H), 3.35 (s, 3H), 2.94 (t, J = 6.5 Hz, 3H), 1.75 (q, J = 7.0 Hz, 2H), 1.56 (q, J = 6.8 Hz, 2H), 1.47 (s, 2H), 1.32 (s, 14H). ^{13}C NMR (226 MHz, MeOD) δ 160.75, 146.17, 131.86, 129.01, 124.81, 115.10, 103.05, 73.68, 72.99, 72.38, 71.65, 71.59, 71.57, 71.53, 71.51, 71.42, 71.38, 71.31, 71.19, 71.13, 70.44, 69.73, 69.03, 65.55, 59.10, 51.77, 51.31, 40.76, 40.72, 30.74, 30.70, 30.58, 30.52, 30.42, 27.23, 10.96. HRMS (ESI+): Found: 1357.6988m/z Calc: 1357.6952m/z for $[C_{60}H_{100}O_{18}N_8F_6Na]^+$.

PCS: **PCS** was prepared fresh from **CS12-DITFAA** 1-2 days before its use and kept at -80^{o}C. It was prepared by adding **CS12-DITFAA** (2.6mg) to a solution of NaOH (aq., 6M, 4.3µL). The mixture was stirred for 5 minutes before the addition of 22µL 1,4-dioxane. The mixture was allowed to sit at room temperature for 2 hours before the addition of phosphate buffer (233.7µL, 0.2M at pH 8.3) to yield a 10mg/mL stock solution which was used directly in all subsequent assays. The formation of PCS was confirmed by TLC and HRMS. HRMS (ESI+): Found: 1143.7493m/z Calc: 1143.7487m/z for $[C_{56}H_{103}O_{16}N_8]^+$ (n=6).

Synthesis of TCS

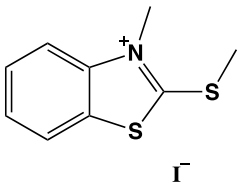


Scheme 10: Synthetic scheme for the preparation of TO-PEG4-N3.



8

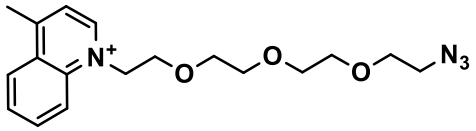
2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (**8**): Compound **8** was synthesized according to previously reported procedures. Briefly, to a solution of **6** (2.55g, 11.6mmol, 1eq) in DCM (100mL) was added TsCl (5.6249g, 29.5mmol, 2.5eq) and TEA (4.1mL, 2.97g, 29.4mmol, 2.5eq). The reaction was allowed to stir at room temperature overnight. The solution was then added saturated NH4Cl (50mL), and the organic phase was collected. The organic phase was washed with brine (50mL), and the organic phase was further dried with Na2SO4 and concentrated by rotary evaporation. The resulting crude was purified by silica flash chromatography to yield **8** as a colorless oil (3.0558g, 8.2mmol, 71% yield). The NMR and HRMS data matched with that previously reported.



9

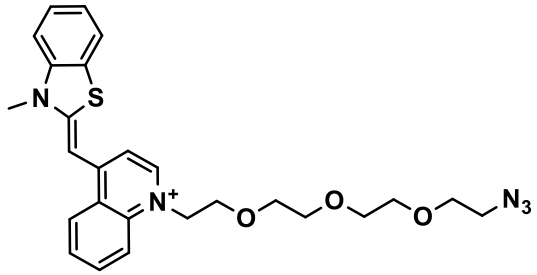
3-methyl-2-(methylthio)benzo[d]thiazol-3-ium iodide (**9**): To a round bottom flask containing 2-(methylthio)benzo[d]thiazole (0.5g, 2.76mmol, 1eq) was added iodomethane (515µL, 1.17g, 8.27mmol, 3eq), and the reaction was allowed to stir at 100°C overnight. The resulting mixture

was filtered, and the solid was washed with cold diethyl ether. The residue was isolated to give **9** as an off-white fine powder (256.1mg, 0.79mmol, 29% yield). ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.39 (d, *J* = 8.1 Hz, 1H), 8.20 (d, *J* = 8.4 Hz, 1H), 7.92 – 7.79 (m, 1H), 7.73 (t, *J* = 7.7 Hz, 1H), 4.11 (s, 3H), 3.12 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 181.25, 142.54, 129.17, 128.26, 127.01, 124.00, 115.73, 36.51, 18.14. HRMS (ESI+): Found: 196.0249m/z Calc: 196.0249m/z for $[C_9H_{10}NS_2]^+$.



10

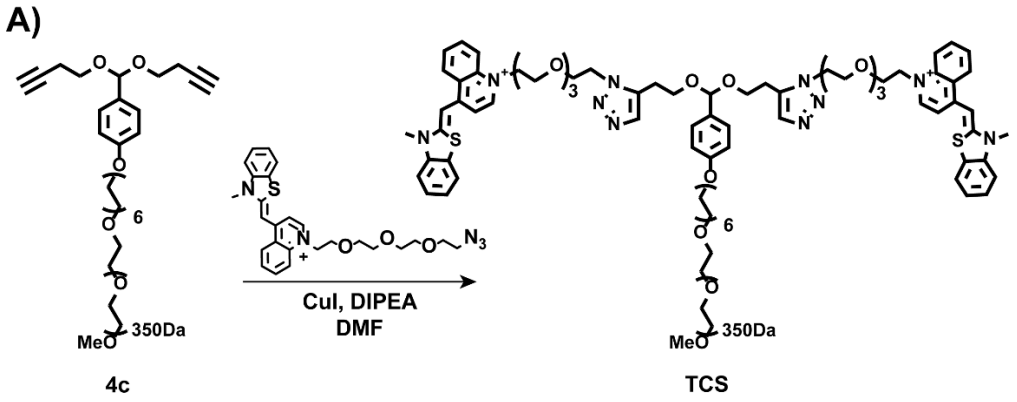
1-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-4-methylquinolin-1-ium iodide (**10**): To a solution of lepidine (2.6mL, 2.82g, 19.7mmol, 5eq) was added **8** (1.5009g, 4mmol, 1eq) in ethyl acetate (4mL), and the reaction was allowed to stir at 50°C overnight. The mixture was then concentrated, and directly purified by neutral alumina flash chromatography (2-4% MeOH in DCM) to yield **10** as a blue/purple oil (727mg, 2.1mmol, 53% yield). ^{1}H NMR (900 MHz, MeOD) δ 9.20 (d, *J* = 6.0 Hz, 1H), 8.58 (dd, *J* = 8.5, 1.3 Hz, 2H), 8.26 (ddd, *J* = 8.6, 6.9, 1.4 Hz, 1H), 8.11 – 8.03 (m, 1H), 7.98 (d, *J* = 6.0 Hz, 1H), 5.24 (t, *J* = 4.9 Hz, 2H), 4.11 – 4.03 (m, 2H), 3.63 – 3.58 (m, 4H), 3.55 – 3.50 (m, 4H), 3.50 – 3.48 (m, 2H), 3.34 (t, *J* = 4.9 Hz, 2H), 3.10 – 3.06 (m, 3H). ^{13}C NMR (226 MHz, MeOD) δ 161.06, 150.29, 138.92, 136.50, 130.98, 130.93, 128.28, 123.43, 120.26, 71.73, 71.56, 71.46, 71.02, 69.20, 58.39, 54.81, 51.75, 20.30. HRMS (ESI+): Found: 345.1919m/z Calc: 345.1921m/z for $[C_{18}H_{25}N_4O_3]^+$.



TO-PEG$_4$-N$_3$

(Z)-1-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-4-((3-methylbenzo[d]thiazol-2(3H)-ylidene)methyl)quinolin-1-ium iodide (**TO-PEG4-N3**): To a solution of **9** (99.8mg, 0.31mmol, 1.07eq) and **10** (99.6mg, 0.29mmol, 1eq) in EtOH (4mL) in a nitrogen-purged pressure tube was added TEA (120µL,). The reaction was allowed to stir at 60°C for two hours, and the reaction was

then concentrated and purified directly by neutral alumina flash chromatography (0-2% MeOH in DCM) to yield **TO-PEG4-N3** as a red tar (89.6mg, 0.18mmol, 62% yield). ^{1}H NMR (900 MHz, MeOD) δ 8.68 (t, *J* = 8.0 Hz, 1H), 8.43 – 8.38 (m, 1H), 8.13 – 8.07 (m, 1H), 7.99 – 7.93 (m, 1H), 7.90 (t, *J* = 7.8 Hz, 1H), 7.79 – 7.73 (m, 1H), 7.66 (dd, *J* = 14.8, 8.2 Hz, 1H), 7.64 – 7.57 (m, 1H), 7.50 – 7.44 (m, 1H), 7.44 – 7.39 (m, 1H), 6.95 – 6.89 (m, 1H), 4.81 – 4.75 (m, 2H), 4.03 – 3.99 (m, 3H), 3.97 (t, *J* = 5.1 Hz, 2H), 3.64 – 3.59 (m, 2H), 3.58 – 3.54 (m, 4H), 3.52 – 3.48 (m, 4H), 3.28 (t, *J* = 4.9 Hz, 2H). ^{13}C NMR (226 MHz, MeOD) δ 159.23, 148.17, 143.21, 139.03, 136.03, 131.39, 126.54, 125.13, 123.70, 123.21, 122.95, 122.74, 120.73, 116.03, 110.76, 106.33, 86.26, 68.89, 68.68, 68.62, 68.53, 68.10, 66.45, 52.66, 48.75, 31.17. HRMS (ESI+): Found: 492.2058m/z Calc: 492.2064m/z for $[C_{26}H_{30}N_5O_3S]^+$.



Scheme 11: Synthetic scheme for the preparation of TCS.

TCS: A mixture of **4c** (10.3mg, 0.014 mmol, 1eq) in DMF (200µL) and DIPEA (6.3µL, 4.6mg, 0.036mmol, 2.5eq) was added to a round bottomed flask, and the solution was purged with nitrogen gas. The solution was then added TO-PEG4-N3 (33mg, 0.07mmol, 5eq) and CuI (cat. amount), and the mixture was allowed to stir at room temperature for 48 hours. The reaction mixture was added toluene (50mL), and the solution was concentrated by rotary evaporation, and the residue was purified by neutral alumina flash chromatography (2-4% MeOH in DCM + 0.1% TEA) to yield **TCS** as a red tar (9mg, 0.005mmol, 38% yield). TCS was stored in the presence of TEA.

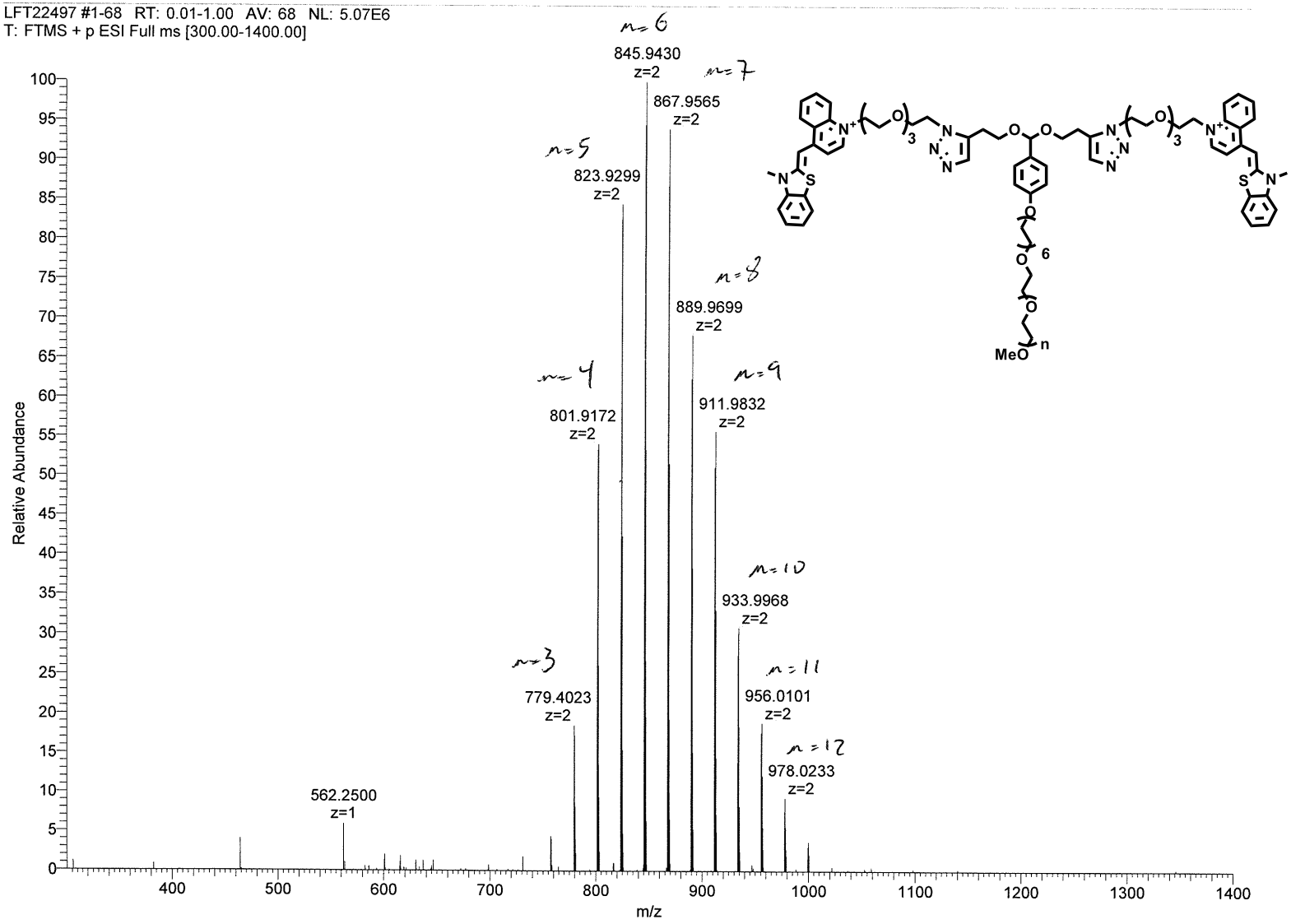
C8-TCS *was prepared and purified using the same protocol. HRMS spectra is shown in the next section. C8-TCS was only used for the DNAse protection assay.*

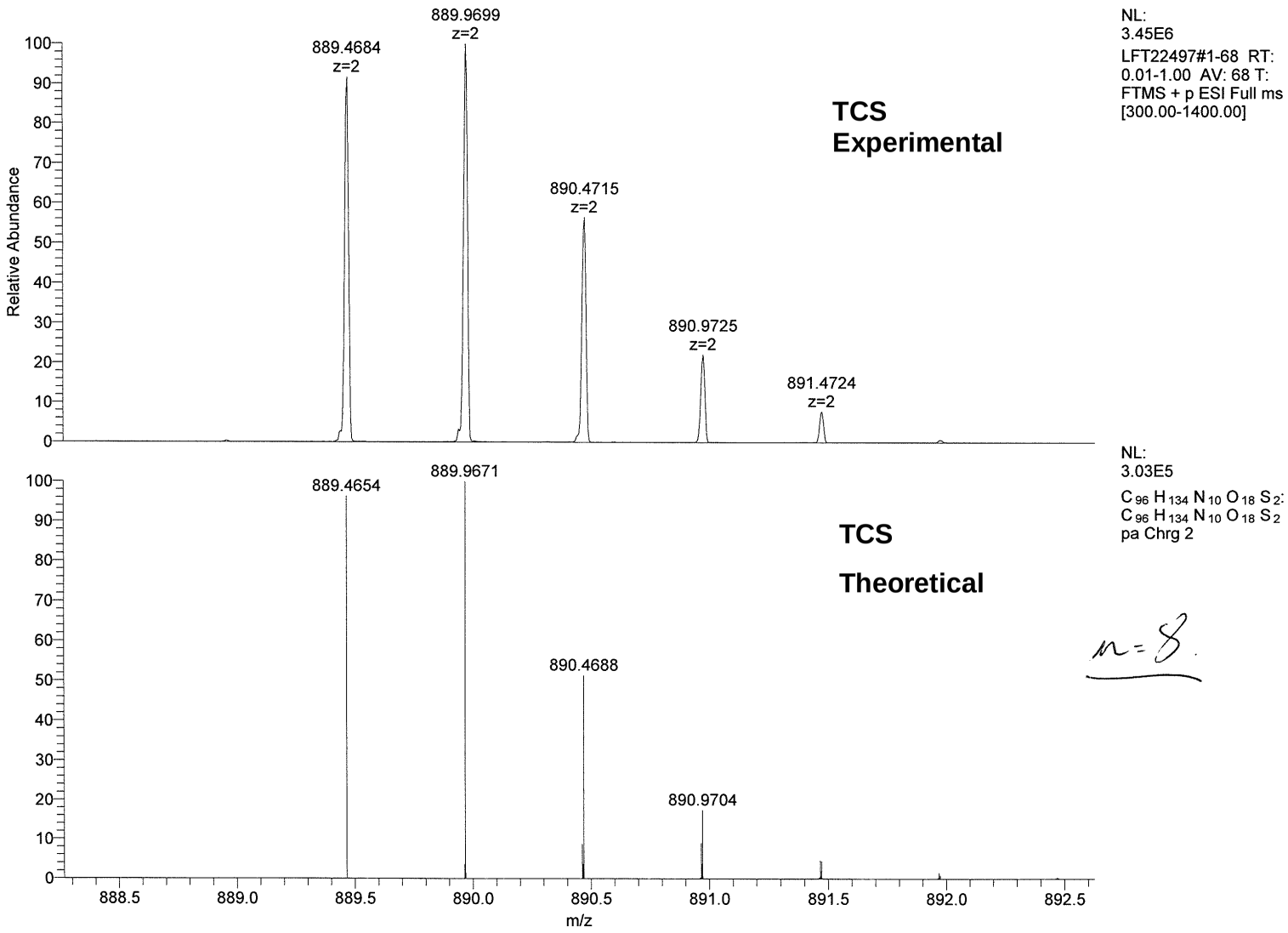
^{1}H NMR (900 MHz, MeOD) δ 8.52 (d, *J* = 8.4 Hz, 2H), 8.29 (d, *J* = 7.1 Hz, 2H), 7.96 (d, *J* = 8.6 Hz, 2H), 7.85 (t, *J* = 7.7 Hz, 2H), 7.79 (d, *J* = 7.8 Hz, 2H), 7.69 (s, 2H), 7.66 (t, *J* = 7.5 Hz, 2H), 7.54 – 7.47 (m, 2H), 7.32 (t, *J* = 7.1 Hz, 2H), 7.22 (d, *J* = 7.0 Hz, 2H), 7.13 (d, *J* = 8.0 Hz, 2H), 6.72 (d, *J* = 8.2 Hz, 2H), 6.70 (s, 2H), 5.50 (s, 1H), 4.65 (d, *J* = 6.0 Hz, 4H), 4.43 (t, *J* = 5.1 Hz, 4H), 3.90 (t, *J* = 4.9 Hz, 4H), 3.87 (s, 6H), 3.80 (t, *J* = 6.4 Hz, 2H), 3.77 (t, *J* = 5.1 Hz, 4H), 3.65 – 3.43 (m, 55H), 2.80 (t, *J* = 6.5 Hz, 4H), 1.65 (p, *J* = 6.6 Hz, 2H), 1.54 (p, *J* = 6.8 Hz, 2H).

^{13}C NMR (226 MHz, MeOD) δ 161.91, 150.61, 146.29, 146.10, 141.86, 138.91, 134.48, 129.62, 129.10, 128.25, 126.72, 126.02, 125.76, 124.81, 123.84, 119.07, 115.17, 113.84, 109.28, 103.09, 89.29, 78.36, 73.13, 72.51, 71.94, 71.73, 71.65, 71.55, 71.50, 71.34, 70.49, 69.45, 69.12, 65.71, 59.25, 56.40, 55.70, 53.75, 51.38, 34.36, 30.90, 30.84, 30.74, 30.64, 30.51, 27.38, 27.32, 7.77.

HRMS (ESI+): Found: 889.9699 m/z Calc: 889.9671m/z for $[C_{96}H_{134}N_{10}O_{18}S_2]^{2+}$ for n=8 (PEG chain length).

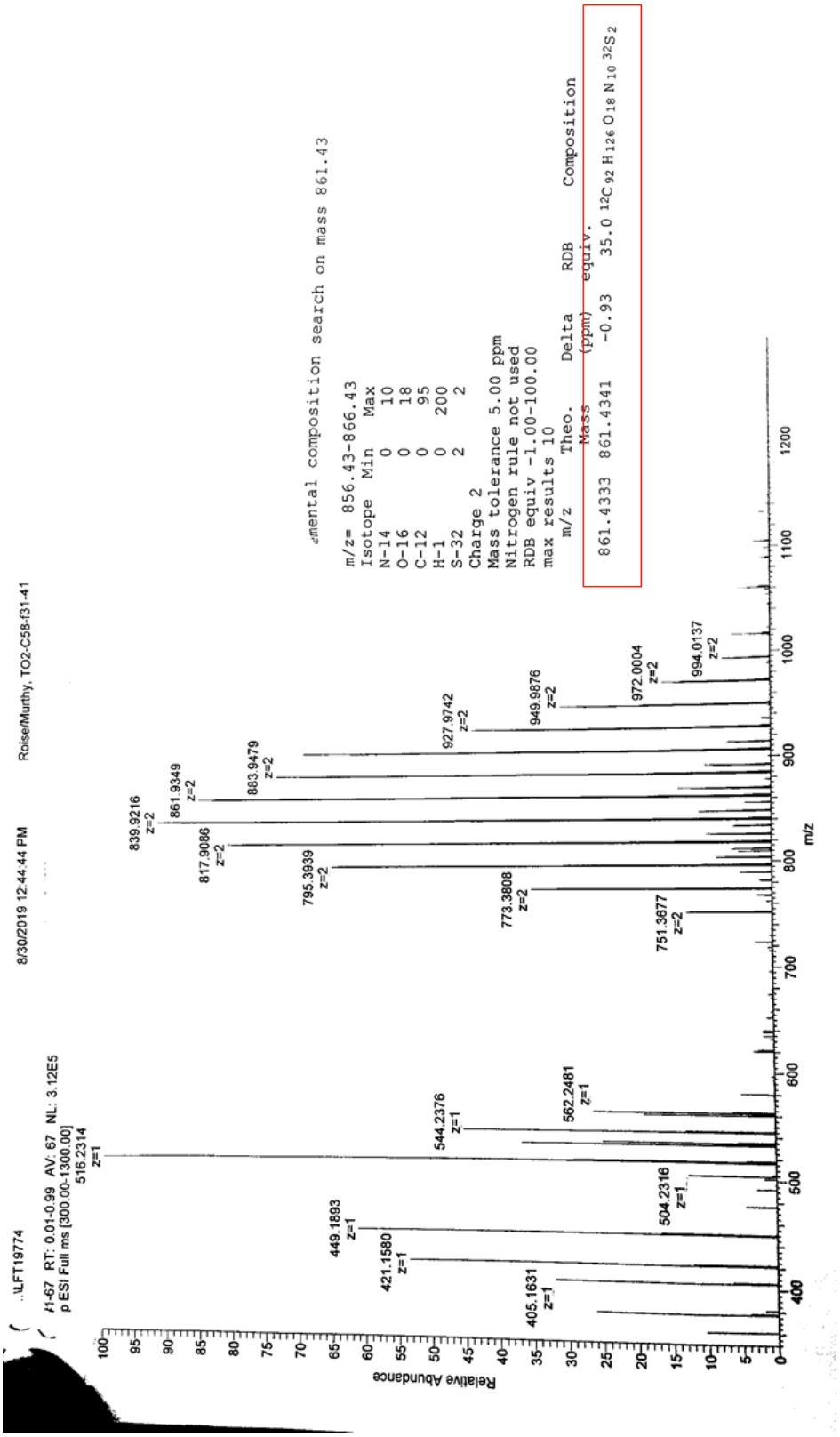
LFT22497 #1-68 RT: 0.01-1.00 AV: 68 NL: 5.07E6
T: FTMS + p ESI Full ms [300.00-1400.00]



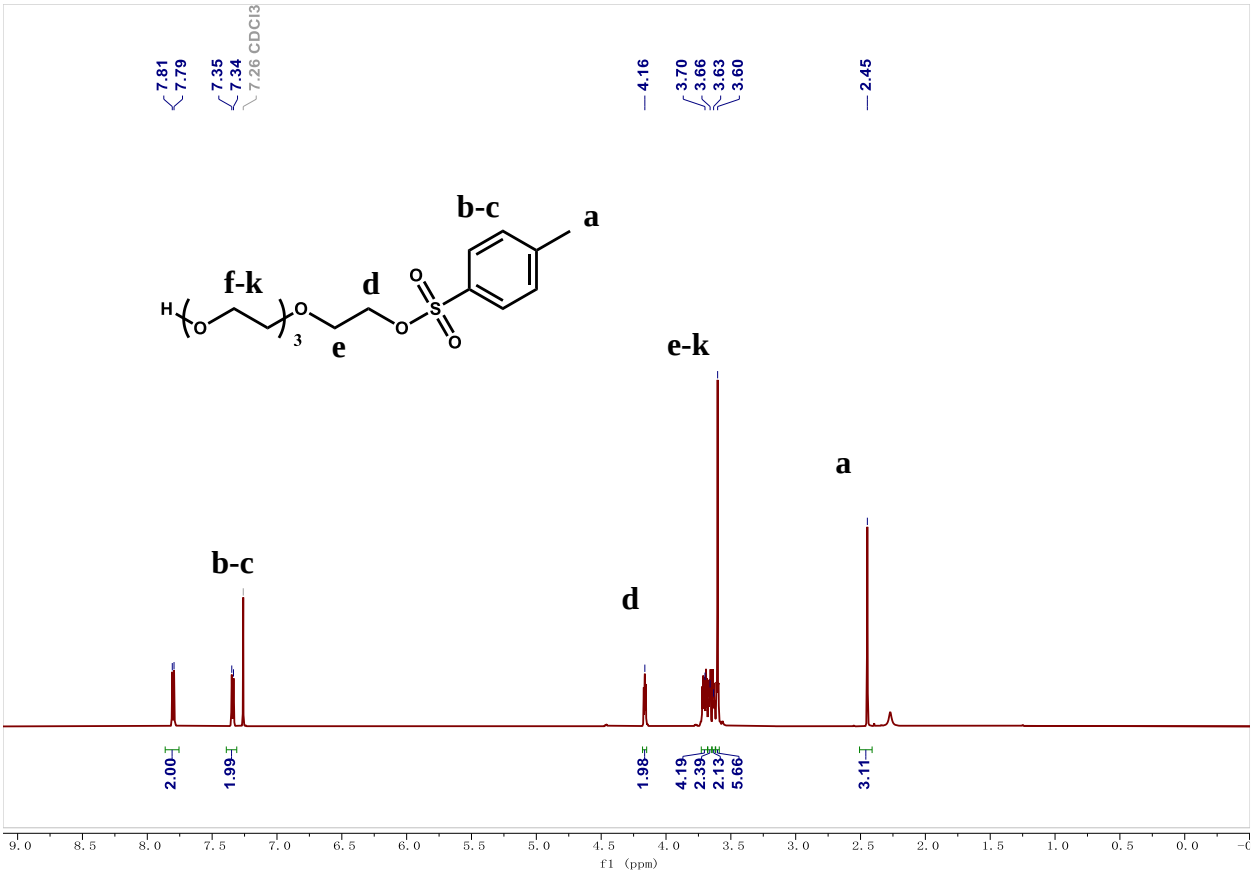

889.9699
z=2
889.4684
z=2
890.4715
z=2
890.9725
z=2
891.4724
z=2
Relative Abundance
TCS
Experimental
NL:
3.45E6
LFT22497#1-68 RT:
0.01-1.00 AV: 68 T:
FTMS + p ESI Full ms
[300.00-1400.00]
889.9671
889.4654
890.4688
890.9704
TCS
Theoretical
NL:
3.03E5
C96 H134 N10 O18 S2:
C96 H134 N10 O18 S2
pa Chrg 2
888.5
889.0
889.5
890.0
890.5
891.0
891.5
892.0
892.5
m/z

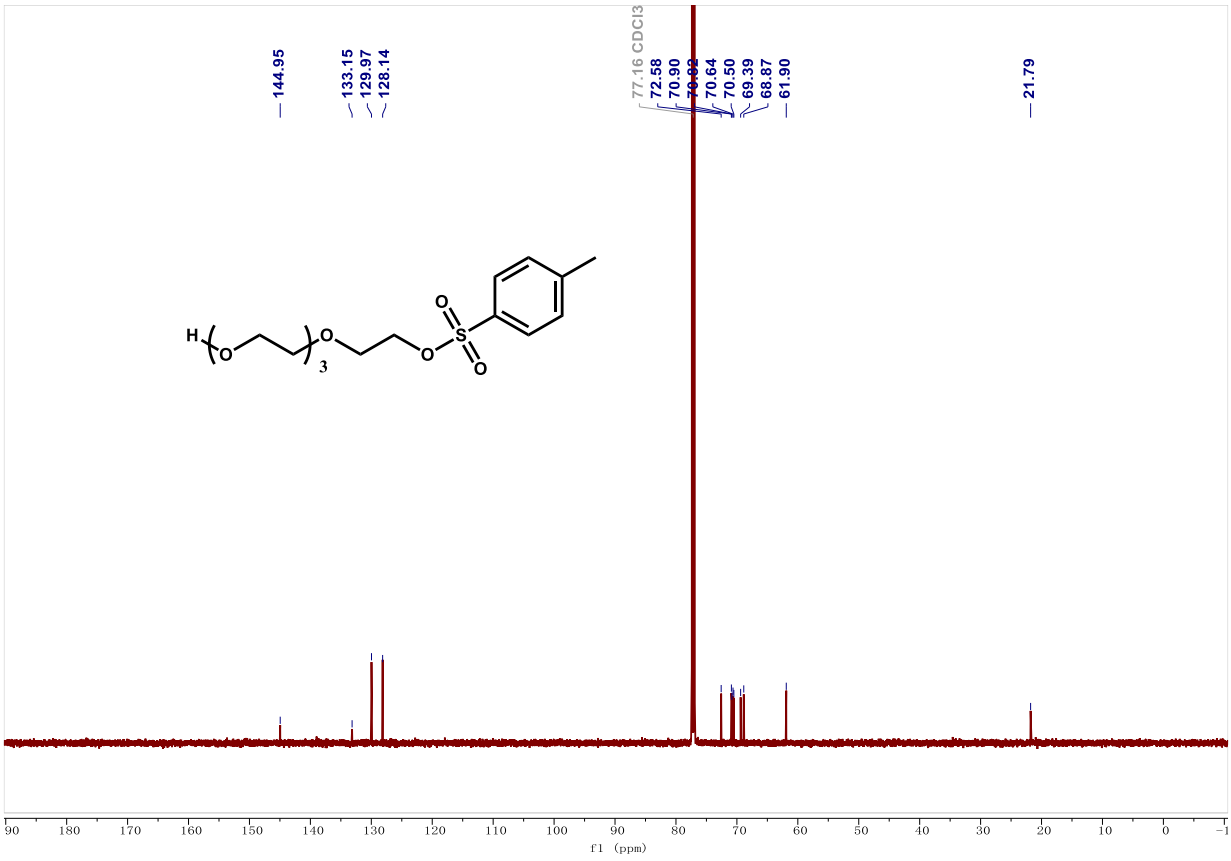
n = 8.

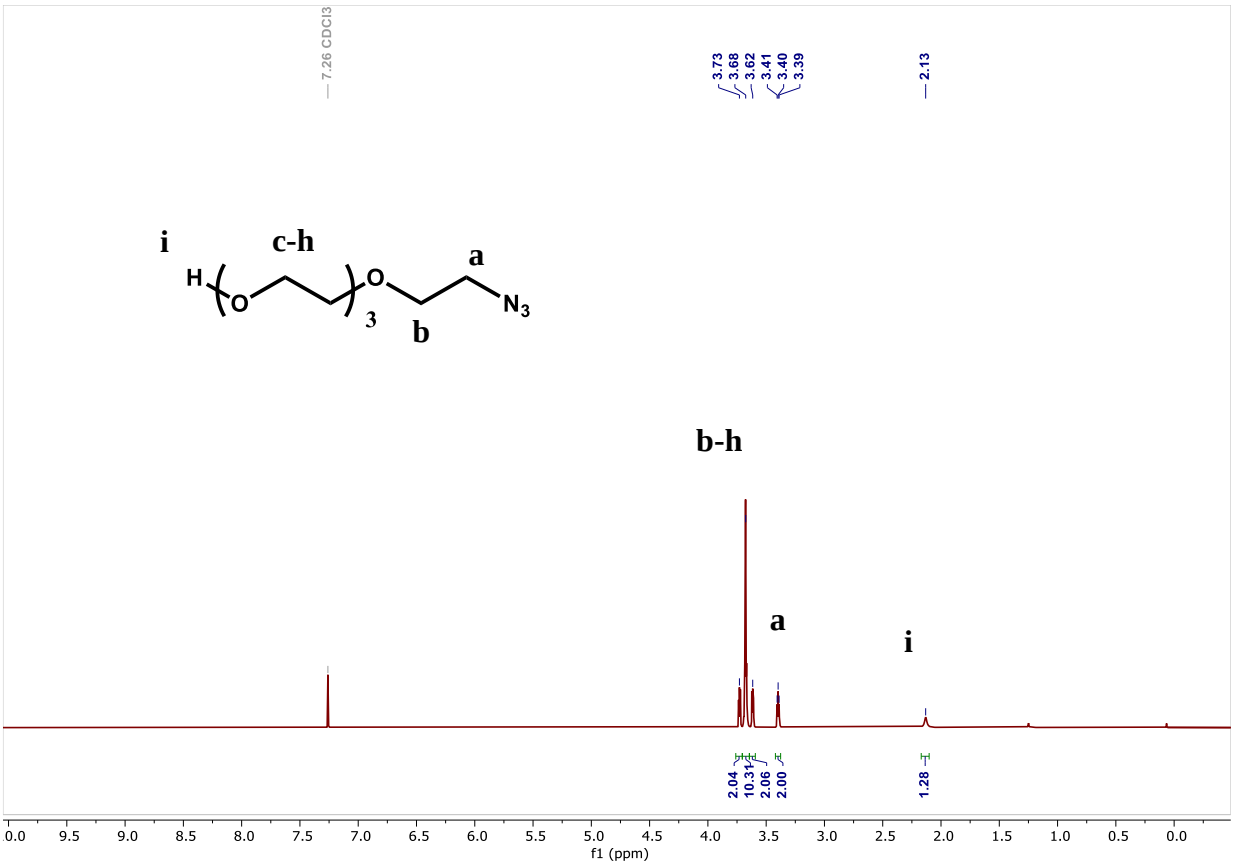
A.10 C8-TCS HRMS

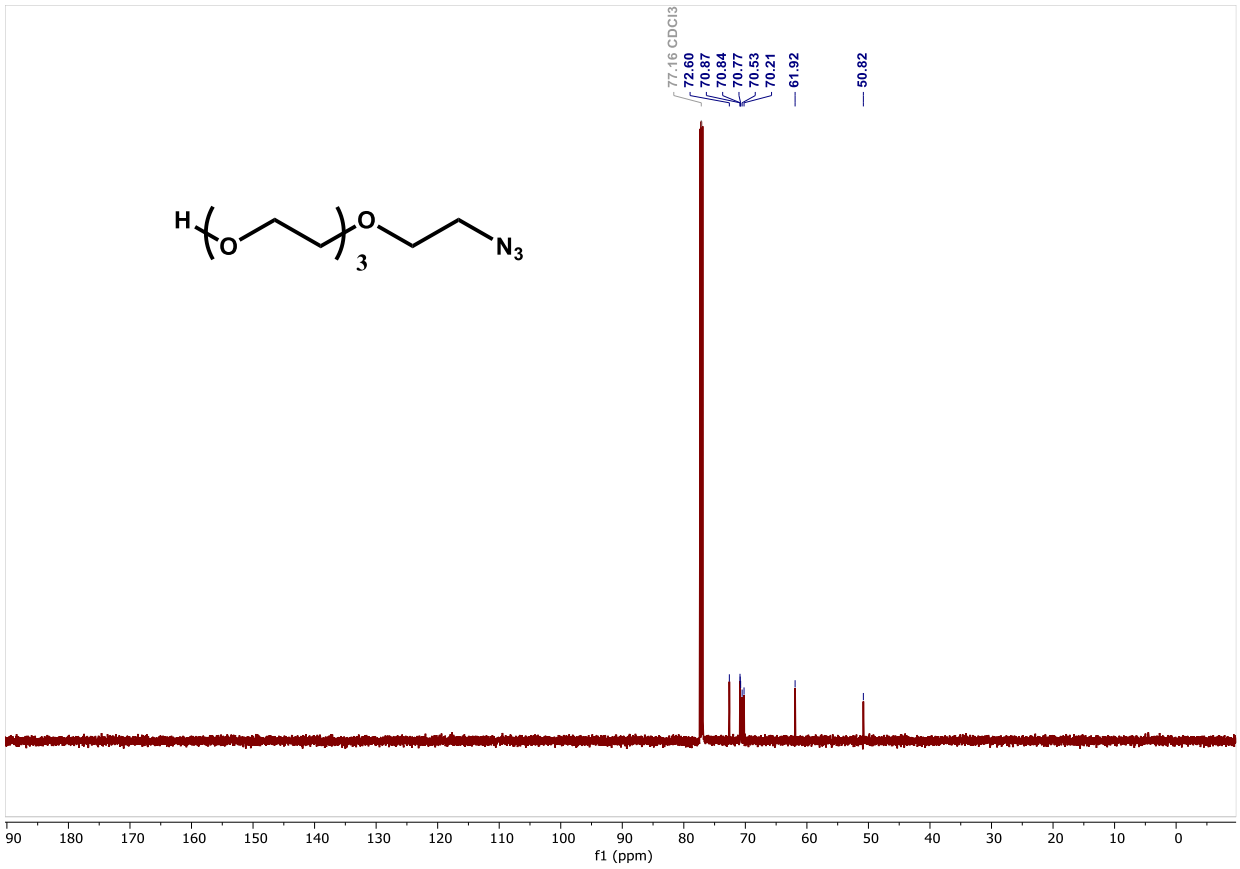


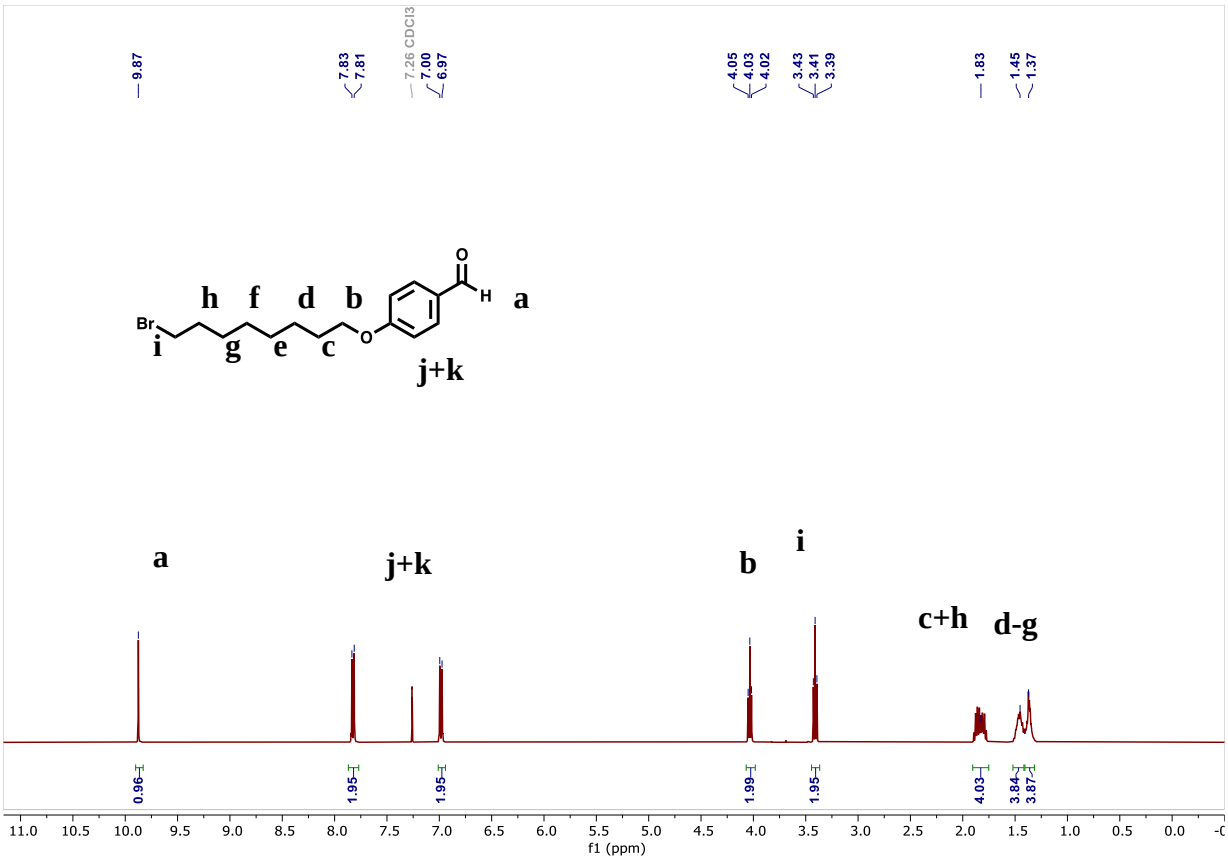
A.11 NMR Spectra

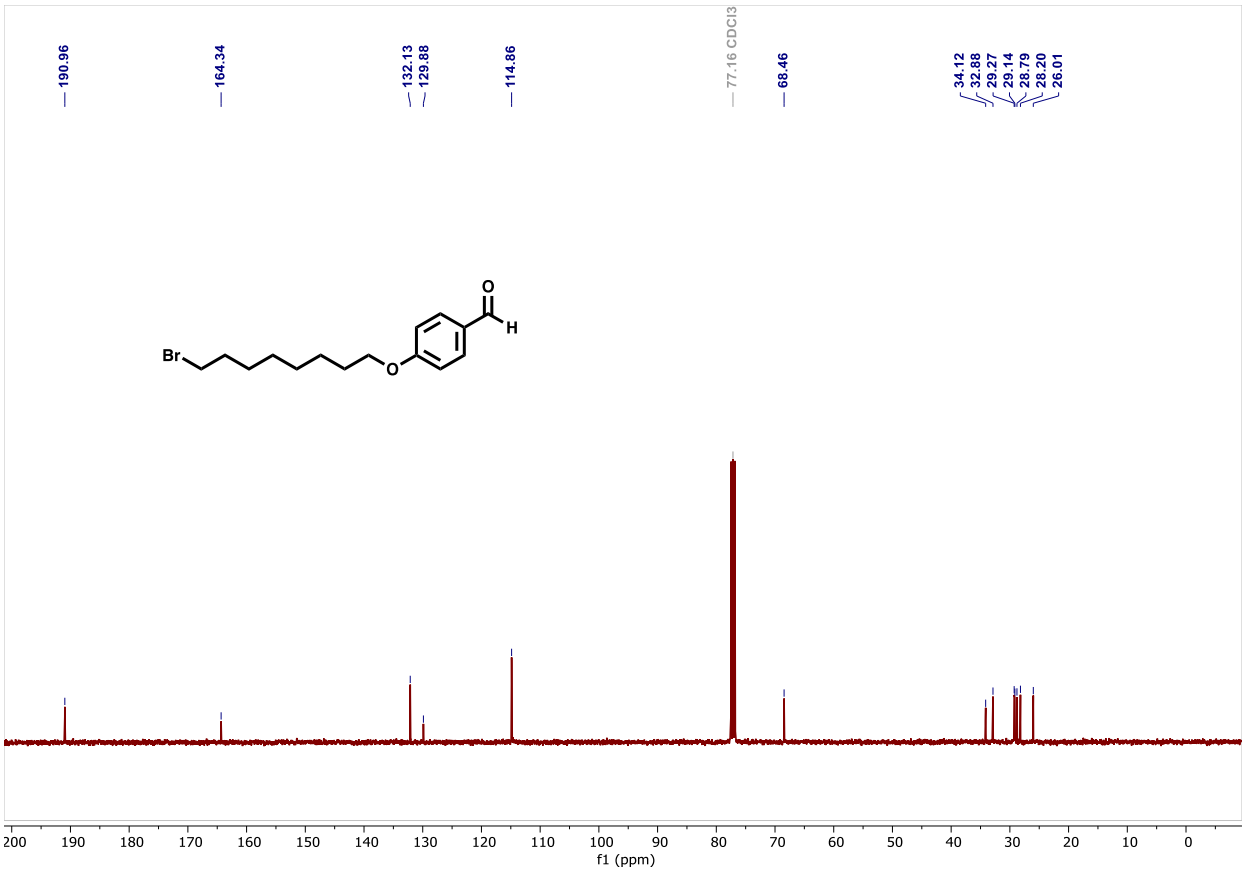


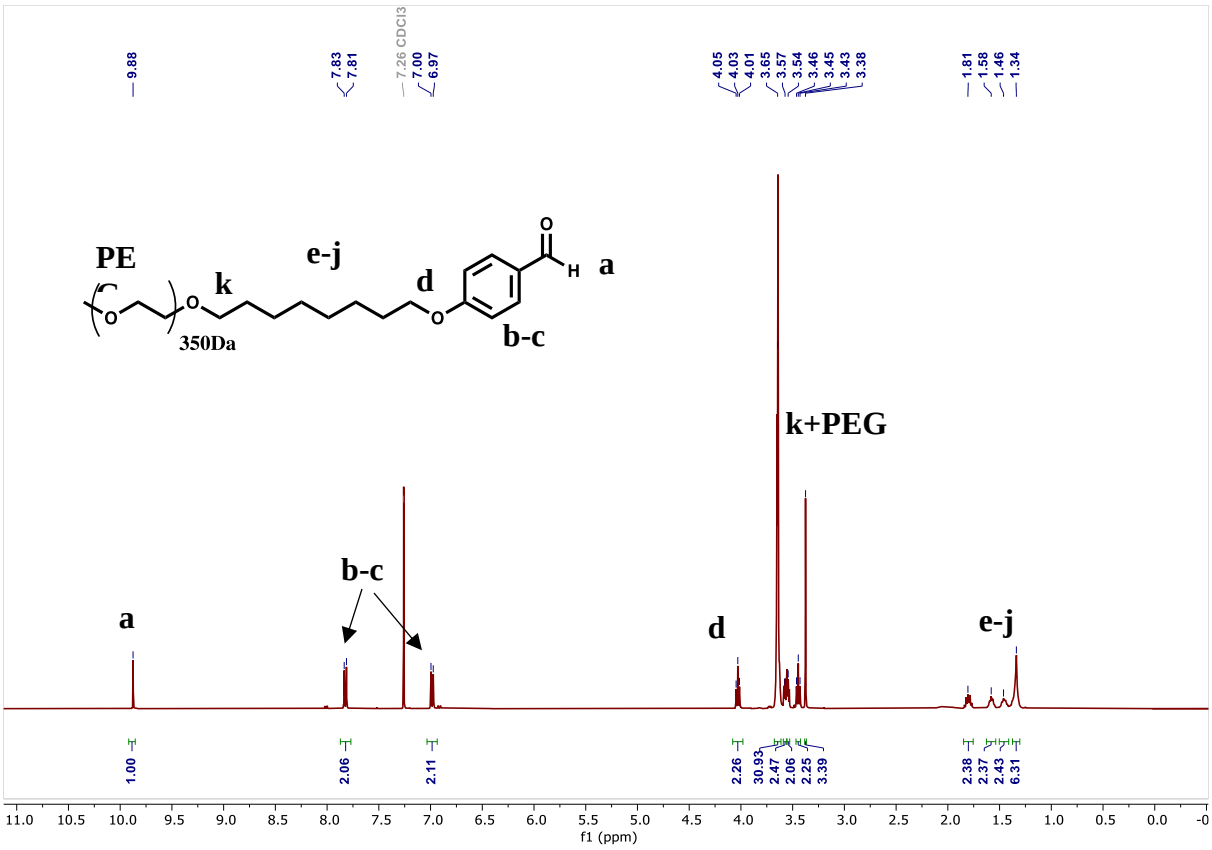

144.95
133.15
129.97
128.14
77.16 CDCl3
72.58
70.90
70.69
70.64
70.50
69.39
68.87
61.90
21.79
H
O
O
3
O
S
O
O
f1 (ppm)

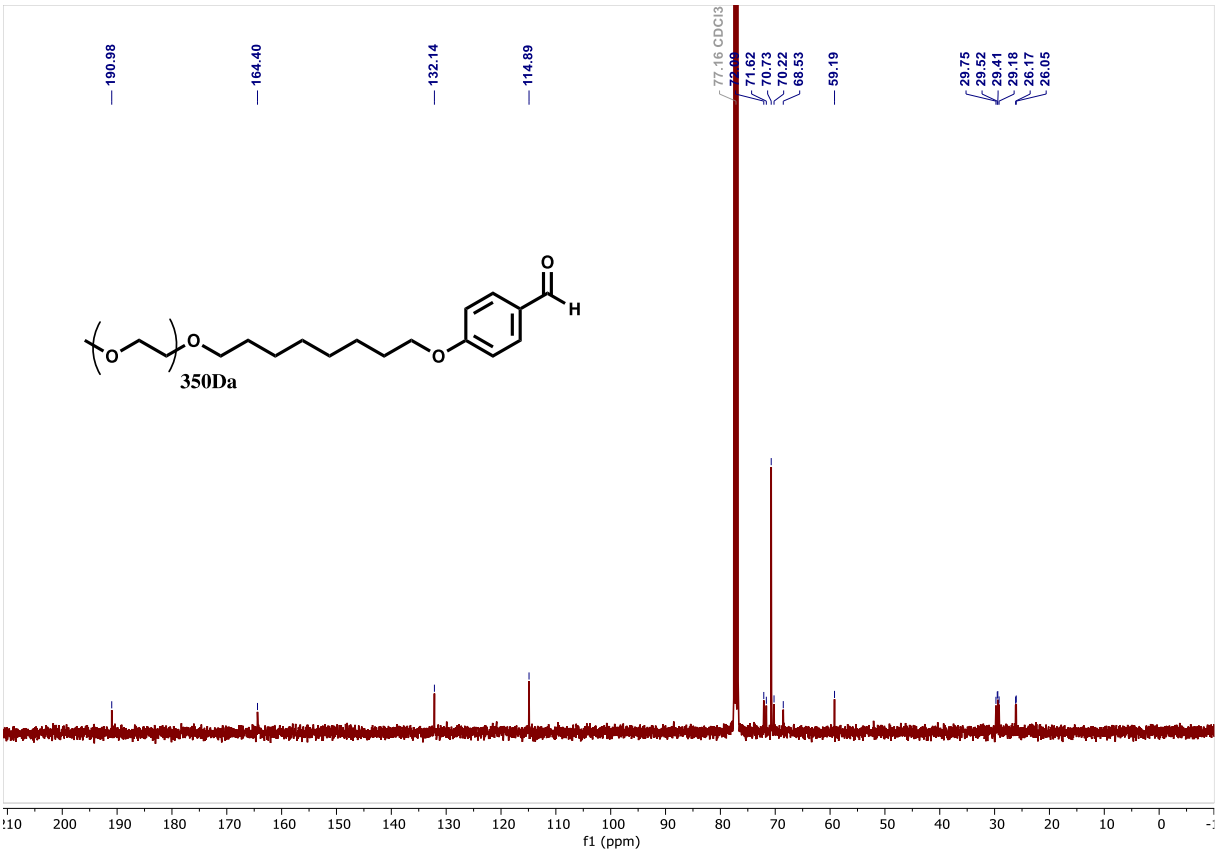

7.26 CDCl3
3.73
3.68
3.62
3.41
3.40
3.39
2.13
i
c-h
a
H
O
O
3
N3
b
b-h
a
i
2.04
10.31
2.06
2.00
1.28
0.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

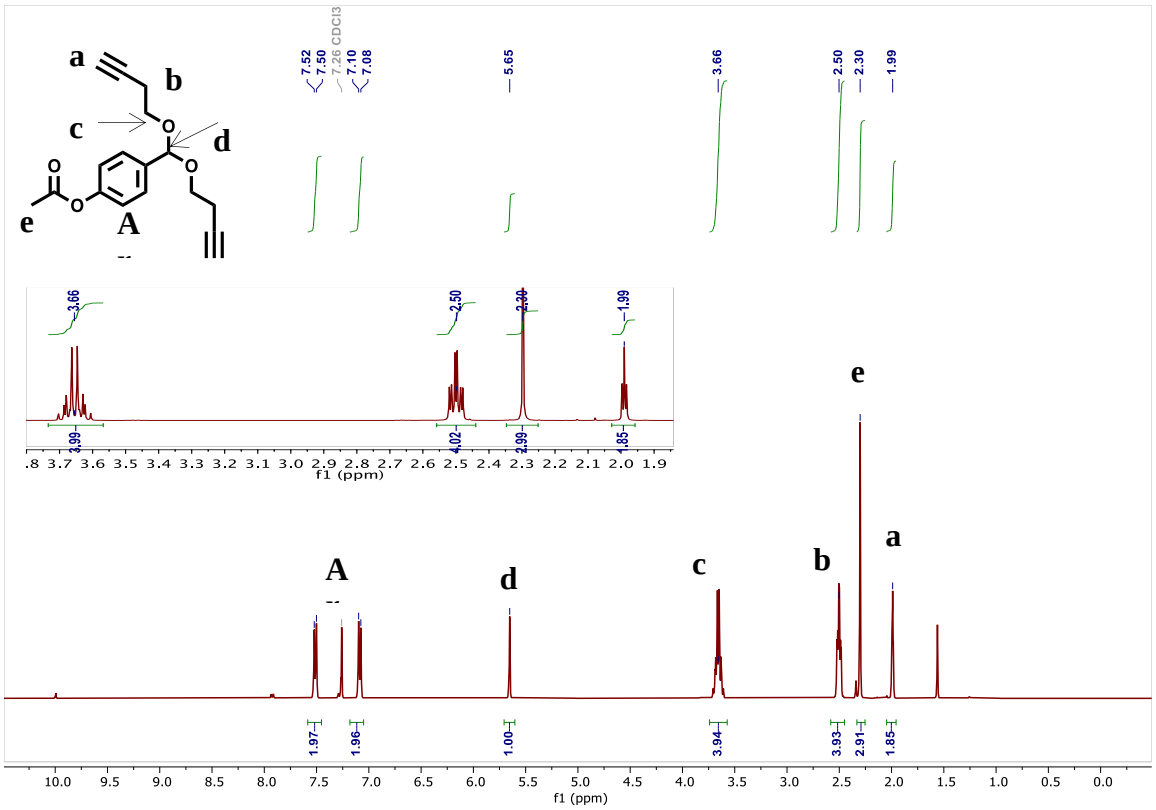

77.16 CDCl3
72.60
70.87
70.84
70.77
70.53
70.21
61.92
50.82
H
O
O
3
N3
90 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

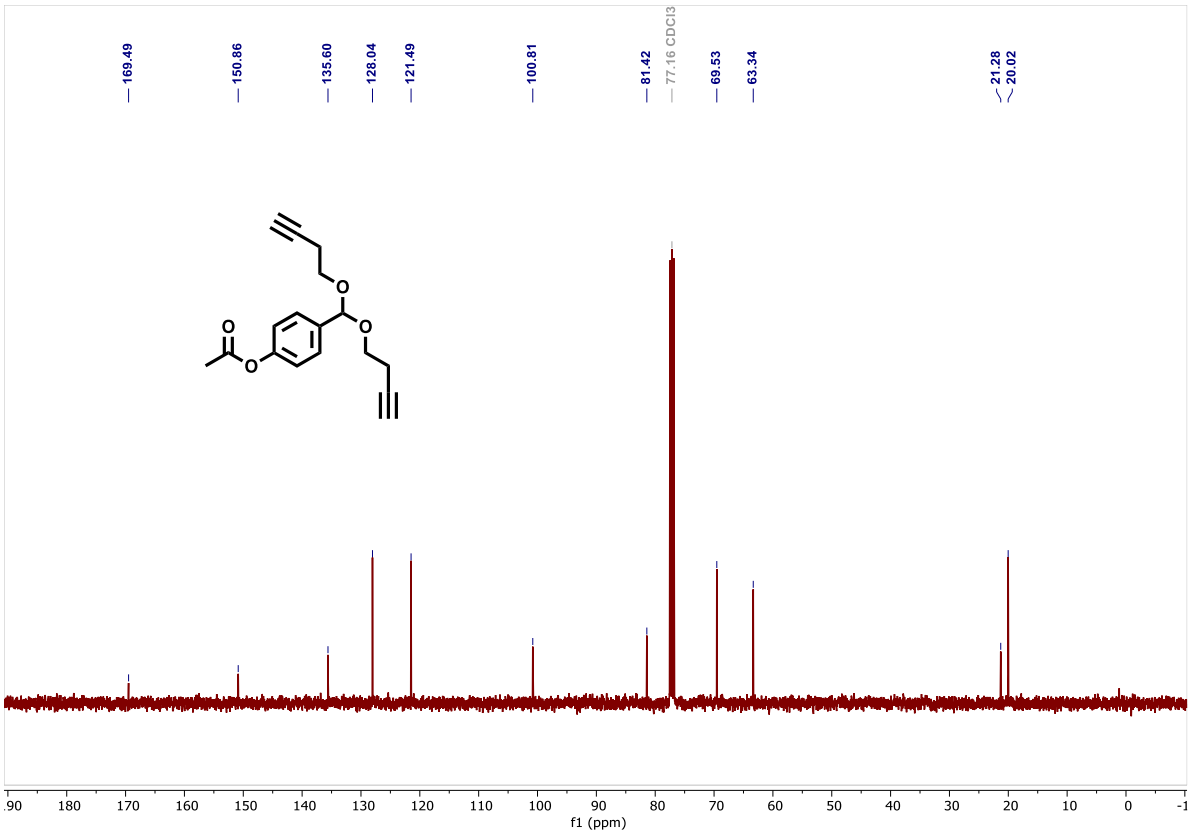

9.87
7.83
7.81
7.26 CDCl3
7.00
6.97
4.05
4.03
4.02
3.43
3.41
3.39
1.83
1.45
1.37
h f d b
a
Br
i
g e c
O
H
O
j+k
a
j+k
b
i
c+h
d-g
0.96
1.95
1.95
1.99
1.95
4.03
3.84
3.87
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

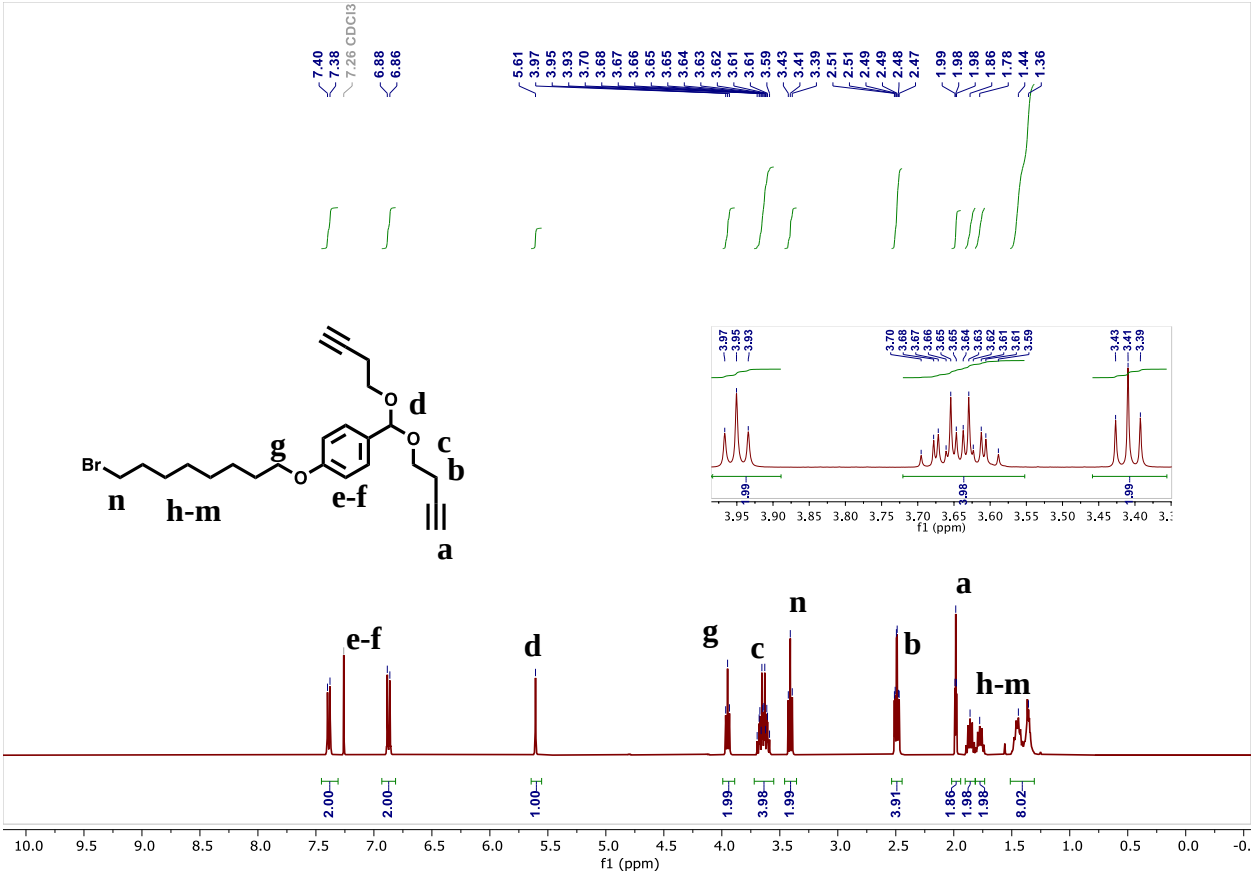

190.96
164.34
132.13
129.88
114.86
77.16 CDCl3
68.46
34.12
32.88
29.27
29.14
28.79
28.20
26.01
O
H
Br
O
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

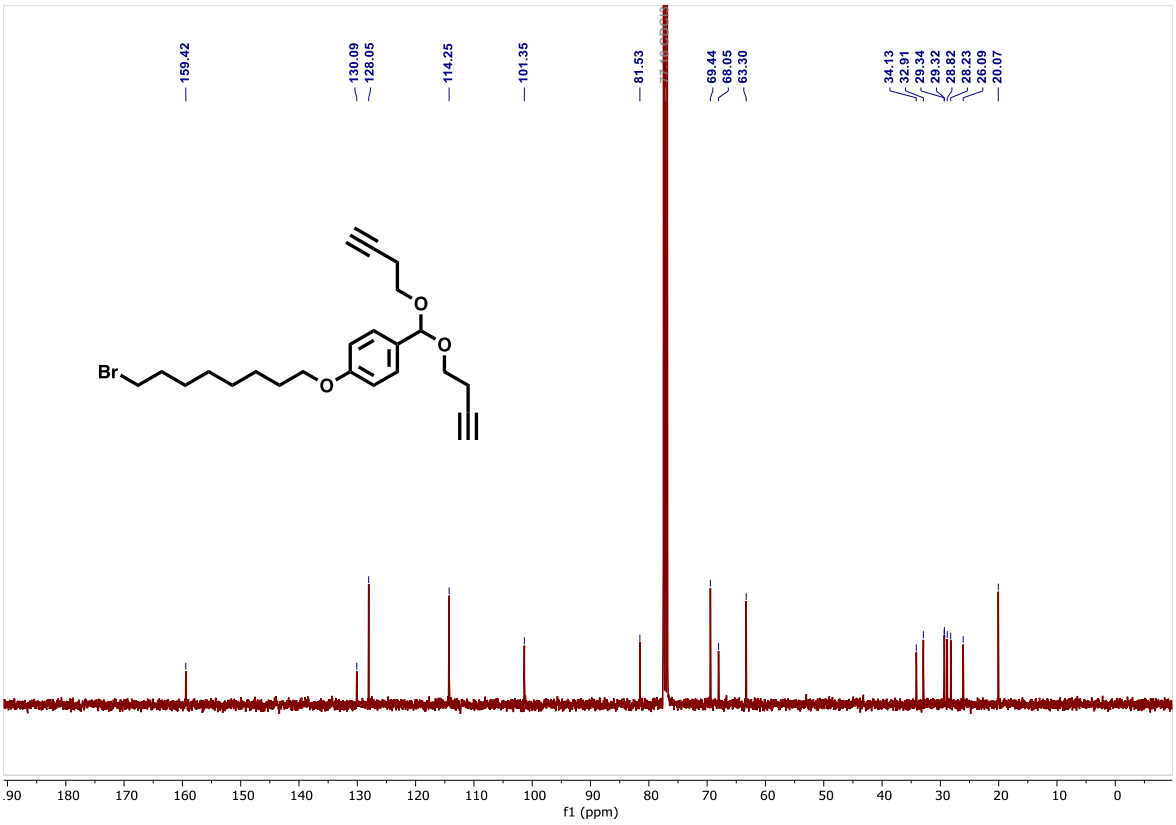

9.88
7.83
7.81
7.26 CDCl3
7.00
6.97
4.05
4.03
4.01
3.65
3.57
3.54
3.46
3.45
3.43
3.38
1.81
1.58
1.46
1.34
PE
k
e-j
d
a
b-c
350Da
k+PEG
b-c
a
d
e-j
1.00
2.06
2.11
2.26
30.93
2.47
2.06
2.25
3.39
2.38
2.37
2.43
6.31
f1 (ppm)

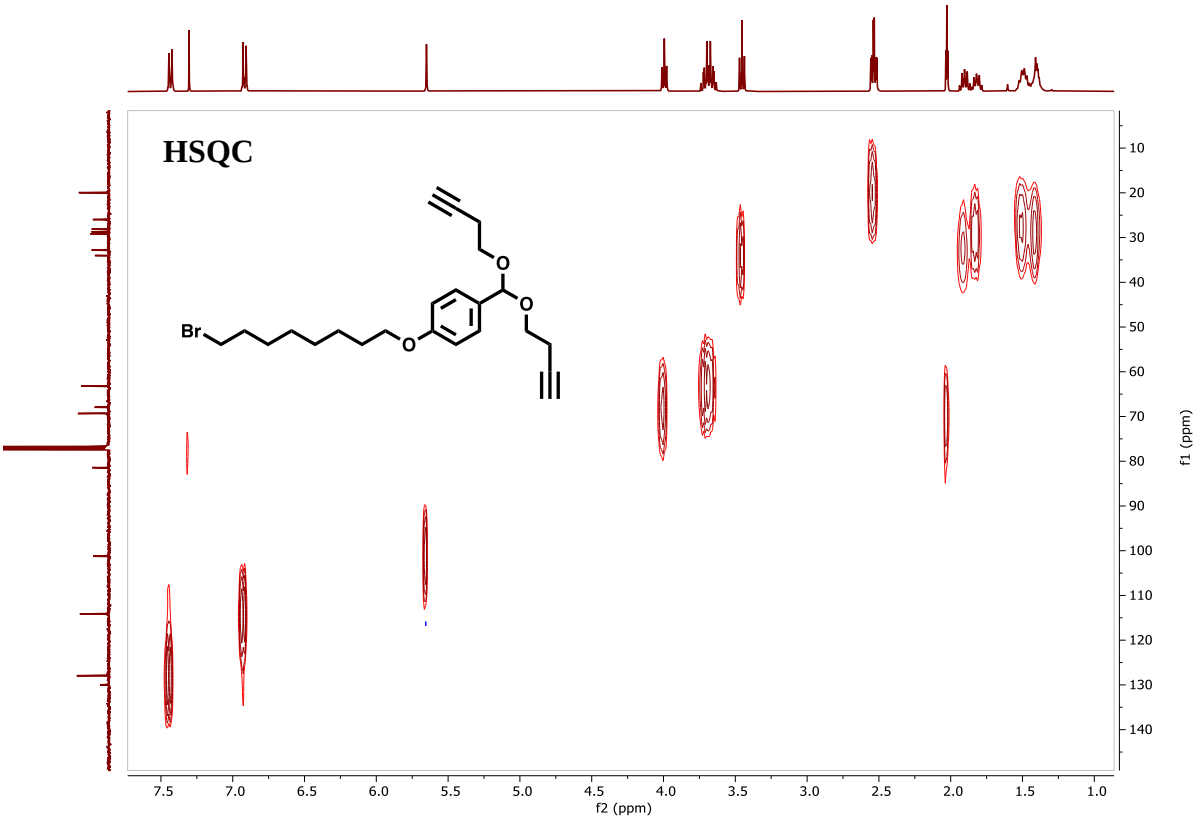

190.98
164.40
132.14
114.89
77.16 CDCl3
71.62
70.73
70.22
68.53
59.19
29.75
29.52
29.41
29.18
26.17
26.05
O
H
O
O
O
350Da
f1 (ppm)

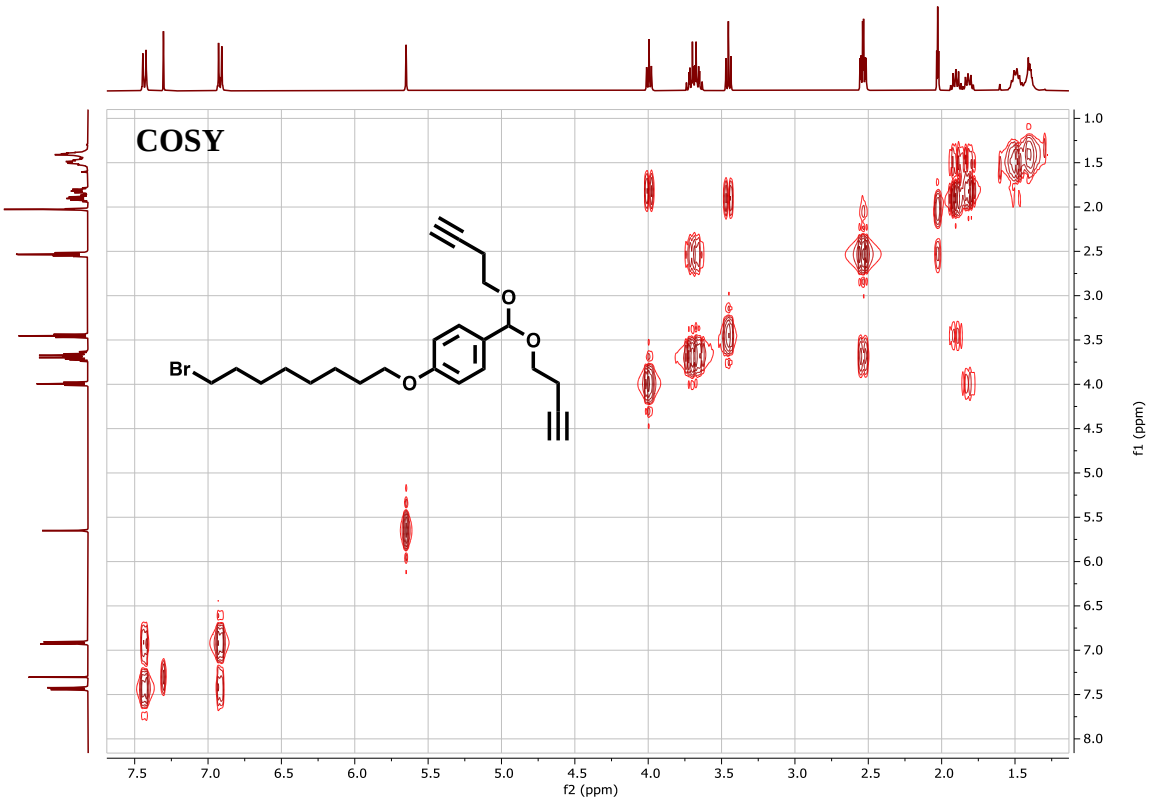

a
b
c
d
e
A
7.52
7.50
7.26 CDCl3
7.10
7.08
5.65
3.66
2.50
2.30
1.99
f1 (ppm)
1.97
1.96
1.00
3.94
3.93
2.91
1.85

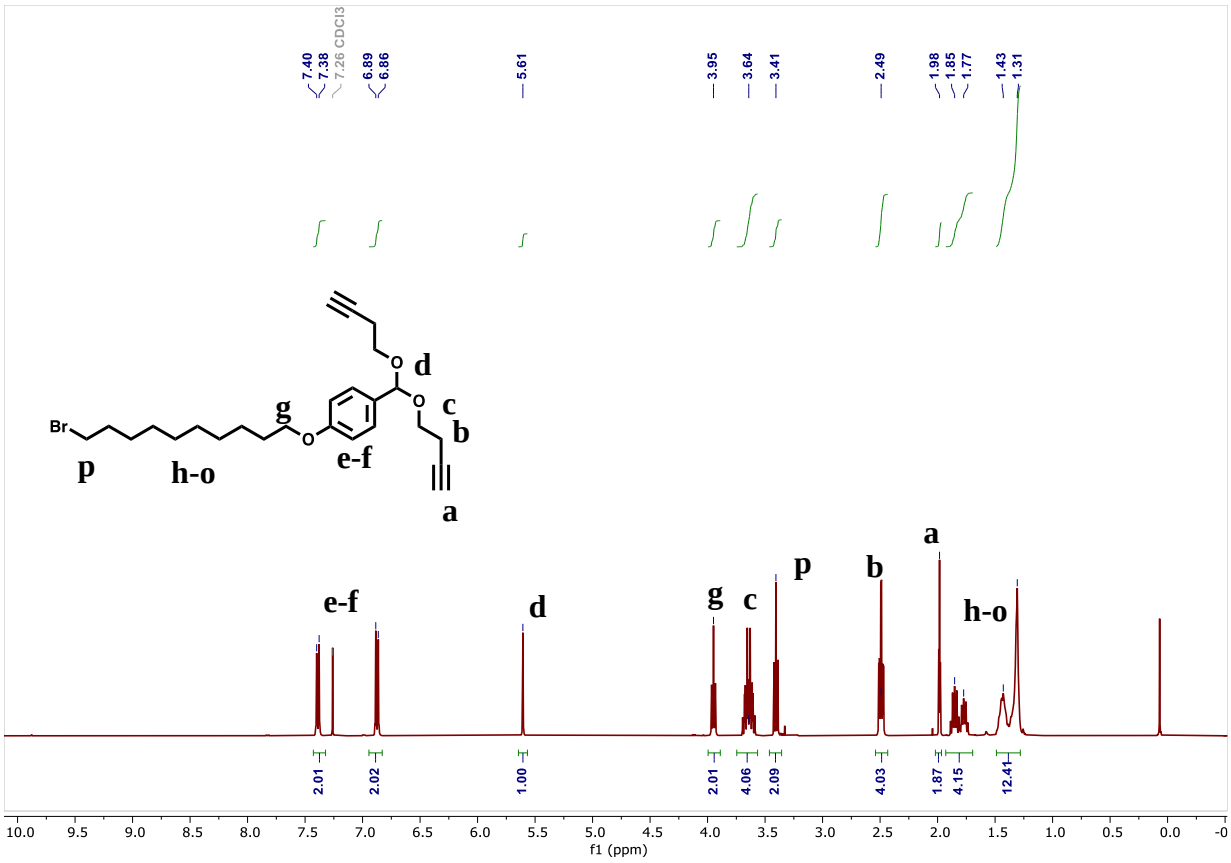

169.49
150.86
135.60
128.04
121.49
100.81
81.42
77.16 CDCl3
69.53
63.34
21.28
20.02
O
O
O
O
f1 (ppm)

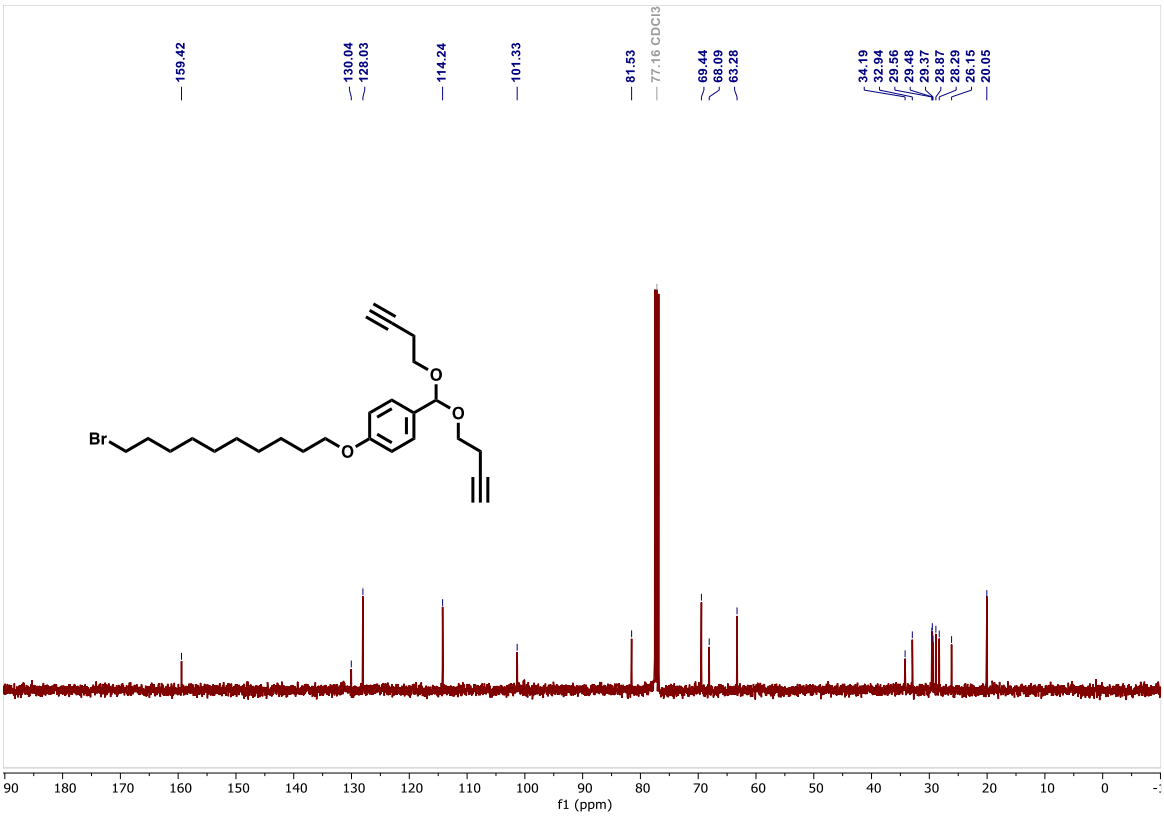

7.40
7.38
7.26 CDCl3
6.88
6.86
5.61
3.97
3.95
3.93
3.70
3.68
3.67
3.66
3.65
3.65
3.64
3.63
3.62
3.61
3.61
3.59
3.43
3.41
3.39
2.51
2.49
2.49
2.48
2.47
1.99
1.98
1.98
1.86
1.78
1.44
1.36
Br
n
h-m
g
e-f
O
d
O
c
b
a
f1 (ppm)
e-f
d
g
c
n
b
a
h-m
2.00
2.00
1.00
1.99
3.98
1.99
3.91
1.86
1.98
1.98
8.02
f1 (ppm)

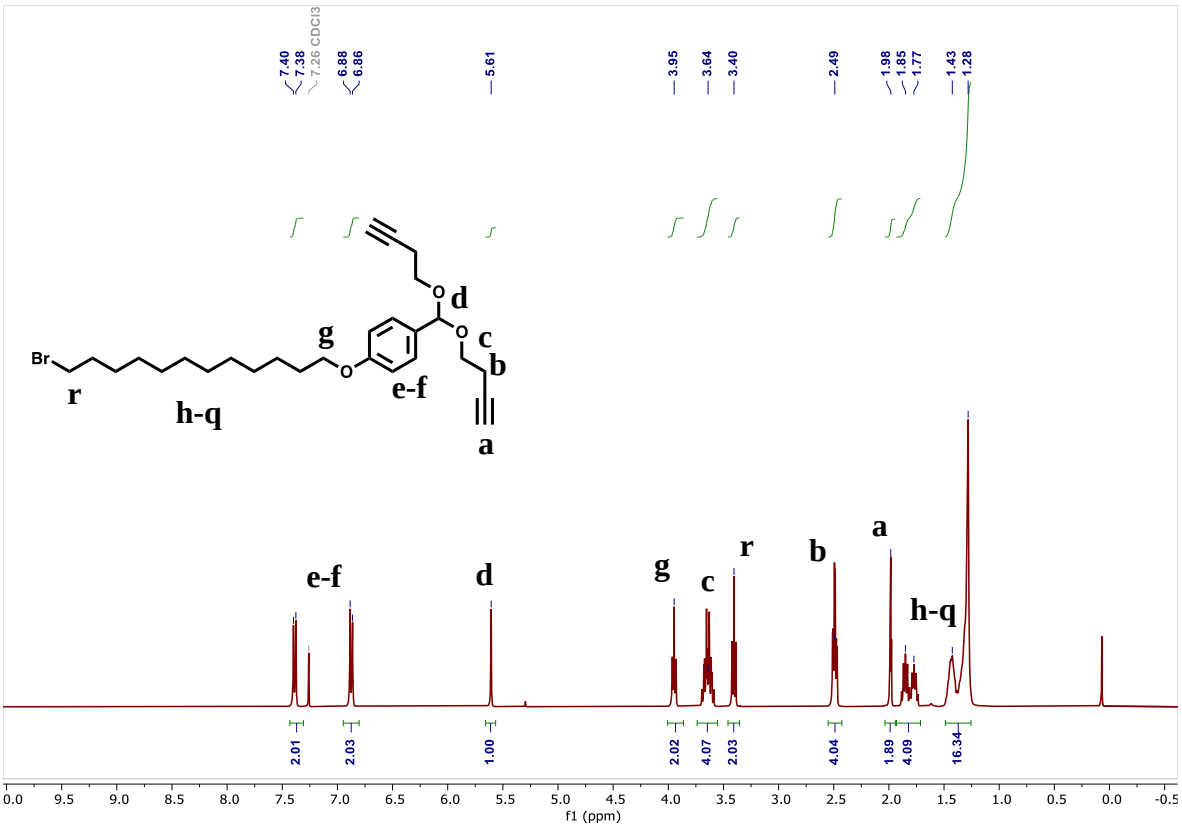

159.42
130.09
128.05
114.25
101.35
81.53
77.40
69.44
68.05
63.30
34.13
32.91
29.34
29.32
28.82
28.23
26.09
20.07
Br
O
O
O
f1 (ppm)

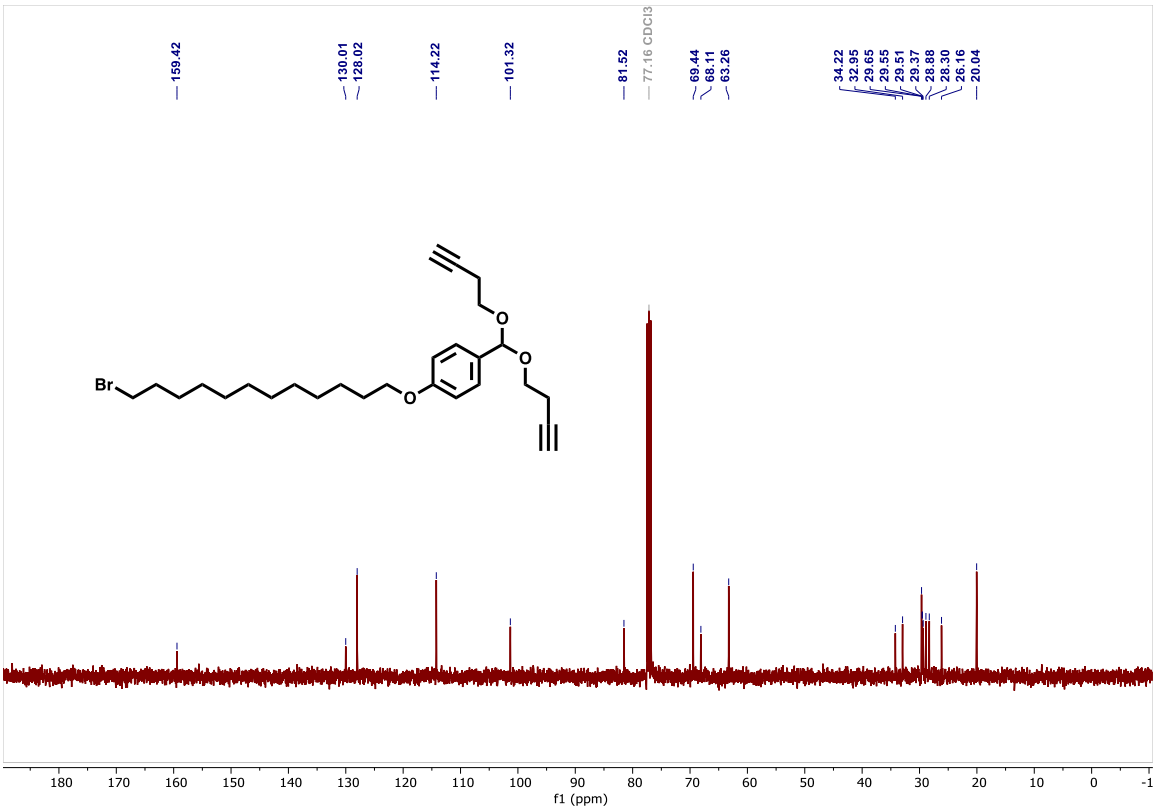

HSQC
Br
O
O
O
f2 (ppm)
f1 (ppm)

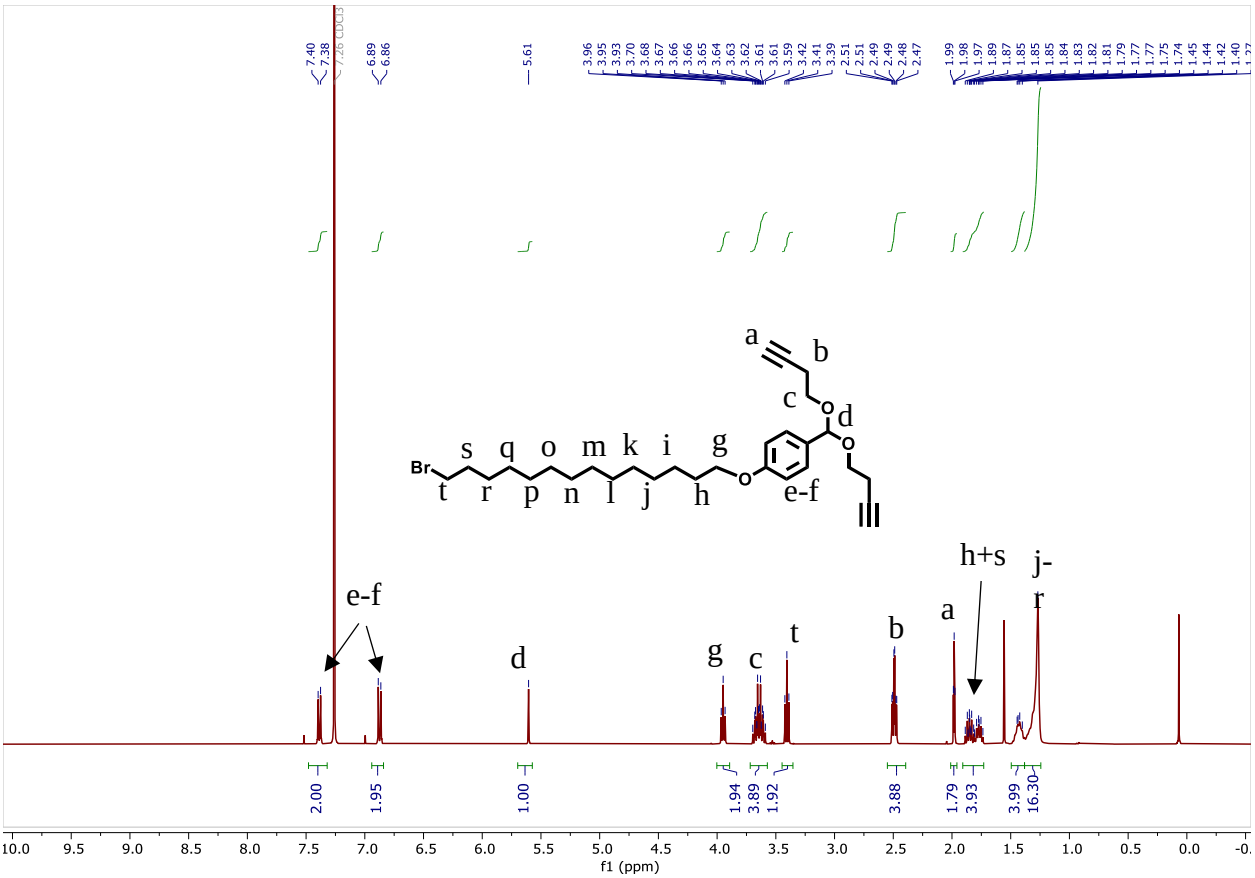

COSY
Br
O
O
O
f2 (ppm)
f1 (ppm)

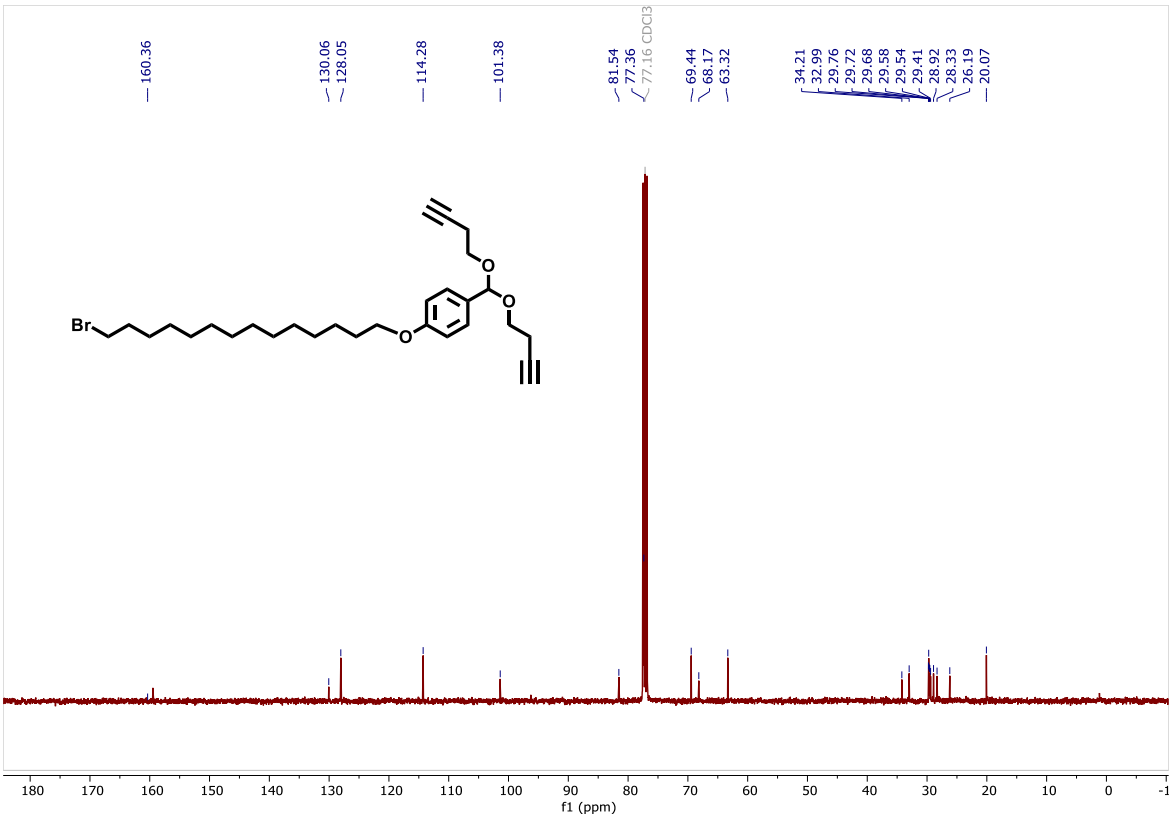

7.40
7.38
7.26 CDCl3
6.89
6.86
5.61
3.95
3.64
3.41
2.49
1.98
1.85
1.77
1.43
1.31
Br
O
O
O
g
d
c
b
a
e-f
p
h-o
e-f
d
g
c
p
b
a
h-o
2.01
2.02
1.00
2.01
4.06
2.09
4.03
1.87
4.15
12.41
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0
f1 (ppm)

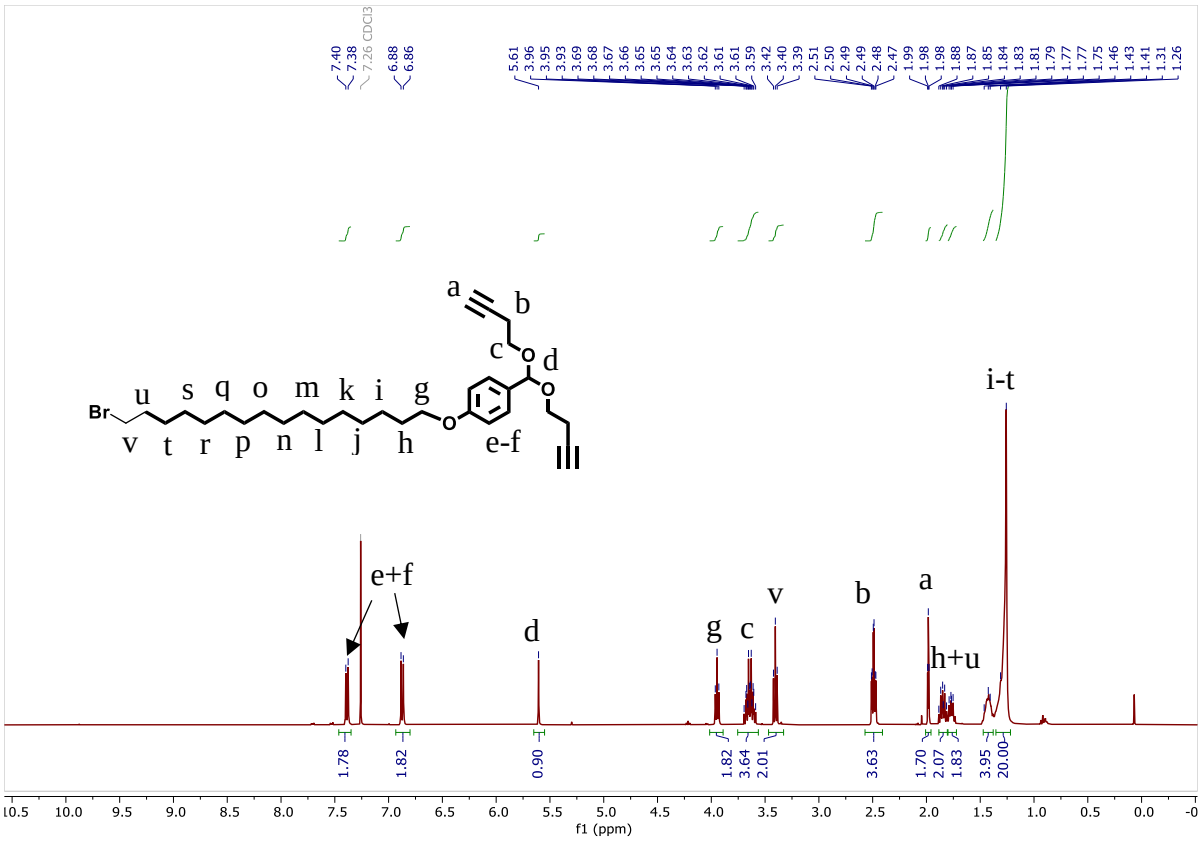

159.42
130.04
128.03
114.24
101.33
81.53
77.16 CDCl3
69.44
68.09
63.28
34.19
32.94
29.56
29.48
29.37
28.87
28.29
26.15
20.05
Br
O
O
O
f1 (ppm)

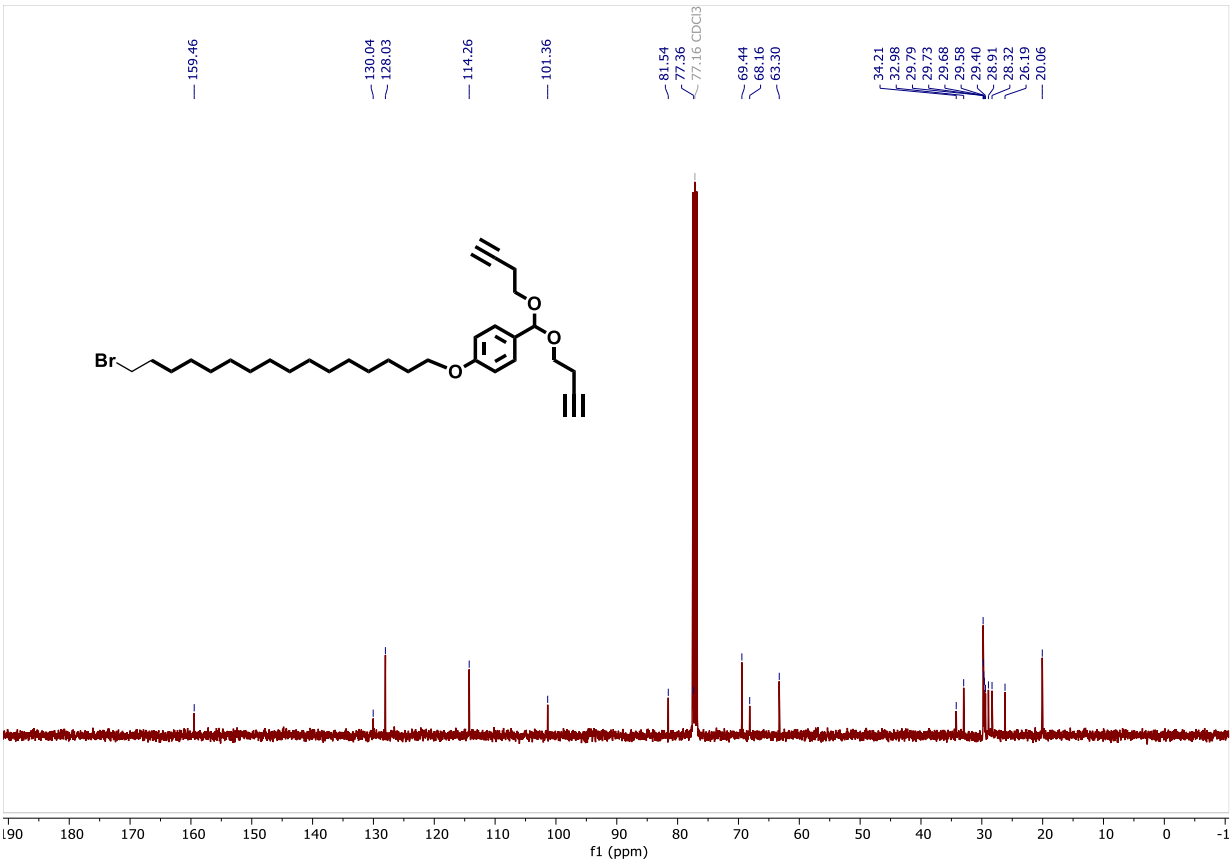

7.40
7.38
7.26 CDCl3
6.88
6.86
5.61
3.95
3.64
3.40
2.49
1.98
1.85
1.77
1.43
1.28
Br
r
h-q
g
e-f
d
c
b
a
e-f
d
g
c
r
b
a
h-q
2.01
2.03
1.00
2.02
4.07
2.03
4.04
1.89
4.09
16.34
f1 (ppm)

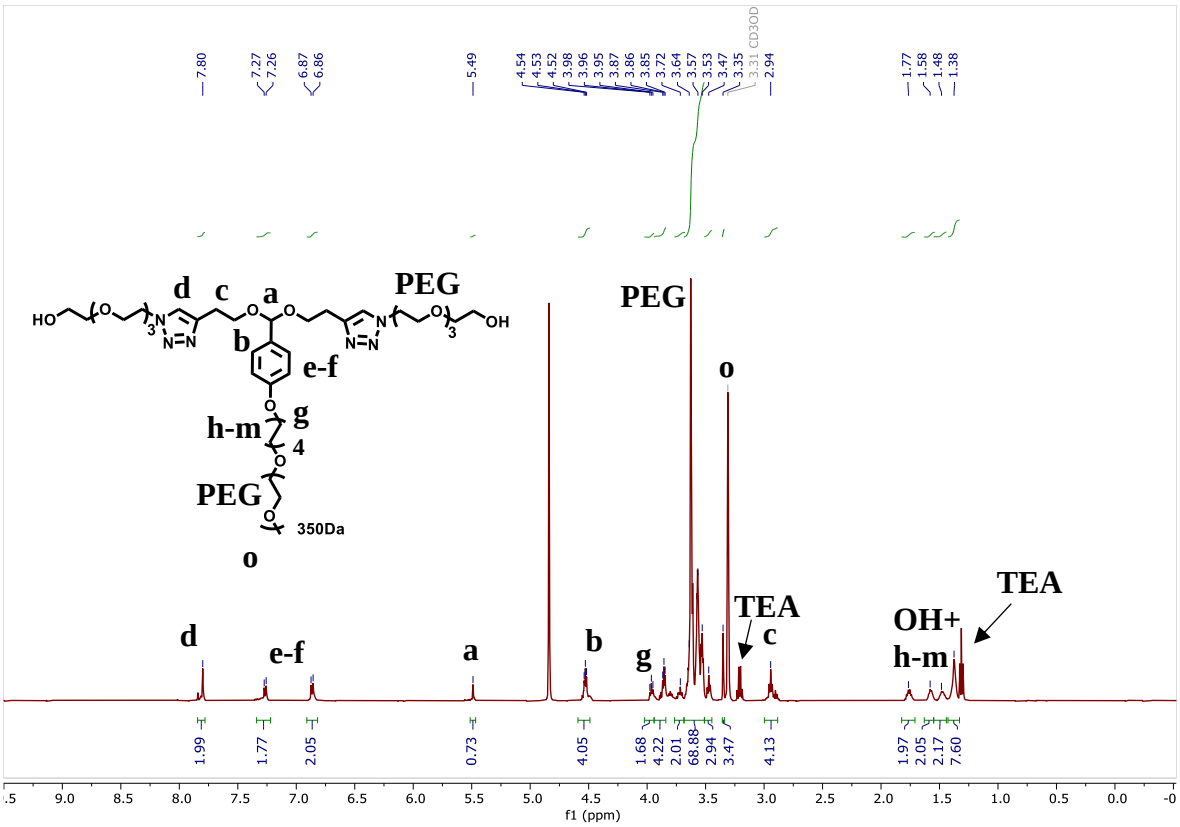

159.42
130.01
128.02
114.22
101.32
81.52
77.16 CDCl3
69.44
68.11
63.26
34.22
32.95
29.65
29.55
29.51
29.37
28.88
28.30
26.16
20.04
Br
O
O
O
O
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-1
f1 (ppm)

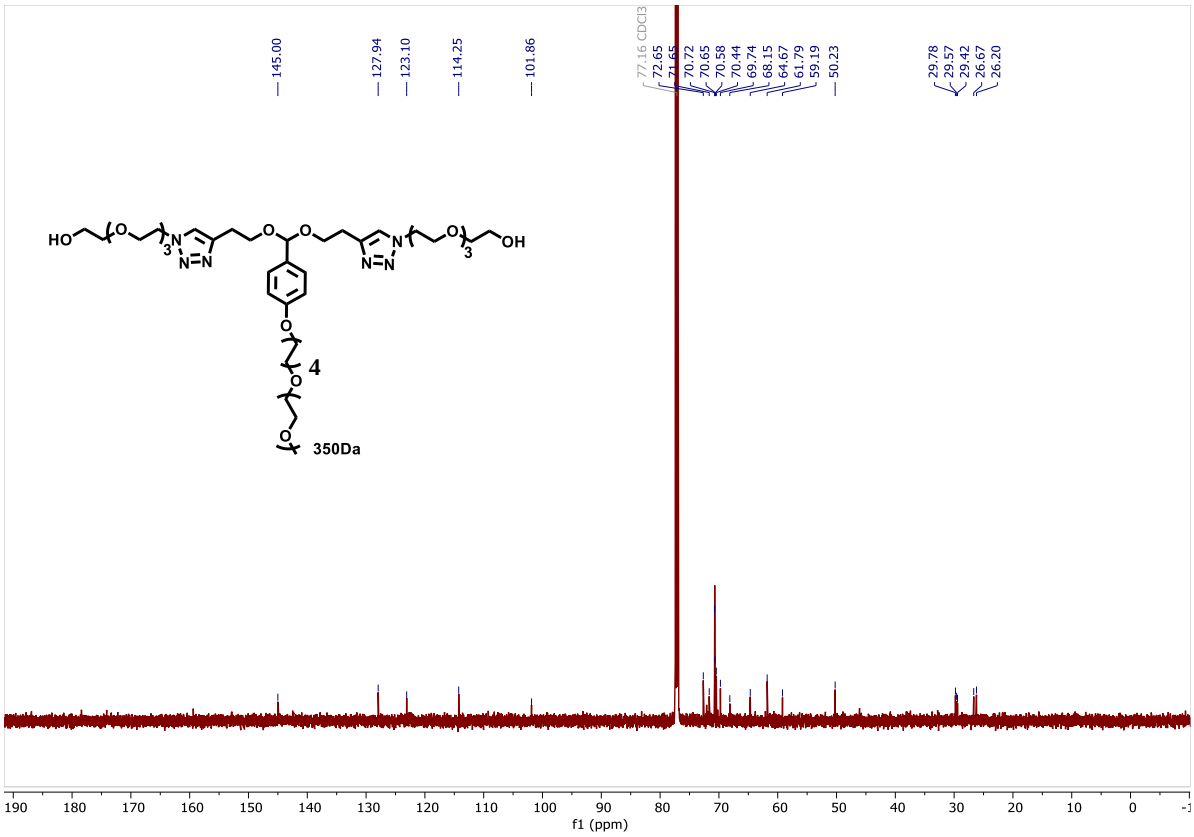

a
b
c
d
s q o m k i g
Br
t r p n l j h
e-f
e-f
d
g
c
t
b
a
h+s
j-r
2.00
1.95
1.00
1.94
3.89
1.92
3.88
1.79
3.93
3.99
16.30
f1 (ppm)

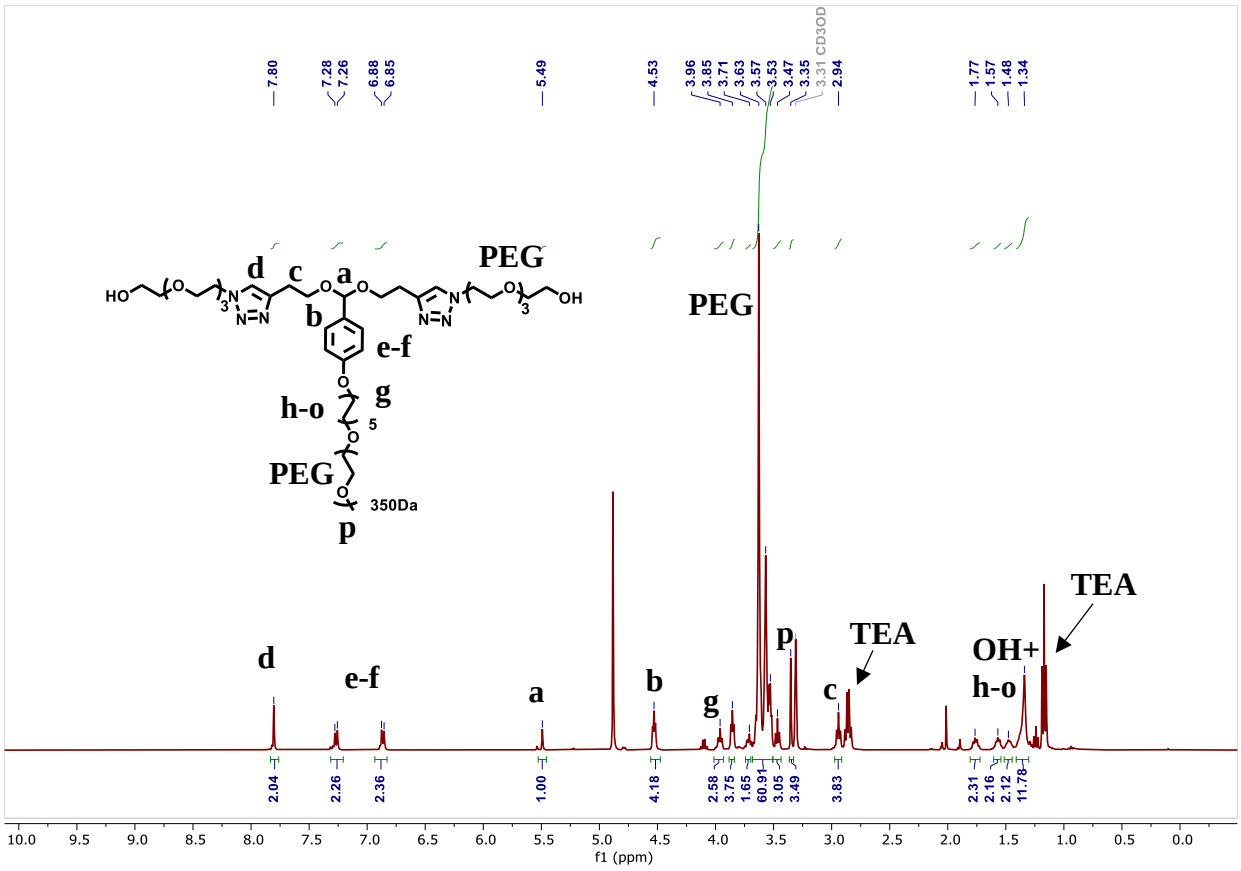

160.36
130.06
128.05
114.28
101.38
81.54
77.36
77.16 CDCl3
69.44
68.17
63.32
34.21
32.99
29.76
29.72
29.68
29.58
29.54
29.41
28.92
28.33
26.19
20.07
Br
O
O
O
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-1
f1 (ppm)

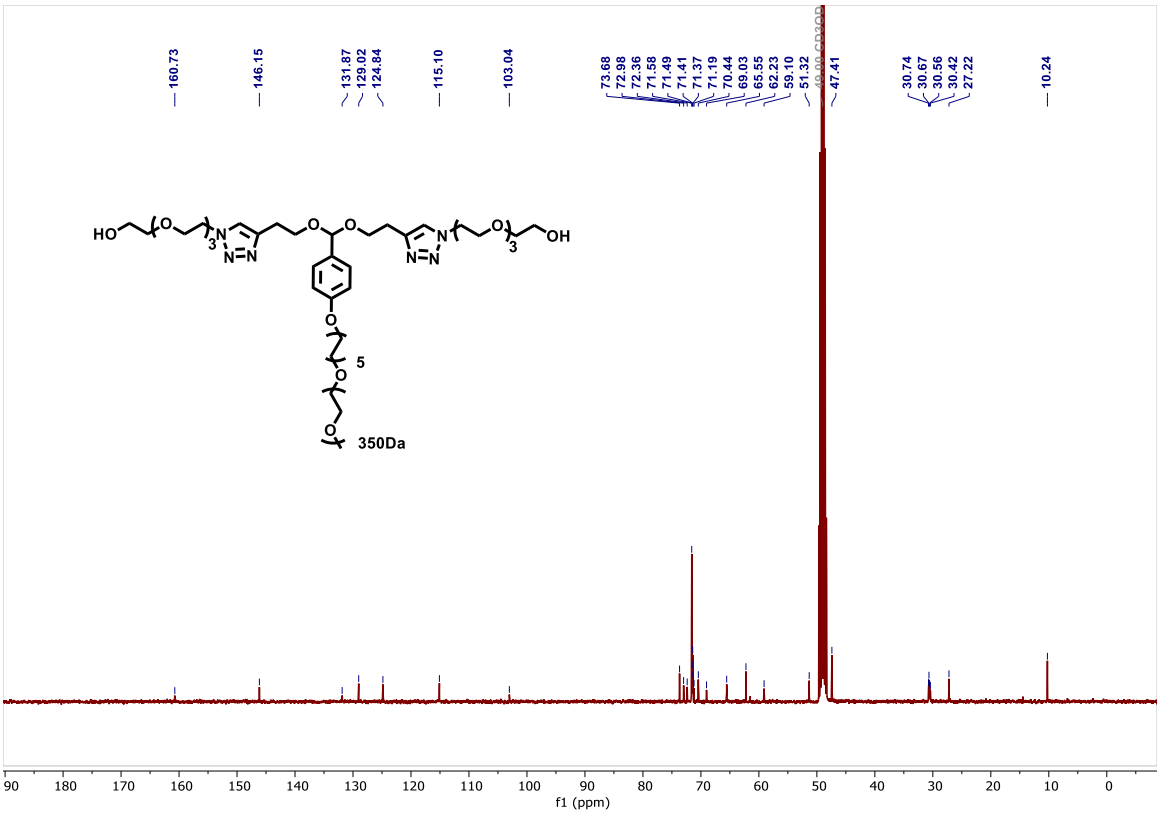

7.40
7.38
7.26 CDCl3
6.88
6.86
5.61
3.96
3.95
3.93
3.69
3.68
3.67
3.66
3.65
3.65
3.64
3.63
3.62
3.61
3.61
3.59
3.42
3.40
3.39
2.51
2.50
2.49
2.49
2.48
2.47
1.99
1.98
1.98
1.88
1.87
1.85
1.84
1.83
1.81
1.79
1.77
1.77
1.75
1.46
1.43
1.41
1.31
1.26
a
b
c
d
Br
u s q o m k i g
v t r p n l j h
e-f
i-t
e+f
d
g
c
v
b
a
h+u
1.78
1.82
0.90
1.82
3.64
2.01
3.63
1.70
2.07
1.83
3.95
20.00
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0
f1 (ppm)

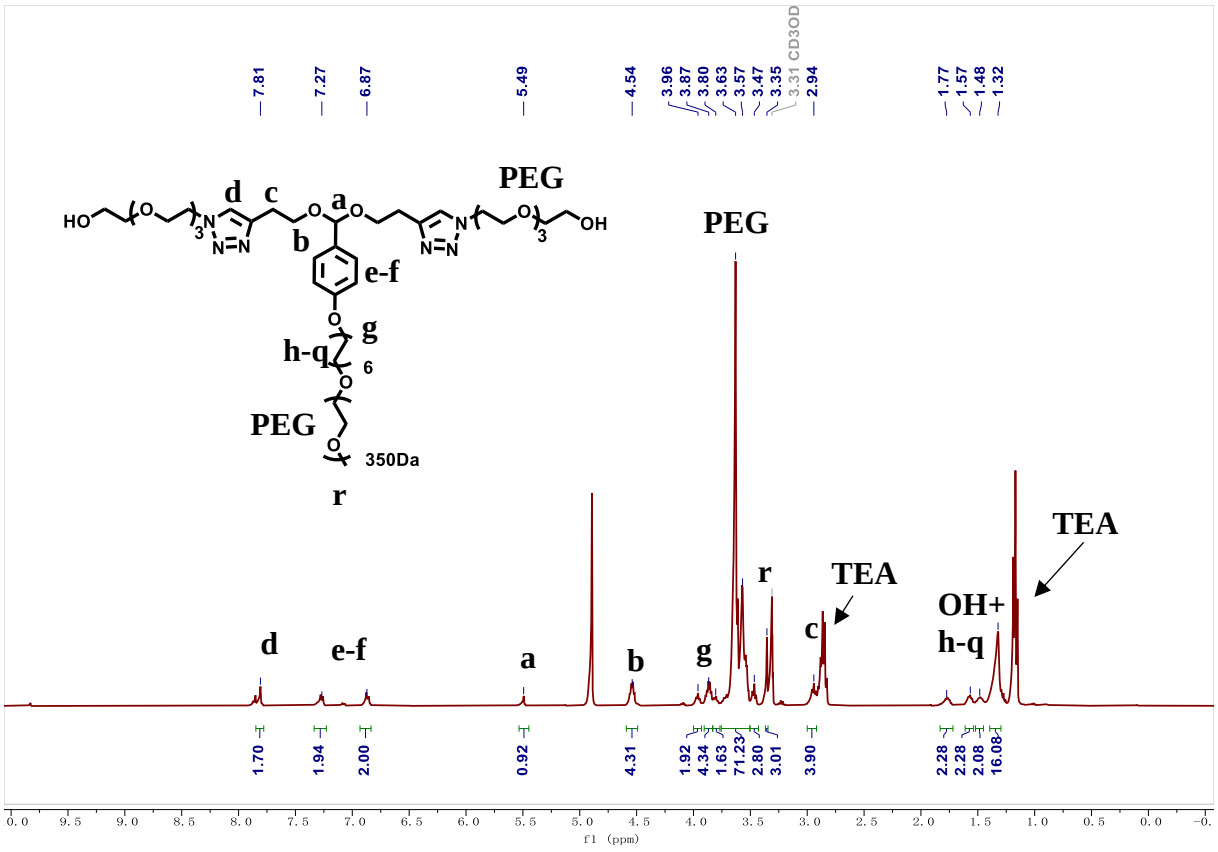

159.46
130.04
128.03
114.26
101.36
81.54
77.36
77.16 CDCl3
69.44
68.16
63.30
34.21
32.98
29.79
29.73
29.68
29.58
29.40
28.91
28.32
26.19
20.06
Br
O
O
O
f1 (ppm)

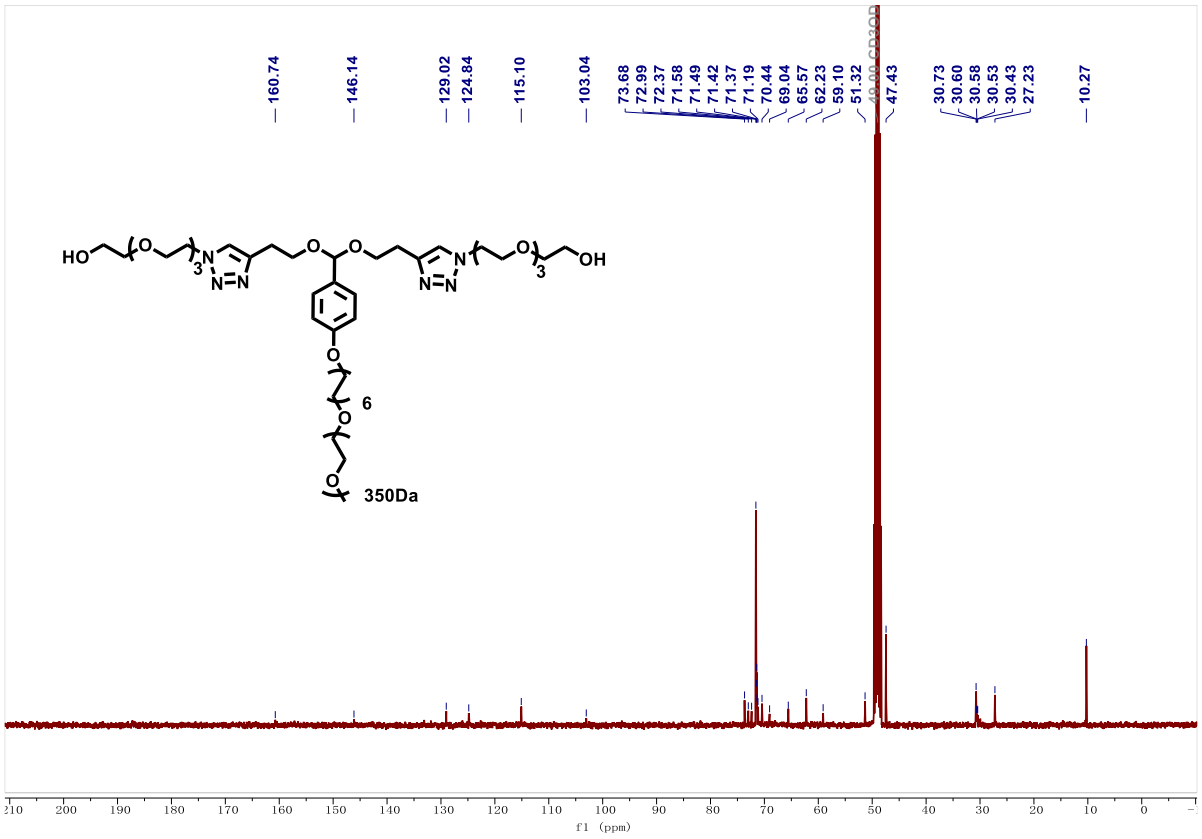

7.80
7.27
7.26
6.87
6.86
5.49
4.54
4.53
4.52
3.98
3.96
3.95
3.87
3.86
3.85
3.72
3.64
3.57
3.53
3.47
3.35
3.31
CD3OD
2.94
1.77
1.58
1.48
1.38
HO
OH
PEG
d
c
a
b
e-f
g
h-m
4
PEG
350Da
o
TEA
OH+
h-m
1.99
1.77
2.05
0.73
4.05
1.68
4.22
2.01
68.88
2.94
3.47
4.13
1.97
2.05
2.17
7.60
f1 (ppm)

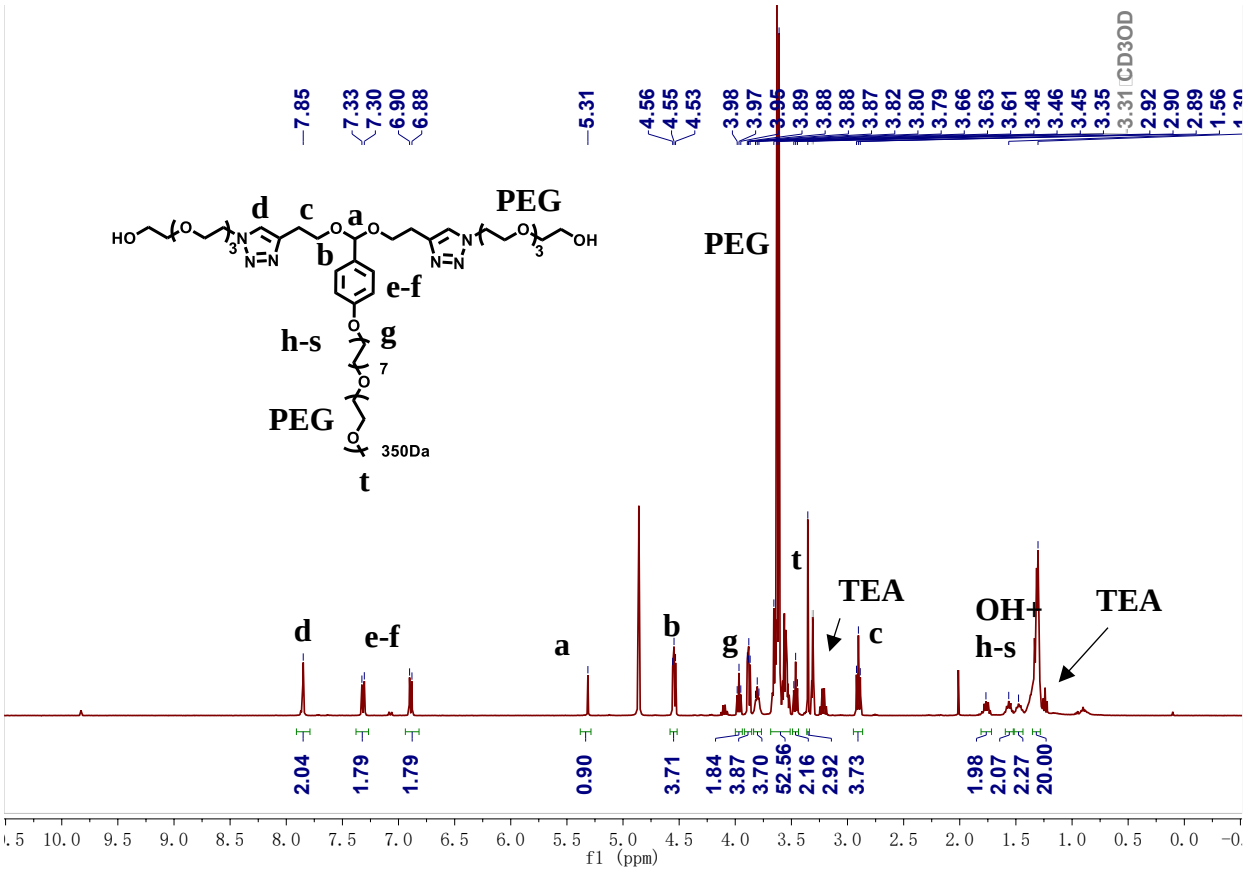

145.00
127.94
123.10
114.25
101.86
77.16 CDCl3
72.65
71.65
70.72
70.65
70.58
70.44
69.74
68.15
64.67
61.79
59.19
50.23
29.78
29.57
29.42
26.67
26.20
HO
OH
4
350Da
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

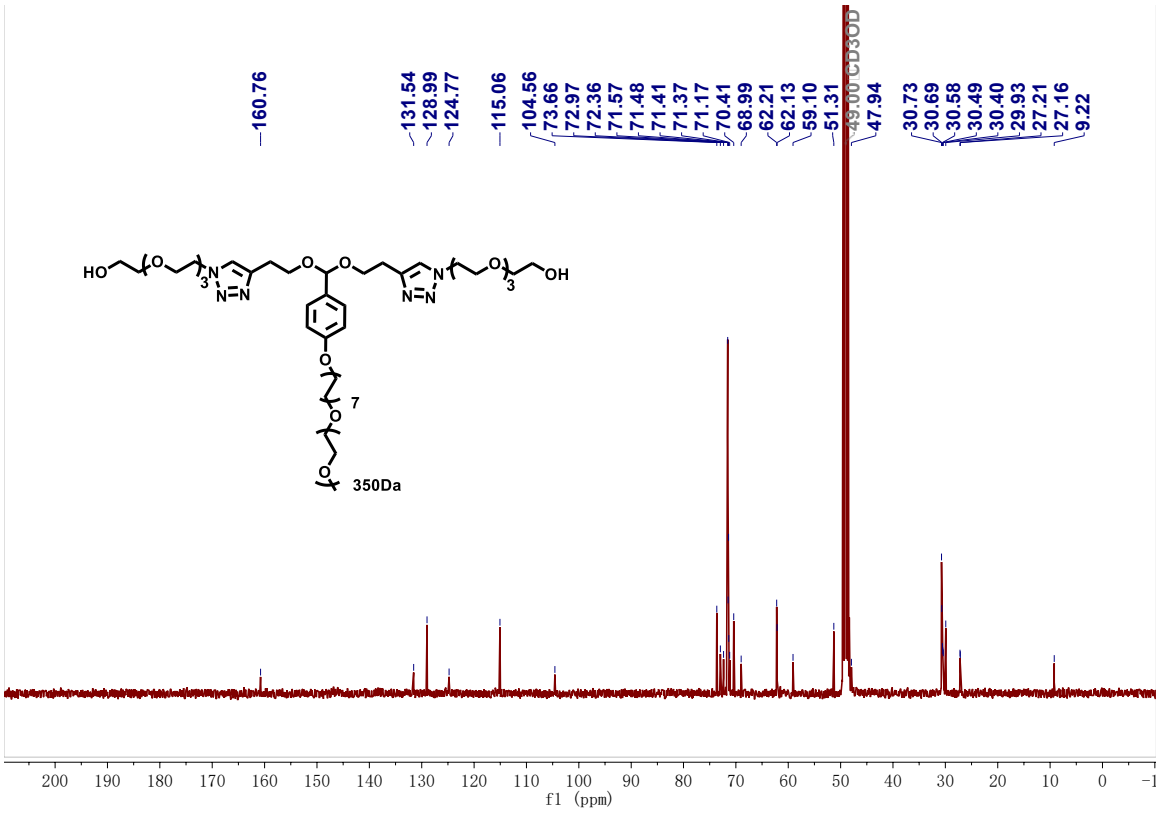

7.80
7.28
7.26
6.88
6.85
5.49
4.53
3.96
3.85
3.71
3.63
3.57
3.53
3.47
3.35
3.31 CD3OD
2.94
1.77
1.57
1.48
1.34
HO
OH
PEG
d
c
a
b
e-f
h-o
g
5
PEG
350Da
p
PEG
d
e-f
a
b
g
p
c
TEA
OH+
h-o
TEA
2.04
2.26
2.36
1.00
4.18
2.58
3.75
1.65
60.91
3.05
3.49
3.83
2.31
2.16
2.12
11.78
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

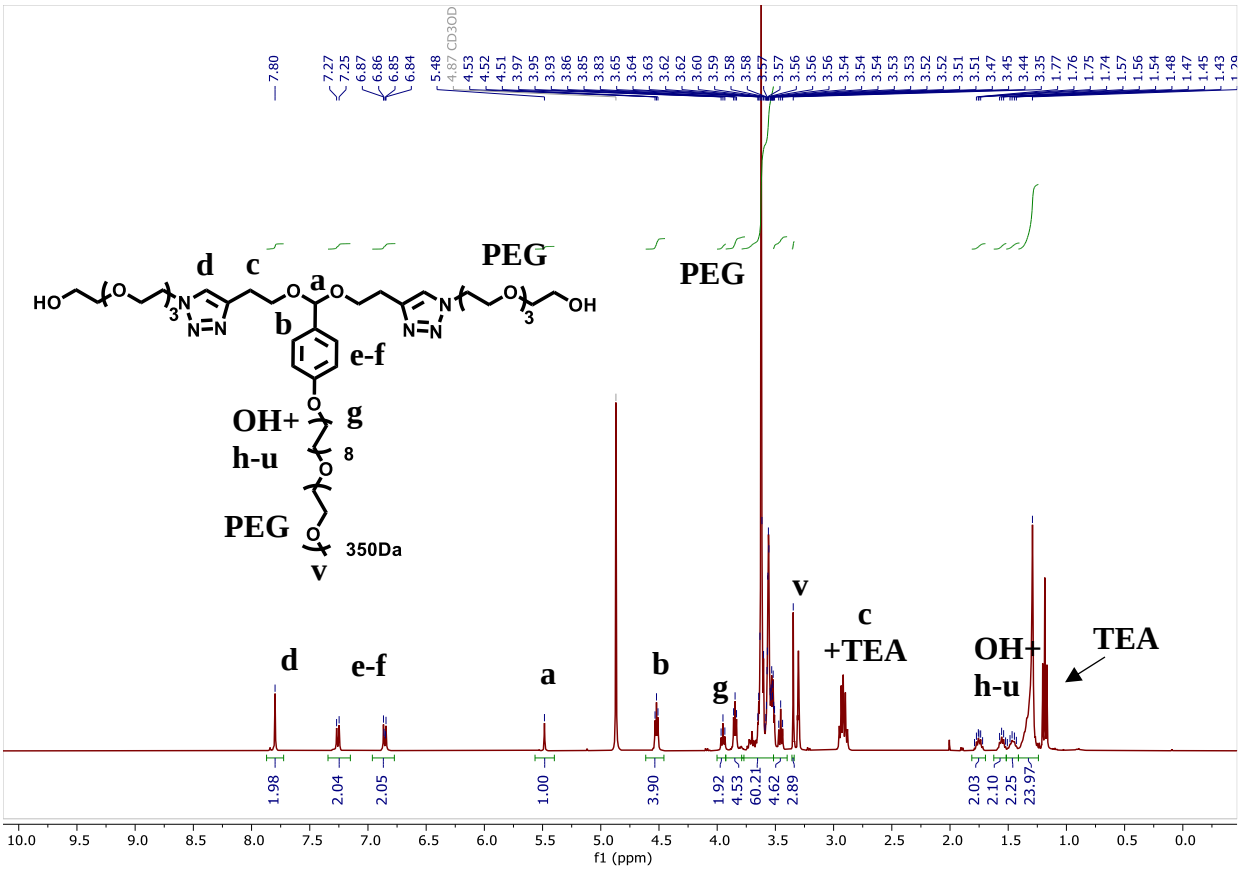

160.73
146.15
131.87
129.02
124.84
115.10
103.04
73.68
72.98
72.36
71.58
71.49
71.41
71.37
71.19
70.44
69.03
65.55
62.23
59.10
51.32
47.41
30.74
30.67
30.56
30.42
27.22
10.24
HO
OH
5
350Da
f1 (ppm)

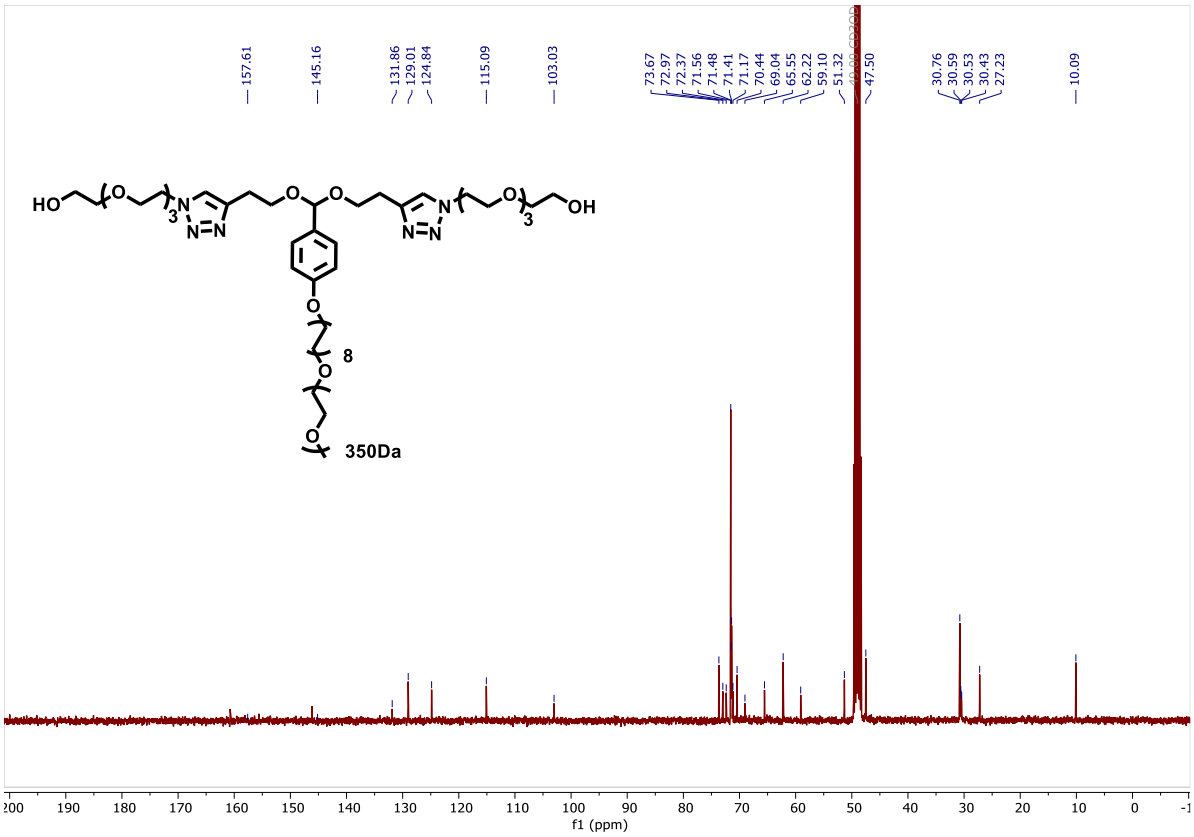

7.81
7.27
6.87
5.49
4.54
3.96
3.87
3.80
3.63
3.57
3.47
3.35
3.31 CD3OD
2.94
1.77
1.57
1.48
1.32
HO
OH
PEG
PEG
d
c
a
b
e-f
g
h-q
6
PEG
350Da
r
TEA
TEA
OH+
h-q
d
e-f
a
b
g
r
c
1.70
1.94
2.00
0.92
4.31
1.92
4.34
1.63
71.23
2.80
3.01
3.90
2.28
2.28
2.08
16.08
f1 (ppm)

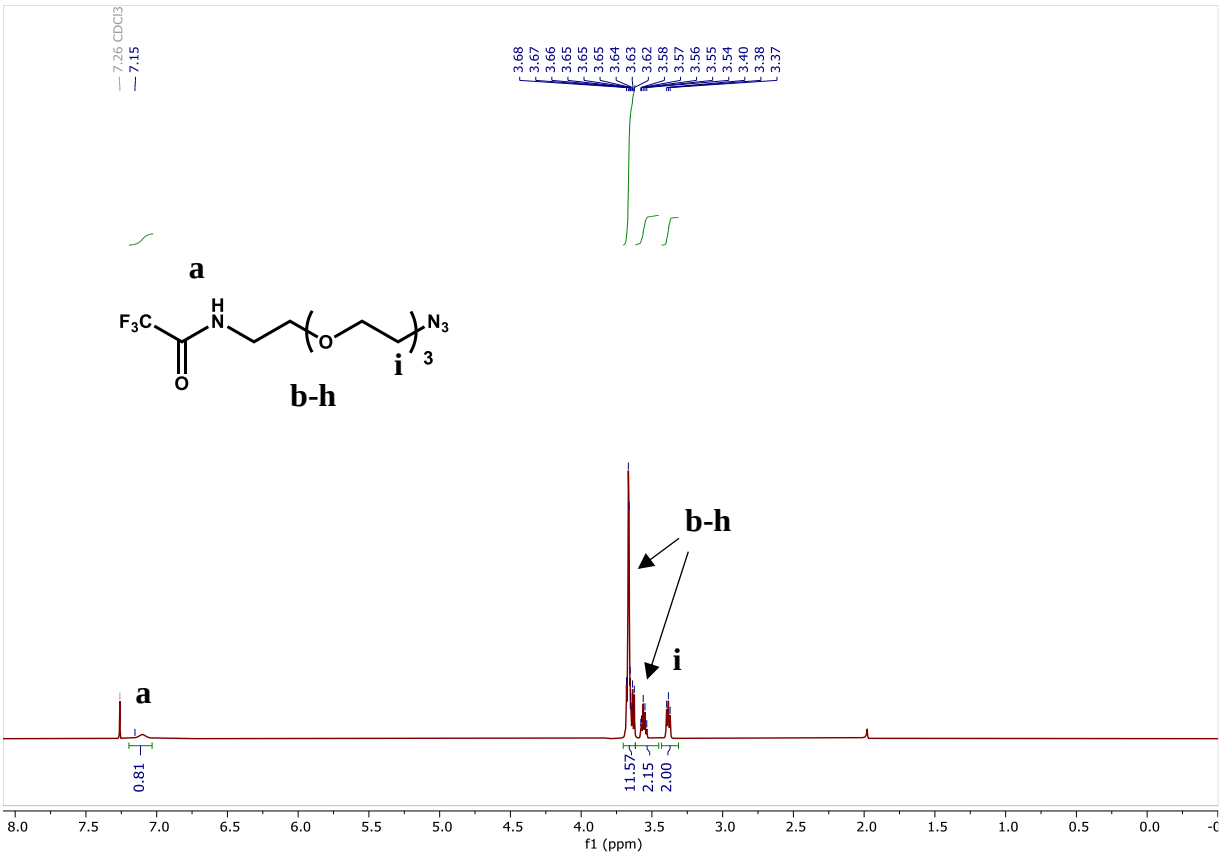

160.74
146.14
129.02
124.84
115.10
103.04
73.68
72.99
72.37
71.58
71.49
71.42
71.37
71.19
70.44
69.04
65.57
62.23
59.10
51.32
47.43
30.73
30.60
30.58
30.53
30.43
27.23
10.27
HO
OH
6
350Da
f1 (ppm)

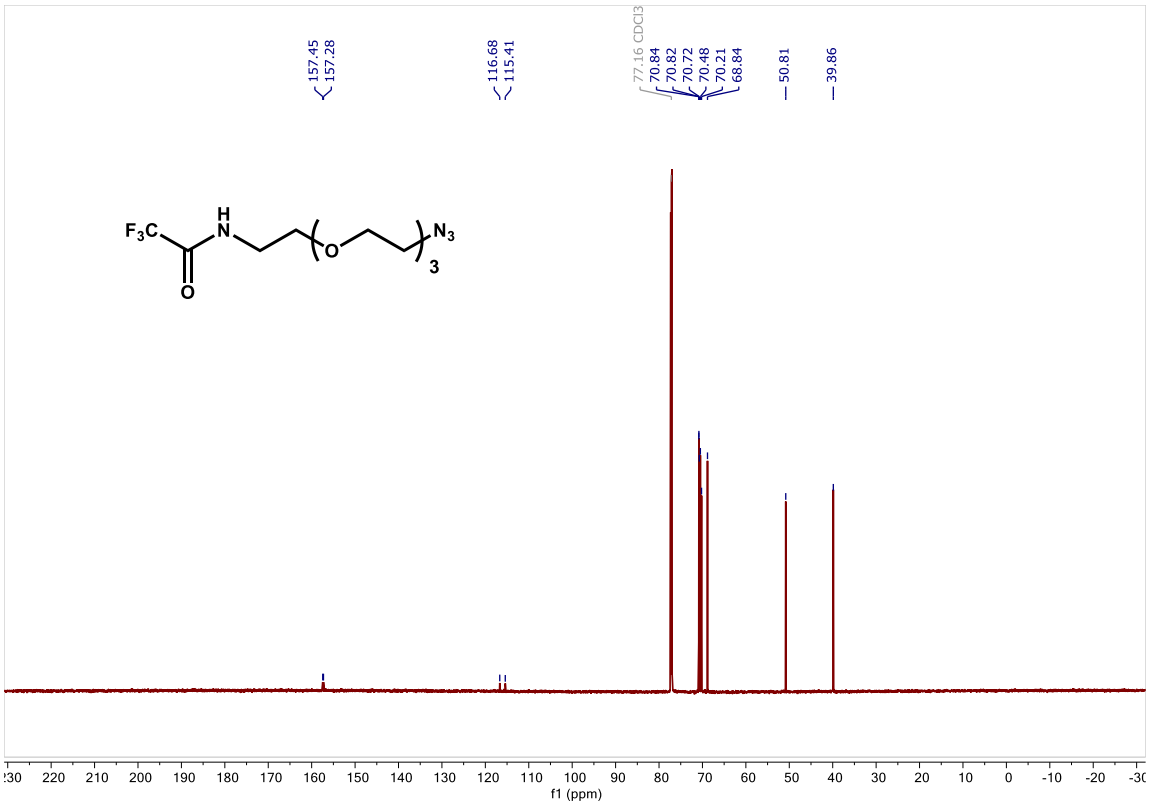

7.85
7.33
7.30
6.90
6.88
5.31
4.56
4.55
4.53
3.98
3.97
3.95
3.89
3.88
3.88
3.87
3.82
3.80
3.79
3.66
3.63
3.61
3.48
3.46
3.45
3.35
3.31
CD3OD
2.92
2.90
2.89
1.56
HO
OH
PEG
PEG
d
c
a
b
e-f
h-s
g
7
PEG
350Da
t
d
e-f
a
b
g
t
TEA
c
OH+
h-s
TEA
2.04
1.79
1.79
0.90
3.71
1.84
3.87
3.70
52.56
2.16
2.92
3.73
1.98
2.07
2.27
20.00
f1 (ppm)

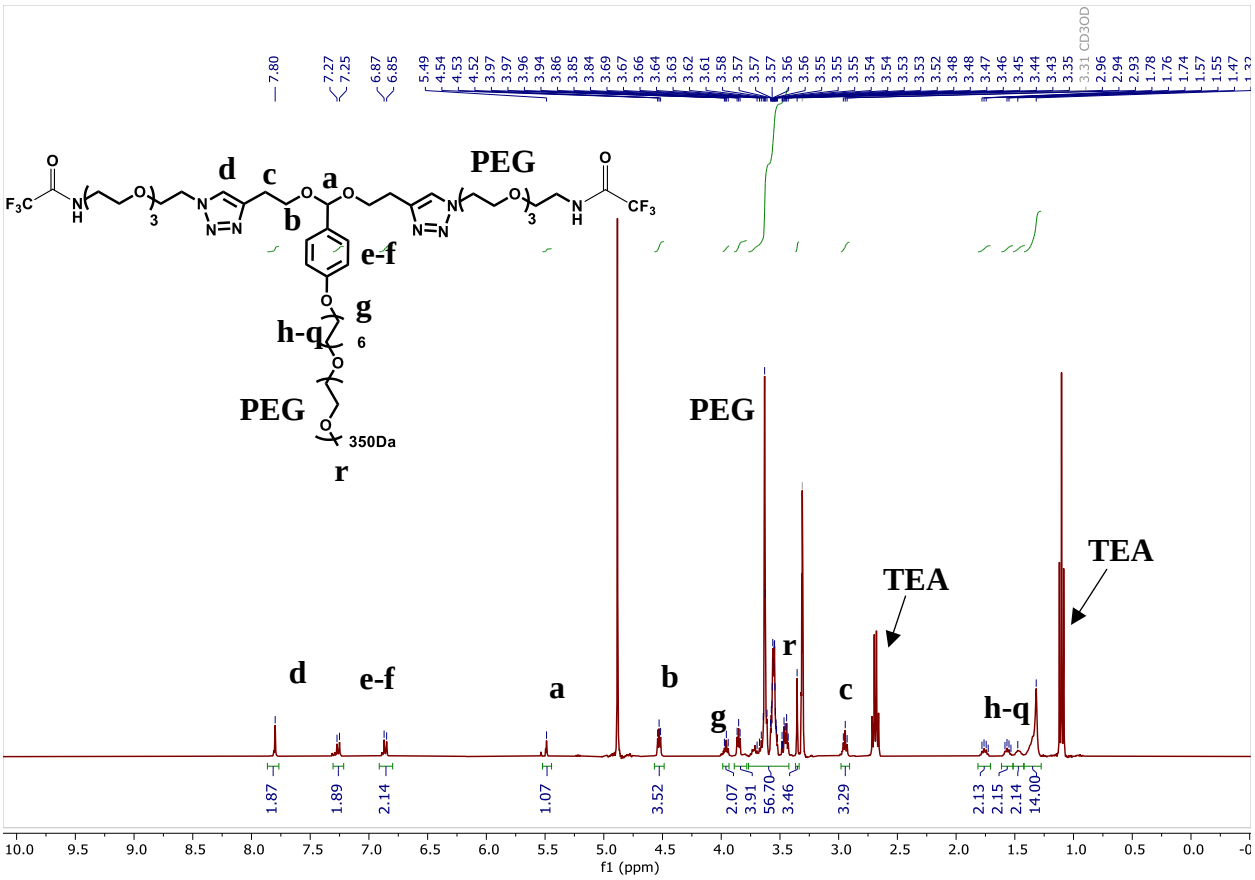

160.76
131.54
128.99
124.77
115.06
104.56
73.66
72.97
72.36
71.57
71.48
71.41
71.37
71.17
70.41
68.99
62.21
62.13
59.10
51.31
49.00 CD3OD
47.94
30.73
30.69
30.58
30.49
30.40
29.93
27.21
27.16
9.22
HO
OH
7
350Da
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1
f1 (ppm)

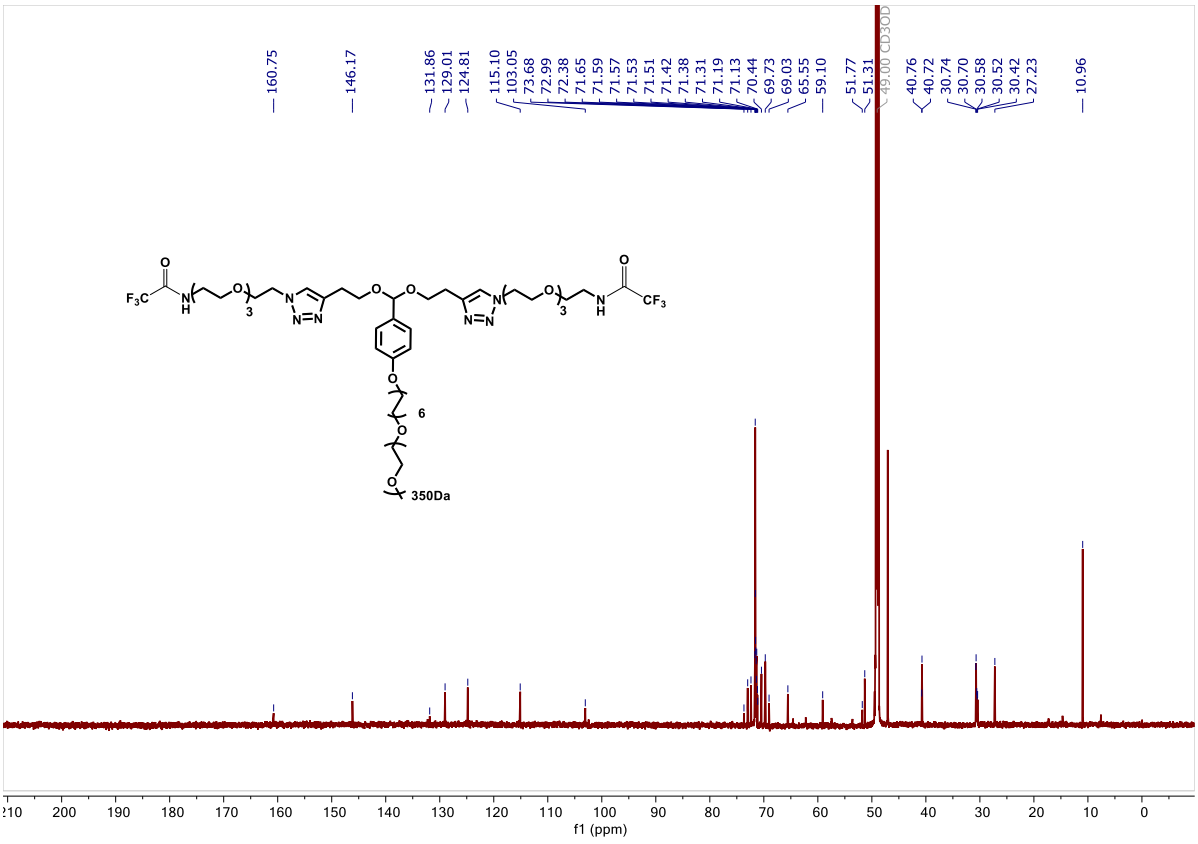

d
c
a
PEG
PEG
HO
OH
b
e-f
OH+
h-u
g
8
PEG
350Da
v
d
e-f
a
b
g
v
c
+TEA
OH+
h-u
TEA
1.98
2.04
2.05
1.00
3.90
1.92
4.53
60.21
4.62
2.89
2.03
2.10
2.25
23.97
f1 (ppm)

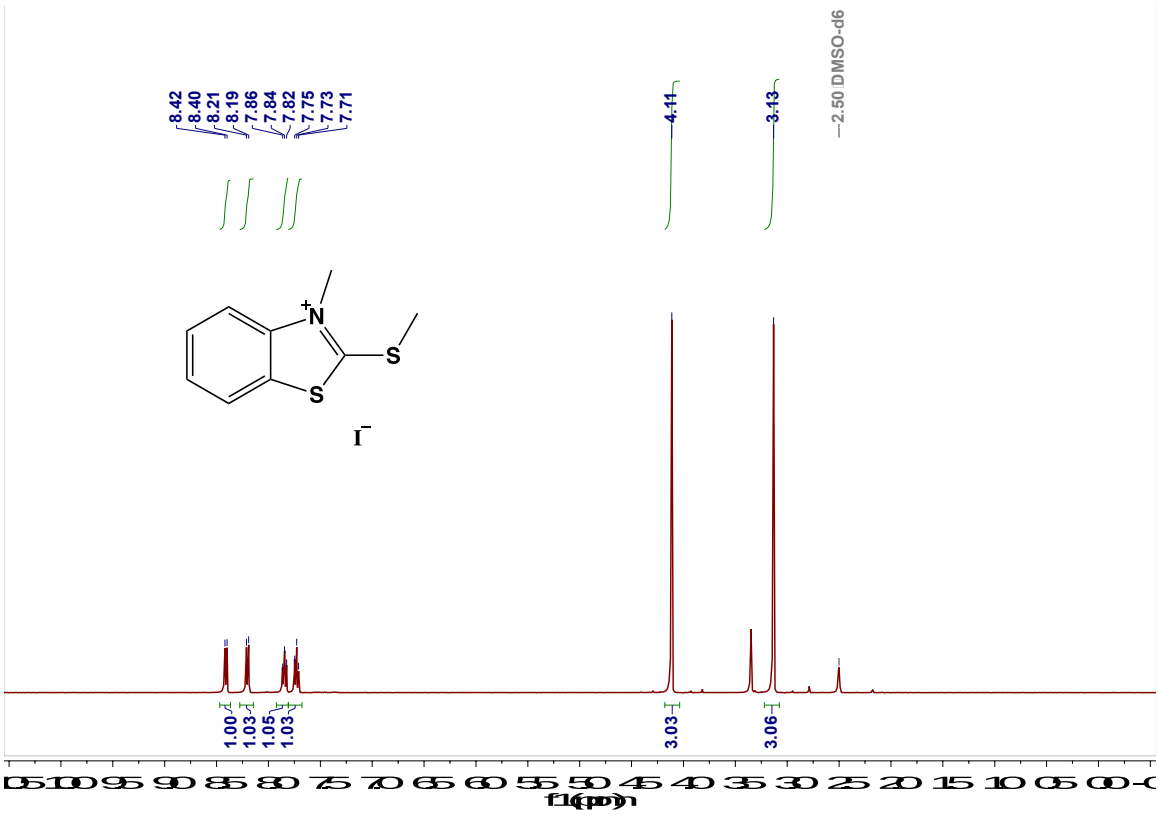

157.61
145.16
131.86
129.01
124.84
115.09
103.03
73.67
72.97
72.37
71.56
71.48
71.41
71.17
70.44
69.04
65.55
62.22
59.10
51.32
47.50
30.76
30.59
30.53
30.43
27.23
10.09
HO
OH
8
350Da
f1 (ppm)

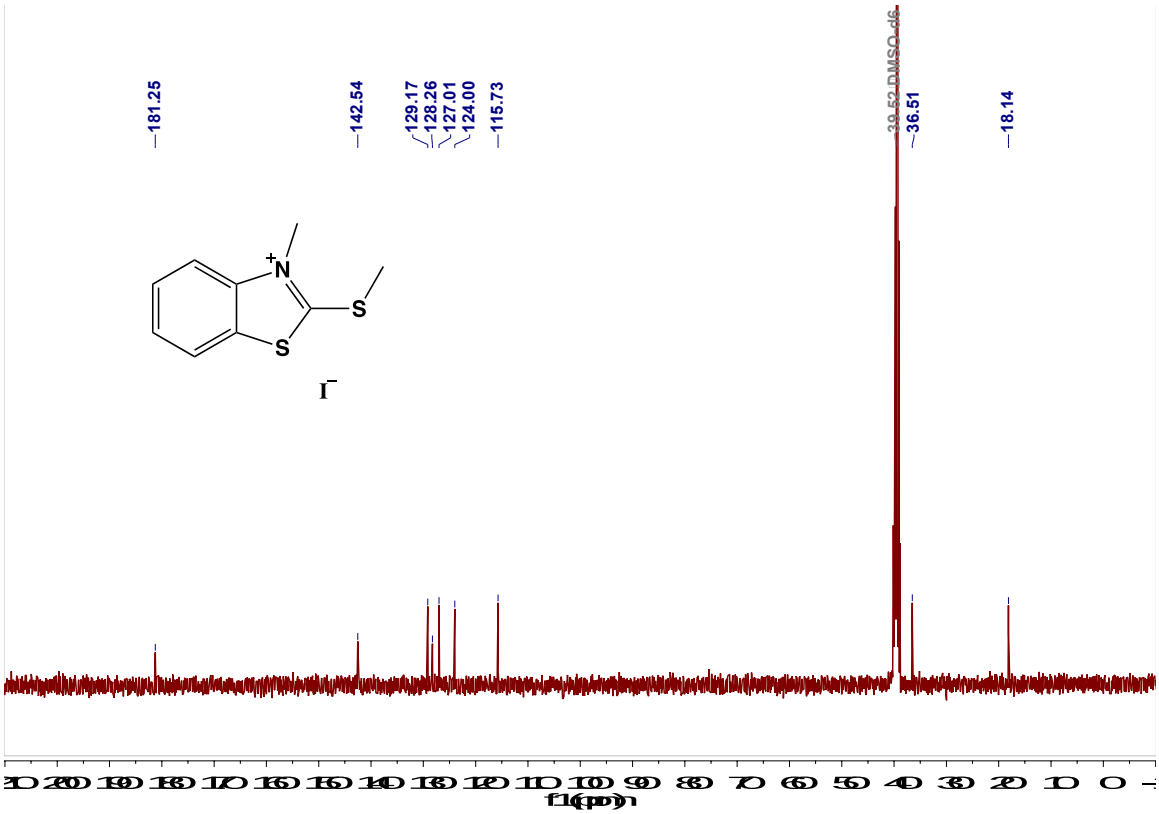

7.26 CDCl3
7.15
3.68
3.67
3.66
3.65
3.65
3.65
3.64
3.63
3.62
3.58
3.57
3.56
3.55
3.54
3.40
3.38
3.37
a
F3C
H
N
O
O
N3
3
i
b-h
b-h
i
a
0.81
11.57
2.15
2.00
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

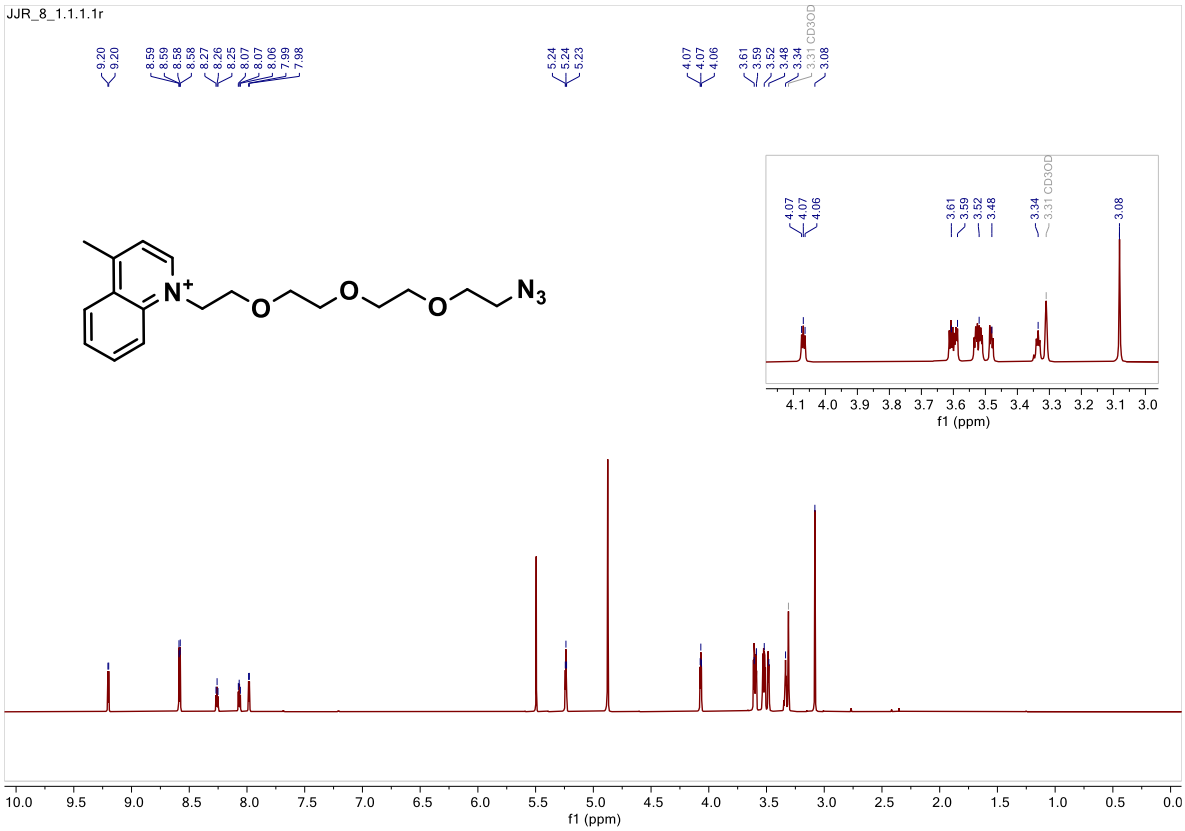

157.45
157.28
116.68
115.41
77.16 CDCl3
70.84
70.82
70.72
70.48
70.21
68.84
50.81
39.86
F3C
H
N
O
O
N3
3
f1 (ppm)

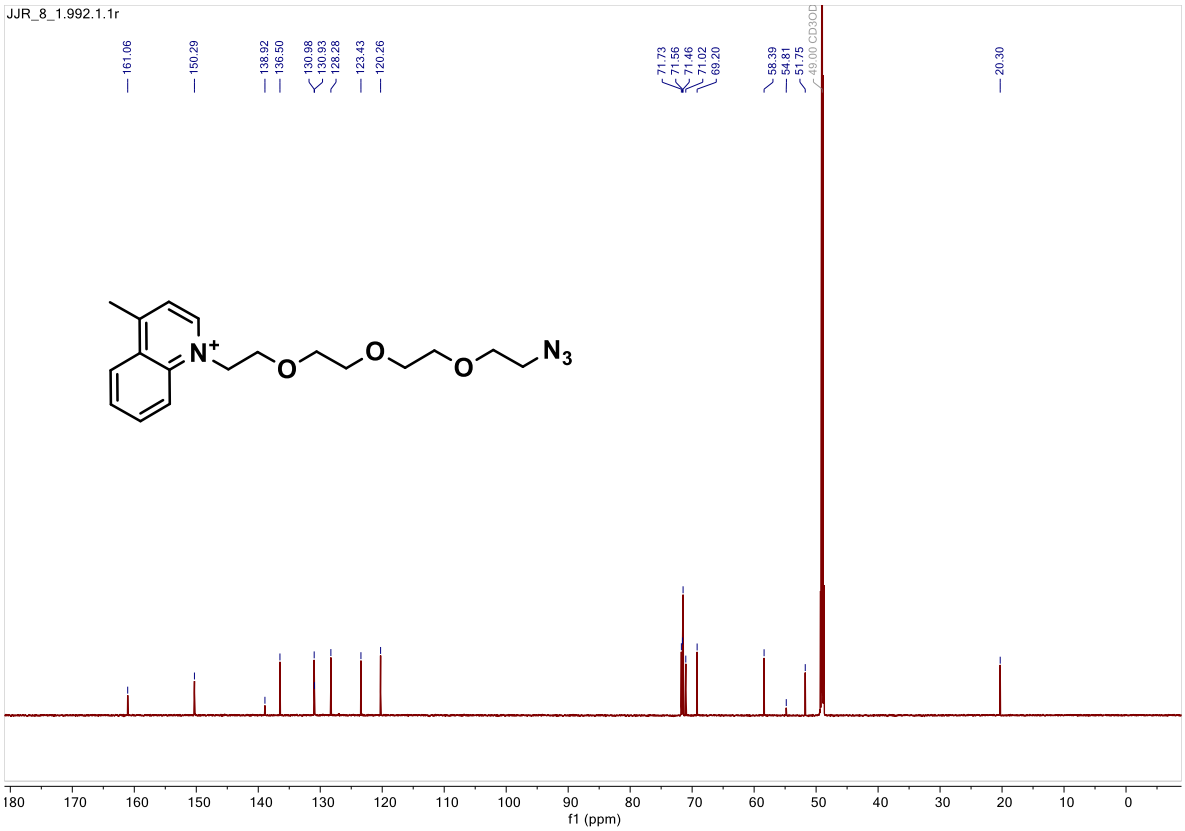

d
c
a
PEG
b
e-f
g
h-q
6
PEG
350Da
r
TEA
TEA
PEG
d
e-f
a
b
g
r
c
h-q
f1 (ppm)

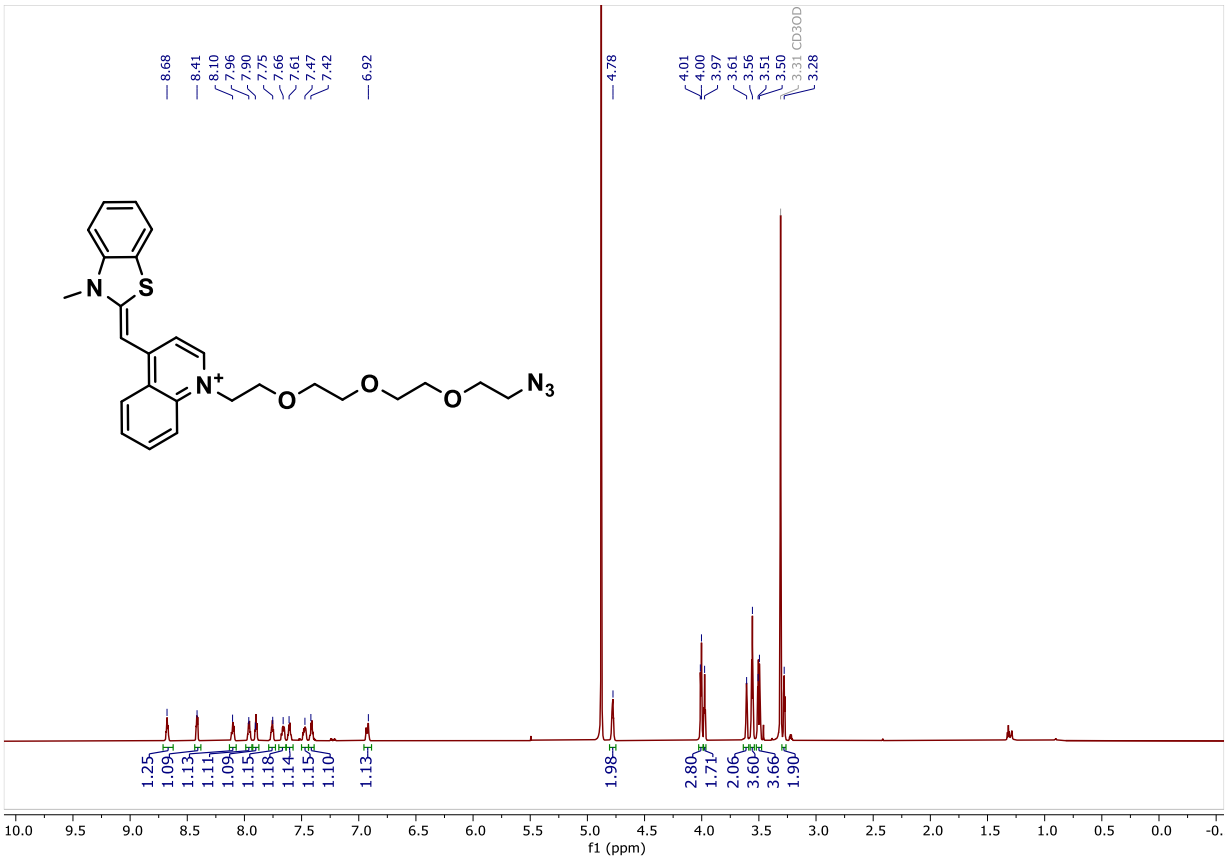

160.75
146.17
131.86
129.01
124.81
115.10
103.05
73.68
72.99
72.38
71.65
71.59
71.57
71.53
71.51
71.42
71.38
71.31
71.19
71.13
70.44
69.73
69.03
65.55
59.10
51.77
51.31
49.00 CD3OD
40.76
40.72
30.74
30.70
30.58
30.52
30.42
27.23
10.96
6
350Da
f1 (ppm)

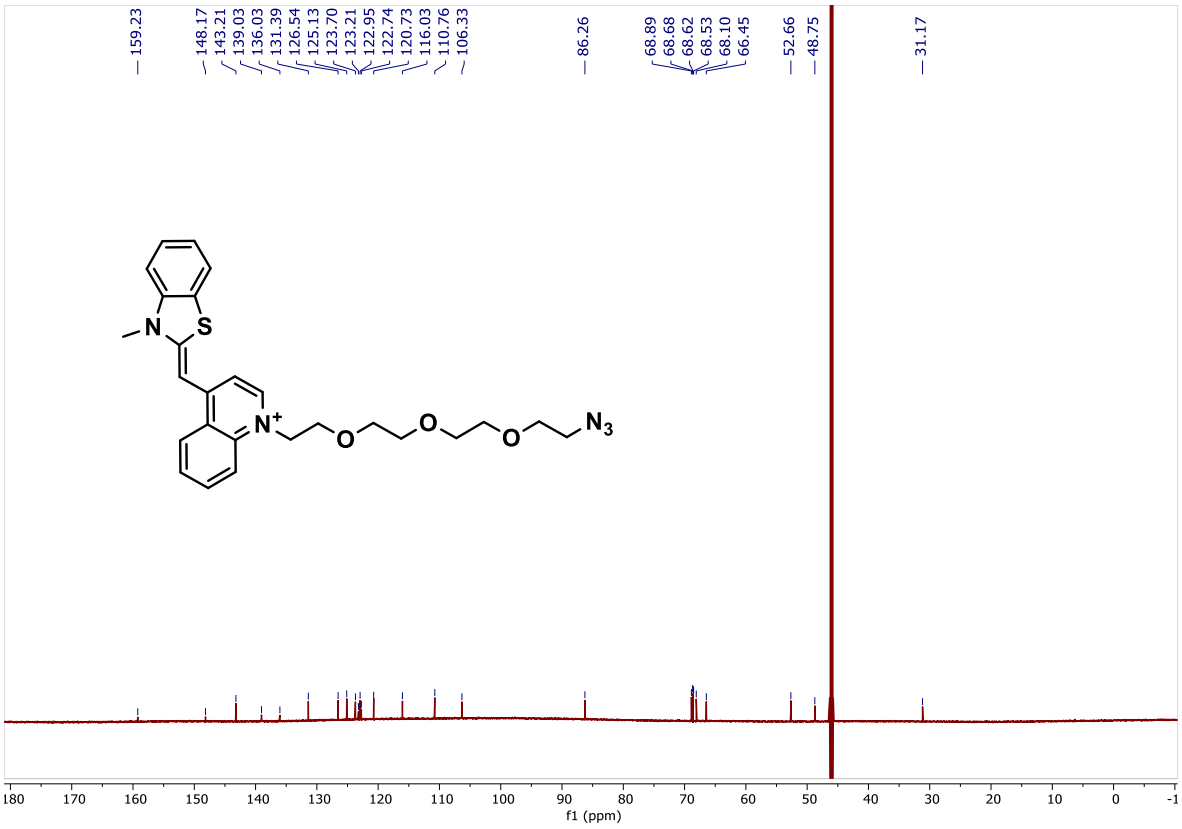

8.42
8.40
8.21
8.19
7.86
7.84
7.82
7.75
7.73
7.71
4.11
3.13
2.50 DMSO-d6
N
S
S
I⁻
1.00
1.03
1.05
1.03
3.03
3.06
f1 (ppm)

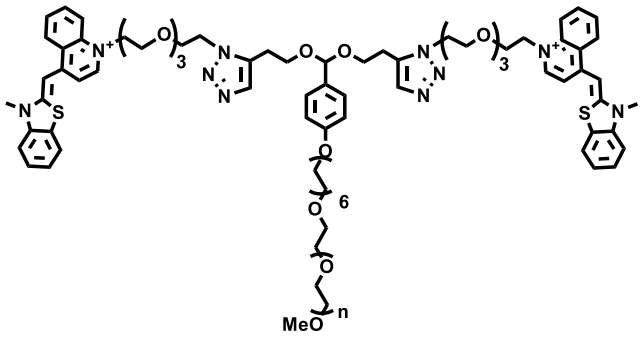

181.25
142.54
129.17
128.26
127.01
124.00
115.73
39.52 DMSO-d6
36.51
18.14
N
S
S
I⁻
f1 (ppm)


JJR_8_1.1.1.1r
9.20
9.20
8.59
8.59
8.58
8.58
8.27
8.26
8.25
8.07
8.07
8.06
7.99
7.98
5.24
5.24
5.23
4.07
4.07
4.06
3.61
3.59
3.52
3.48
3.34
3.31 CD3OD
3.08
N+
O
O
O
N3
4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1 3.0
f1 (ppm)
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

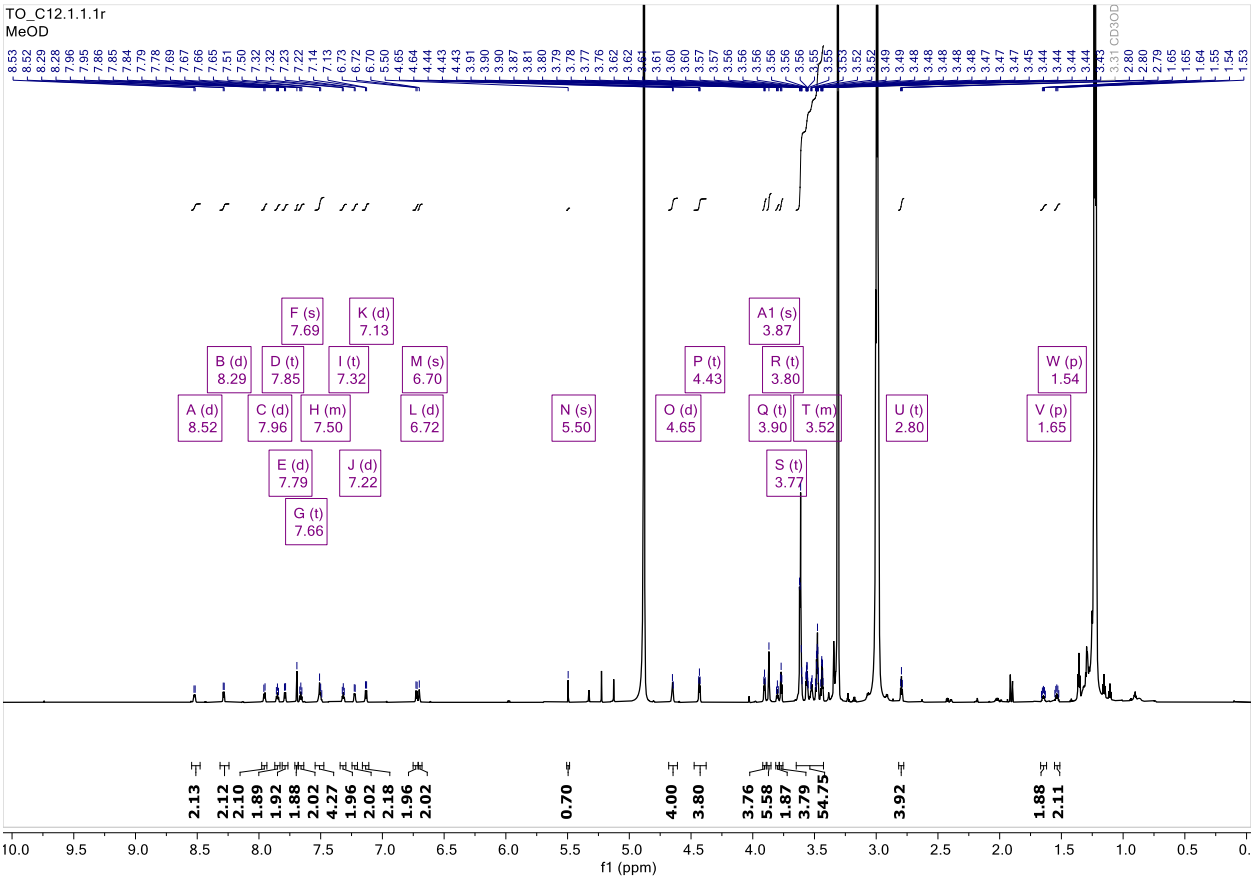

JJR_8_1.992.1.1r
161.06
150.29
138.92
136.50
130.98
130.93
128.28
123.43
120.26
71.73
71.56
71.46
71.02
69.20
58.39
54.81
51.75
49.00 CD3OD
20.30
N+
O
O
O
N3
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

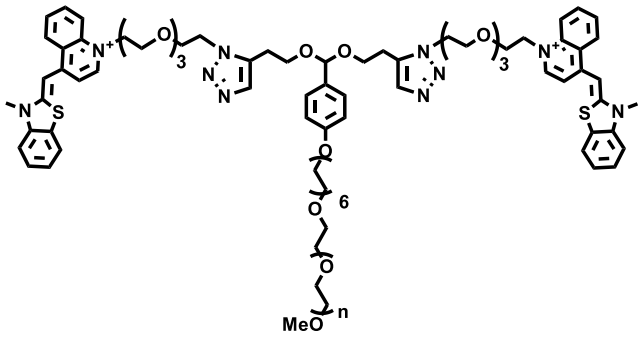

8.68
8.41
8.10
7.96
7.90
7.75
7.66
7.61
7.47
7.42
6.92
4.78
4.01
4.00
3.97
3.61
3.56
3.51
3.50
3.31 CD3OD
3.28
1.25
1.09
1.13
1.11
1.09
1.15
1.18
1.14
1.15
1.10
1.13
1.98
2.80
1.71
2.06
3.60
3.66
1.90
f1 (ppm)


159.23
148.17
143.21
139.03
136.03
131.39
126.54
125.13
123.70
123.21
122.95
122.74
120.73
116.03
110.76
106.33
86.26
68.89
68.68
68.62
68.53
68.10
66.45
52.66
48.75
31.17
N
S
N+
O
O
O
N3
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-1
f1 (ppm)

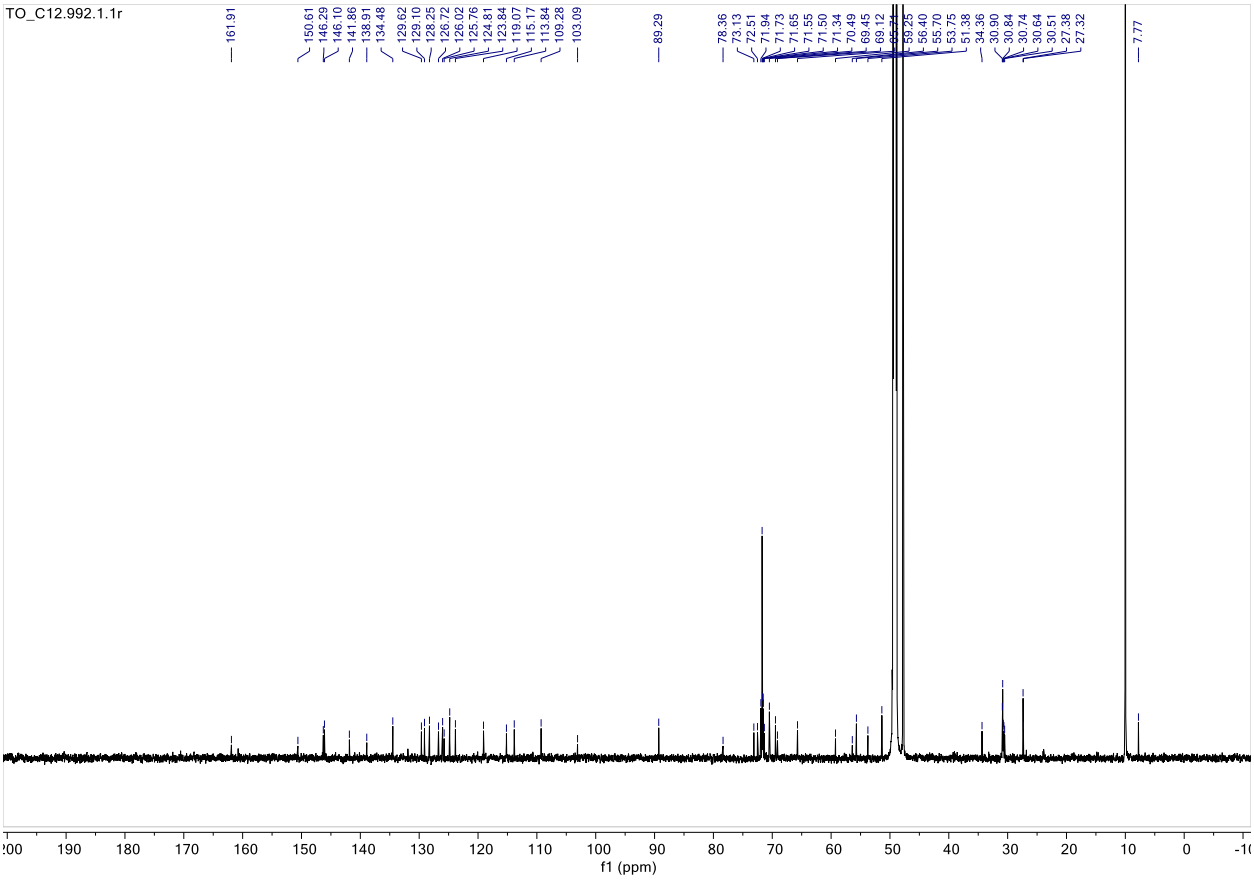


TCS


TO_C12.1.1.1r
MeOD
F (s) 7.69
K (d) 7.13
A1 (s) 3.87
B (d) 8.29
D (t) 7.85
I (t) 7.32
M (s) 6.70
P (t) 4.43
R (t) 3.80
W (p) 1.54
A (d) 8.52
C (d) 7.96
H (m) 7.50
L (d) 6.72
N (s) 5.50
O (d) 4.65
Q (t) 3.90
T (m) 3.52
U (t) 2.80
V (p) 1.65
E (d) 7.79
J (d) 7.22
S (t) 3.77
G (t) 7.66
2.13
2.12
2.10
1.89
1.92
1.88
2.02
4.27
1.96
2.02
2.18
1.96
2.02
0.70
4.00
3.80
3.76
5.58
1.87
3.79
54.75
3.92
1.88
2.11
f1 (ppm)


MeO
n
6
3
3

TCS


TO_C12.992.1.1r
161.91
150.61
146.29
146.10
141.86
138.91
134.48
129.62
129.10
128.25
126.72
126.02
125.76
124.81
123.84
119.07
115.17
113.84
109.28
103.09
89.29
78.36
73.13
72.51
71.94
71.73
71.65
71.55
71.50
71.34
70.49
69.45
69.12
58.25
56.40
55.70
53.75
51.38
34.36
30.90
30.84
30.74
30.64
30.51
27.38
27.32
7.77
f1 (ppm)

Chapter 2 – Delivery of Gene Editing Proteins CRISPR-Cas9 and Cre recombinase

A novel fluorescent single tailed surfactant was synthesized, which can disrupt endosomes, complex lipofectamine, and identify transfected cells. In this report, we designed and synthesized a novel fluorescent single tailed surfactant (termed FEDS), which can disrupt endosomes, complex lipofectamine, and can also identify cells that have been transfected. FEDS was able to increase the gene editing efficiency of lipofectamine/Cas9 ribonucleoprotein by 300% via a combination of fluorescent based enrichment and endosomal disruption.

In addition, we report here a PEG-poly-eosin block copolymer (PEG-pEosin) that can encapsulate proteins and release them in active form under mildly acidic conditions. A PEG-pEosin formulation composed of Cre and the endosomolytic protein LLO efficiently performed gene editing in cells and in the brains of mice after an intracranial injection.

Chapter 2 is based on the following published research studies:

Li J‡, **Røise JJ**‡, Zhang J‡, Yang J, Kerr DL, Han H, & Murthy N. (2019). *A novel fluorescent surfactant enhances the delivery of the Cas9 ribonucleoprotein and enables the identification of edited cells*. Chem Commun, 55, 4562-4565

Reinhard S‡, Han H‡, Tuma J, **Røise JJ**, Li IC, Li J, Lee HY & Murthy N. (2020). *A pH-sensitive eosin-block copolymer delivers proteins intracellularly*. Chem. Commun., 56, 14207-14210

2.1 Introduction

Protein-based therapeutics have the potential to revolutionize the field of therapeutics. With their superior selectivity and complex chemistries, they provide a gateway to treat diseases that small molecules are unable to. Proteins that hold desired properties as therapeutics include transcription factors, metabolic enzymes, and protein antigens; however, recent trends revolve around the use of the newly discovered CRISPR-Cas9 gene editing enzyme, which can be used to treat and possibly cure a plethora of genetic diseases. In addition to Cas9, zinc-finger nucleases (ZFNs) and Cre recombinase, along with their many variants, have also been investigated for *in vivo* gene editing[37]. However, although gene editing shows great potential in the field of biotherapeutics, the main bottleneck for their clinical use remains low delivery efficiencies. Currently, the main method of delivery of intact proteins to cells is through the use of viral vectors, which have the downside of high immunogenicity. As an alternative, mRNA and pDNA vectors encoding for said proteins are commonly utilized, but this strategy is associated with prolonged expression, which often leads to a proportionate increase in off-target editing. Subsequently, there is an urgent need for the development of efficient protein delivery vehicles for the field of protein therapeutics to reach its clinical potential.

In an effort to develop and explore new methods for protein delivery, this chapter will highlight two such efforts. Firstly, a fluorescent surfactant that is masked by a carboxylic acid was developed to successfully aid in the delivery of Cas9 ribonucleoprotein (Cas9 RNP) by Lipofectamine. Lipofectamine can be used for the delivery of Cas9 RNP due to the highly negative charge of the intact Cas9-sgRNA complex but has the downside of the toxic characteristics of Lipofectamine. The developed surfactant, **FEDS**, successfully increased the delivery efficiency of Cas9 RNP in the presence of Lipofectamine. In addition, **FEDS**, due to its fluorescent property, allowed for the selective sorting of edited cells, as cells which saw high levels of **FEDS** was found to have a proportionate increase in gene editing, and could therefore be sorted using fluorescent cell sorting (FACS).

Secondly, a polymeric material consisting of multiple Eosin Y moieties connected to the backbone through ketal linkers (**PEG-pEosin**) was prepared to aid in the delivery of Cre recombinase. Eosin Y is a dye which has broad protein-binding properties and was utilized here to bind Cre recombinase along with the protein Listeriolysin O, a cholesterol-dependent cytolysin (CDC) which disrupts cell membranes at endosomal conditions. The PEG-pEosin/Cre/LLO complex is taken up in cells through endocytosis, where the acetal backbone will hydrolyze along with the acidification of the endosome. When PEG-Eosin hydrolyzes, its binding efficacy to Cre and LLO will diminish, causing the release of free LLO and Cre in the endosome. The free LLO will then efficiently disrupt the endosomal membrane, causing cytosolic release of Cre recombinase. This strategy was used to deliver Cre both *in vitro* and *in vivo*.

2.2 Cas9 Delivery using a Fluorescently Traceable Surfactant (FEDS)

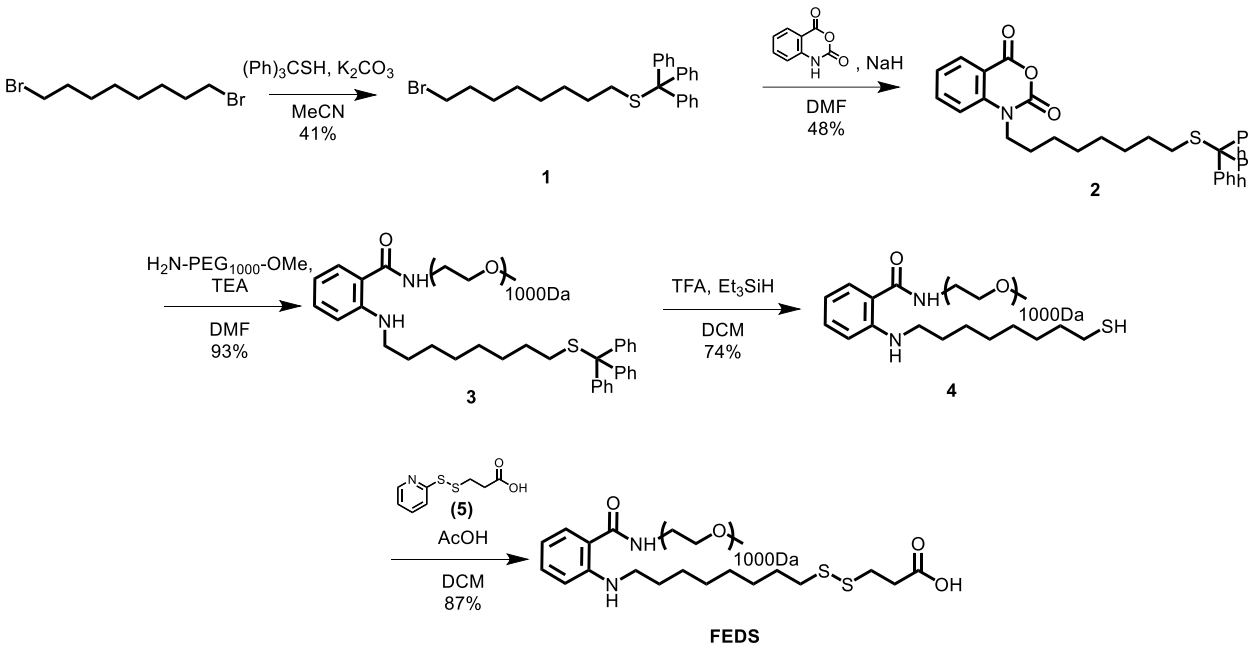
Introduction

Although therapeutic proteins have the potential to cure a large range of currently untreatable diseases, their clinical use has been largely hindered by a lack of efficient delivery strategies. Currently, one of the most promising enzymes for use in the clinical sector is the gene-editing enzyme CRISPR-Cas9, which can be engineered with a single guide RNA (sgRNA) to selectively edit parts of the genome through either Non-Homologous End Joining (NHEJ) or Homology Directed Repair (HDR) following a double-stranded break[37], to achieve anything from single-point mutations to replacements of full genes. However, despite the promise of Cas9, its clinical use has been limited due to its low delivery efficiencies[38]. Cas9 RNP is a protein complex larger than 160kDa, and although smaller variants exist, successful delivery of these are still mostly limited to viral vectors[39]. In addition, when editing genomes, it is often necessary to include markers to sort for edited cells. Currently, the common way of achieving this is through the inclusion of a fluorescent protein marker; however, this introduces additional bulk which can potentially lead to lower delivery efficacies.

Due to the current challenges in Cas9 RNP delivery, we therefore developed a new adjuvant that both increases Cas9 delivery efficiencies as well as serves as a fluorescent marker, allowing for cell sorting through FACS of edited cells. This compound, **FEDS**, consists of a surfactant similar to 9-nonoxynol, an FDA approved contraceptive, modified with a disulfide-carboxylic acid group on its termini (Scheme 12). This carboxylic group is designed to mask the hydrophobic portion of FEDS in physiological conditions, but be protonated in acidic endosomal conditions, leading to disruption of the endosomal membrane. In addition, FEDS is designed with a negative charge in order to associate with positive lipids commonly used for nucleic acid delivery, such as Lipofectamine. Importantly, the anthranilic acid amine scaffold has fluorescent properties, which should allow for cell sorting of cells with high levels of Cas9 RNP/FEDS uptake using FACS.

Synthesis and Characterization

FEDS was synthesized in six steps according to the scheme shown in Scheme 12. Firstly, isatoic anhydride was alkylated with 1-bromo-8-tritylmercaptooctane in the presence of sodium hydride to form **2**. A ring-opening reaction on **2** with amino-PEG (Mw = 1000Da) was utilized to form **3**. This was followed by deprotection of the trityl thiol group to form **4**, and finally **FEDS** was synthesized through a disulfide-thiol exchange with 3-(pyridin-2-yldisulfaneyl)propanoic acid. **FEDS** contain an anthranilic amide moiety which has an excitation maximum at 350 nm and an emission maximum at 460 nm. The quantum yield of FEDS was also measured, using quinine sulfate in 0.05 M H_2SO_4 as a standard, and was found to have a quantum yield of 0.15.



Scheme 12: Synthetic scheme for the preparation of FEDS.

Hemolysis

FEDS is designed to selectively enhance the transfection ability of lipofectamine by disrupting endosomes before degradation of Cas9 RNP. However, FEDS must also not cause membrane disruption at pH 7.4, as this would cause FEDS to enter and disrupt the outer cell membrane. It is therefore necessary that FEDS shows pH-dependent membrane disruption. A hemolysis assay was utilized to determine the membrane disruptive efficiencies of FEDS at pH 7.4 and pH 5.5, which is the rough pH of a late endosome. FEDS was added to rabbit red blood cells at concentrations ranging from 0-10 mg/mL. The samples were then incubated at 37°C for 30 minutes, before they were centrifuged at 100g for 10 minutes, and absorption of the supernatant was measured at 541nm to probe for hemoglobin release. As seen in Figure 39a, FEDS shows great pH dependency, with no significant hemolysis observed up to 10 mg/mL in pH 7.4, but up to 70% hemolysis at pH 5.5. Further, we performed a control study with the reduced FEDS precursor molecule **4**, which should show pH-independent hemolysis due to the lack of a carboxyl group. Figure 39b shows that **4** is membrane disruptive at both pH 5.5 and 7.4. This supports the hypothesis that FEDS is shows pH-dependent membrane disruption mostly due to its carboxylic acid moiety. These results demonstrate that FEDS should selectively disrupt endosomal membranes, as it is inactive at extracellular pH values (pH 7.4), and active at endosomal pH conditions (pH 5.5).

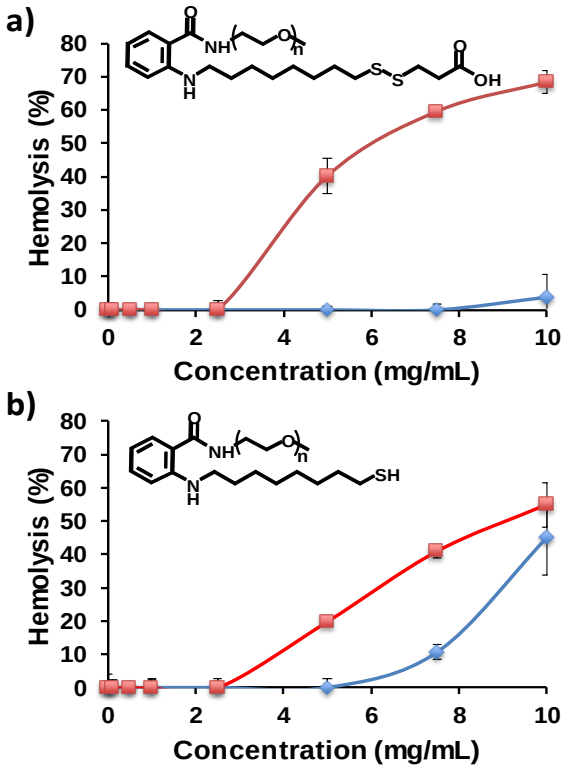


Figure 39: Hemolysis efficiency of FPS under pH 5.5 (□) and pH 7.4 (□) at different concentration. a) FPS caused strong hemolysis at pH 5.5, but no hemolysis at pH 7.4. b) The reduced FPS is membrane disruptive at both pH 5.5 and pH 7.4. Absorbance at 541 nm of rabbit RBCs incubated in water followed by three freeze-thaw cycles was set as 100% hemolysis standard. Figure reprinted from Li et al[5].

Transfection and Cell Sorting

Upon endocytosis of Cas9 RNP, it is sorted through the endocytic pathway and finally degraded in lysosomes. FEDS is designed to disrupt endosomes, releasing Cas9 RNP into the cytosol before it is degraded. For FEDS to associate with Cas9 RNP, it is necessary to use it as an adjuvant to a cationic lipid. In this study we therefore investigated if FEDS could enhance the delivery of Cas9 RNP/Lipofectamine in HEK cells. A GFP-expressing HEK cell line with a doxycycline P_{tet} promoter was used as a model cell line (doxGFP-HEK), and the Cas9 RNP was designed to knock out the GFP gene.

Cas9 RNP and Lipofectamine 2000 (5 µg Cas9 per 1µL Lipofectamine) were added 100µg of FEDS in 150 µL Opti-MEM medium and added to the doxGFP-HEK cells, and the percentage of GFP negative cells was determined after doxycycline activation via flow cytometry. As seen in Figure 40, FEDS successfully enhances Cas9 RNP delivery in HEK cells. Cas9 RNP with lipofectamine alone lead to 13.2% GFP knockdown, while in the presence of FEDS, 24.2% GFP knockdown was observed. These numbers were further corroborated by a T7E1 endonuclease assay.

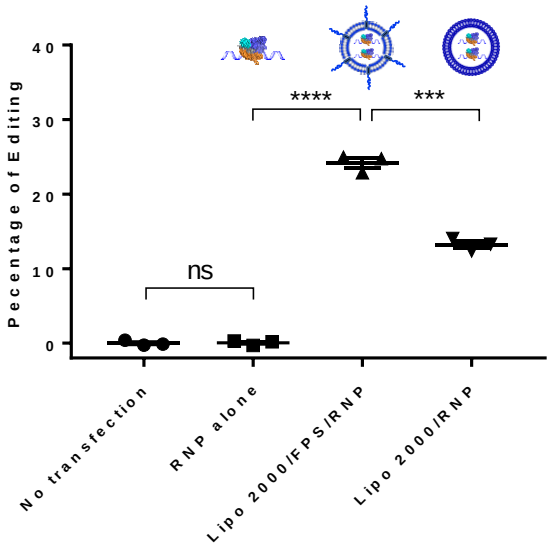


Figure 40: FPS enhances the ability of lipofectamine to deliver the Cas9 RNP. Knockdown of GFP in RT-HEK cells treated with different Cas9 RNP formulations. FPS dramatically improved the ability of lipofectamine 2000 (Lipo 2000 /FPS/RNP vs Lipo 2000/RNP) to deliver the Cas9 RNP. Figure reprinted from Li et al[5].

An important feature of FEDS is its fluorescent properties. We hypothesized that cells with high levels of internalized FEDS should show a proportionate increase in GFP knockdown. This would enable us to enrich for cells with high levels of gene editing. To probe for this, doxGFP-HEK cells were transfected as described above, and the cells were sorted using FACS by high levels of FEDS-associated fluorescence. As seen in Figure 41, cells treated with Cas9 RNP using only Lipofectamine were shown to have a knockdown rate of 16.7%. This falls in line with the knockdown efficiency of cell with low levels of FEDS (low-FEDS) which showed 20.6% of GFP knockdown. In contrast to this, high-FEDS cells showed a drastic improvement, with 50.9% GFP knockdown. This illustrates that FEDS allows for fluorescence-based cell sorting of gene edited cells.

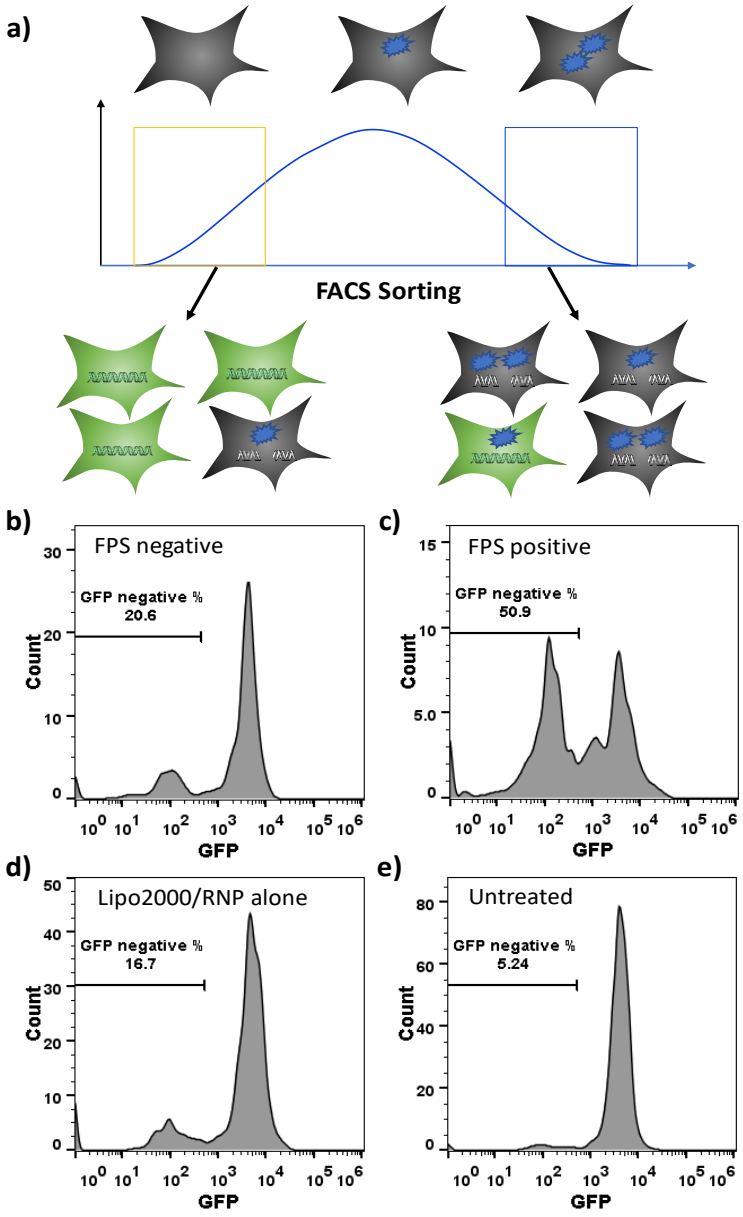


Figure 41: FPS can enrich for cells that have been gene edited. a) Schematic illustration of FACS sorting and enrichment of edited cells. Cells that have internalized large amounts of FPS complexed to lipofectamine/Cas9 RNP can be identified by flow cytometry. These cells have a higher rate of gene editing, and FPS enables the enrichment of gene edited cells. b) GFP expression of RT HEK cells sorted from FPS negative cells after Dox activation. c) GFP expression of FPS positive RT HEK cells after Dox activation. d) GFP expression of RT HEK cells transfected by lipo2000/RNP after Dox. e) GFP expression of RT HEK cells after Dox activation. Figure reprinted from Li et al[5].

Conclusion

Lysosomal degradation of the Cas9 RNP complex remains a major bottleneck in clinical gene editing. In this report we presented FEDS, a novel adjuvant that increases Cas9 RNP delivery in the presence of Lipofectamine 2000. In addition, FEDS allowed for cell sorting to achieve enrichment of cells with high levels of gene editing due to its fluorescent properties.

2.3 Cre Recombinase Delivery using a Protein Binding Polymer

Introduction

Therapeutic proteins hold great potential in the field of clinical medicine, as their complex chemistry allows for reactivity and selectivity that is currently not possible to achieve with small molecule drugs. Over 200 protein and peptide drugs have been approved by the FDA since the 1980s[40]; however, these therapeutics have in large been limited to extracellular targets. The reason for this apparent lack in intracellular protein therapeutics is due mostly to their low delivery efficacies. Although extensive research has been done on developing protein delivery vehicles, including nanoparticles, polymer formulations, micelles, and liposomes.[41], these strategies have yet to cause a significant impact on protein therapeutics in a clinical setting. This is in direct contrast to the field of nucleic acid delivery, which has seen a great amount of success, most recently with the development of mRNA-based vaccines for COVID-19 which utilize lipid nanoparticle technology for delivery[42]. In order for protein therapeutics to reach a similar level of success, there is a great need for new technologies for intracellular therapeutic protein delivery.

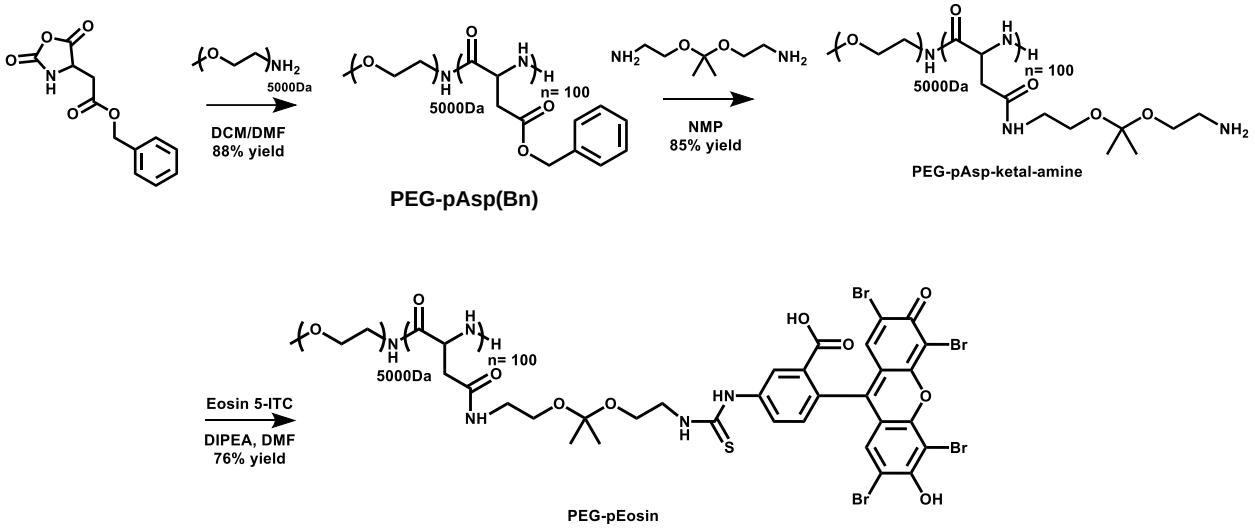
Cre recombinase is gene editing enzyme that cuts out segments of the genome. Although Cre recombinase itself can only be used to remove segments between two 34-basepair specific LoxP sites, limiting it to model systems, efforts to mutate Cre recombinase to target clinically relevant genes have been successful. For example, Cre was mutated to target HIV-1 long terminal repeats[43], highlighting their potential use as a future protein therapeutic.

Listeriolysin O (LLO) is a cholesterol-dependent cytolysin found in *Listeria monocytogenes* that binds to cholesterol-containing membranes and forms pores up to 350nm in diameter[44]. CDCs serve as methods for bacterial vacuolar escape, but they have also been used as delivery vehicles for proteins[45] and nucleic acids[46] due to their pore-forming properties. A distinguishing property of LLO over other CDCs is its optimal activity at pH 5.5,[47] which makes it a prime candidate as an endosomal disruptive agent, seeing as the endosomal pH ranges from 5.0-6.5. This allows for selective pore-formation on the endosomal membrane, causing cytosolic release of its endosomal cargo. However, for LLO to aid in the delivery of Cre recombinase, they must be delivered to the same endosomal compartment.

In an effort to develop a universal protein delivery vehicle, we created a polymeric material (PEG-pEosin) with a protein binding moiety (Eosin Y) attached to the backbone through an acid-degradable linker (Scheme 13). This allows for protein to be bound before cell delivery, and subsequently released under endosomal conditions. Eosin Y binds a range of proteins with dissociation constants of between 10^{-3} to 10^{-6} M,[48,49] which should increase due to the multivalent nature of polymers, similarly to polycationic polymers used for nucleic acid binding. Cre recombinase was used as a model protein for delivery, and LLO was co-encapsulated for endosomal disruption.

Synthesis and Characterization

PEG-pEosin is a block copolymer composed of a PEG-b-poly(aspartic acid) that has multiple Eosin Y moieties conjugated to its aspartic acid segment through a pH-degradable ketal linker. PEG-pEosin was synthesized as shown in Scheme 13 and described in the Appendix. Briefly, **PEG-pAsp(Bn)** was prepared by treatment of benzyl 2-(2,5-dioxooxazolidin-4-yl)acetate with MeO-PEG_{5000}-NH_2 at a 100:1 ratio. **PEG-pAsp-ketal-amine** was then prepared by reaction of 2,2-Bis(aminoethoxy)propane to displace the aspartic benzyl group. Finally, Eosin Y was attached to **PEG-pAsp-ketal-amine** through an isothiocyanate linker to form PEG-pEosin. The final material was purified by an ether/hexane precipitation (50/50), followed by size-dependent exclusion througha 7 kDa spin filter to remove unreacted Eosin Y. ^{1}H-NMR revealed a >90% Eosin Y functionalization rate.



Scheme 13: Synthetic scheme for the preparation of PEG-pEosin.

Protein Binding

A vital property of PEG-pEosin is its reversible binding to proteins. BSA was used as a model protein to investigate PEG-pEosin binding, as well as acid-dependent release, due to its known interactions with Eosin Y[50]. Protein gel electrophoresis was used to probe for binding by looking at molecular weight shifts upon addition of PEG-pEosin to BSA. As shown in Figure 42, addition of PEG-pEosin to BSA leads to concentration-dependent retention on a PAGE gel. Significant retention is seen upon addition of 2eq of PEG-pEosin, and at 6 equivalents the BSA band is nearly completely retained. In comparison, BSA which had been reacted with Eosin 5-isothiocyanate (Eosin-NCS) did not show a significant shift. These results were further corroborated by dynamic light scattering (see *Appendix B2*).

In addition to binding BSA, PEG-pEosin must also be able to release free BSA under endosomal conditions. This property was assayed by adjusting the pH of BSA/Peg-pEosin (6eq of PEG-pEosin) to pH conditions similar to those found in endosomal compartments (pH 5.5-6.8) followed by a 4hr incubation before assaying retention on a PAGE gel. As shown in Figure 42, PEG-pEosin successfully releases BSA in endosomal pH condition. After a 4-hour incubation

period at pH 6.8 BSA is partially released, and at conditions of pH <6.4 BSA is nearly completely recovered. This illustrates that the binding of proteins by PEG-pEosin is dependent on its multivalency, which is lost in endosomal conditions. This should allow for the use of PEG-pEosin as a protein vehicle, opening up possibilities for endosomal disruption and cytosolic delivery of protein therapeutics.

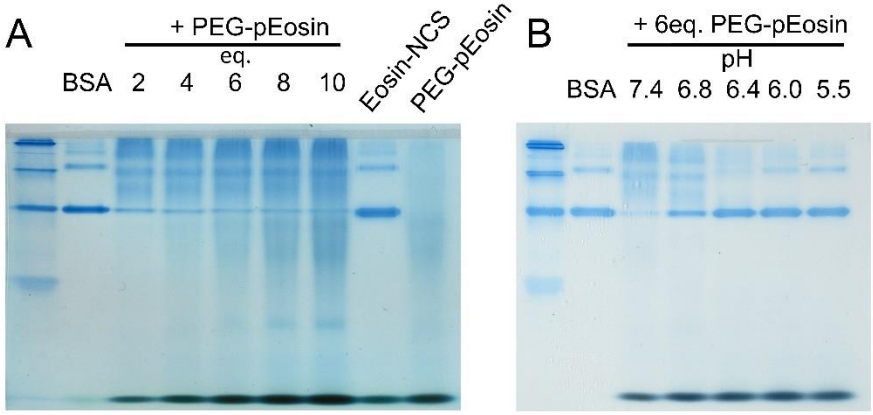


Figure 42: PEG-pEosin binds proteins. (A) Native PAGE showing protein binding of the PEG-pEosin polymer to bovine serum albumin (BSA) at different molar ratios. PEG-pEosin/BSA molar ratios > 6 result in efficient binding of BSA. (B) PEG-pEosin releases BSA under acidic conditions, due to hydrolysis of the ketal linkage. Incubation of BSA+PEG-pEosin at pHs below 6.8 results in release of BSA. Figure reprinted from Reinhard et al[51].

LLO Pore-forming Activity

An important feature of PEG-pEosin mediated protein delivery is the incorporation of LLO as an endosomal disruptive protein. This is made possible by the optimal pH range of LLO, which corresponds to that in endosomal conditions. As seen in Figure 43, LLO does not show hemolytic efficacy at concentrations up to 100ng/mL at pH 7.4, while achieving 80% hemolysis at pH 5.5. However, despite its pH selectivity, LLO still causes considerable toxicity at pH 7.4 at higher concentrations. It is therefore beneficial if PEG-pEosin releases LLO only under endosomal condition. In order to probe for this feature, the activity of PEG-pEosin complexed with LLO was compared to free LLO. A hemolysis assay in which hemoglobin release is measured was utilized as a measure of membrane disruptive ability.

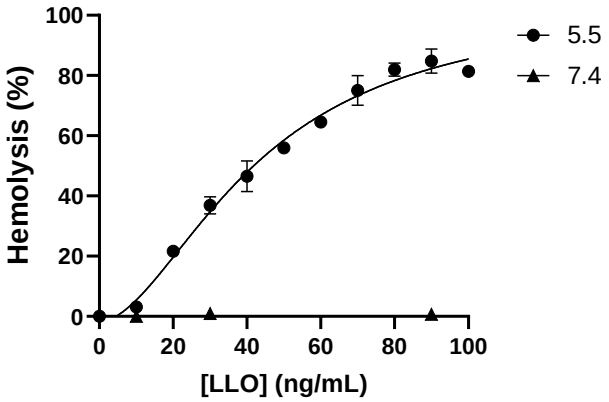


Figure 43: Hemolytic efficacy of LLO at pH 5.5 and pH 7.4. LLO shows pH-dependent hemolysis, with nearly 80% hemolytic efficacy at 100ng/mL of LLO at pH 5.5. Hemolysis was not observed in the same concentration range at pH 7.4.

LLO (0.5µg/mL) was complexed to PEG-pEosin (6 equivalents) under a range of pH conditions (7.4, 6.8, and 5.5), and membrane disruptive efficiencies were measured. As seen in Figure 44, LLO at 0.5µg/mL causes complete hemolysis after only 30 minutes across the investigated pH range. However, when LLO is complexed with PEG-pEosin, hemolysis is considerably slower, with less than 50% hemolysis after a 2-hour incubation at pH 7.4. This shows that LLO is prevented from disrupting cell membranes when it is complexed with PEG-pEosin. In addition, at pH 6.8 and 5.5, the ketal linkages in the PEG-pEosin backbone hydrolyze, allowing for release of LLO and rapid membrane disruption. Importantly, Figure 44 also shows that LLO binds reversibly, and does not lose its activity during the complexation process. This is an important feature for LLO to be an efficient endosomal disruptive agent.

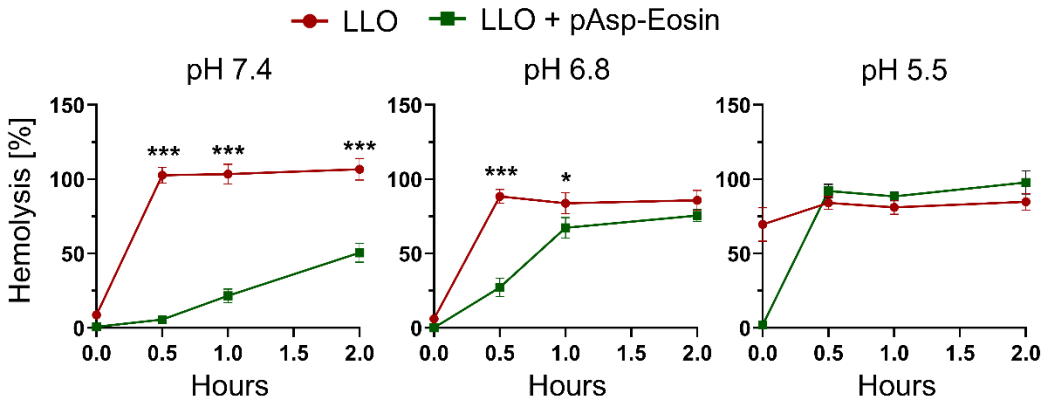


*Figure 44: PEG-pEosin reversibly inhibits the activity of LLO. Hemolysis assay of free and PEG-pEosin-formulated LLO at pH 7.4, 6.8 and 5.5, over a course of 2 hours. PEG-pEosin-formulated LLO has minimal hemolysis as pH 7.4 (30 minutes), but recovers activity at pH 5.5 (30 minutes). Negative control (hemoglobin release from PBS-treated erythrocytes) was set to 0 %. Triton X treatment served as positive control and was set to 100 %. ***: p≤0.001, *: p≤0.05. Figure reprinted from Reinhard et al[51].*

Cre Delivery

The PEG-pEosin polymer was then investigated for its ability to deliver proteins to cells. For this purpose, a 3T3 Ai9 cell line which includes a stop sequence between two *loxP* sequences, which is removed by Cre recombinase. The cell line is engineered so that removal of the stop sequence leads to tdTomato expression. The 3T3 Ai9 cell line functions as a proof-of-concept cell line for intracellular Cre recombinase delivery.

Formulations of Cre recombinase and LLO were added 6 equivalences of PEG-pEosin. The experiments were done with 8µg/mL (143nM) or 12µg/mL (214nM) of LLO, 3.8µg/mL of Cre recombinase (100nM), and 1.46µM or 1.88µM of PEG-pEosin, respectively. The complexes were added to Ai9 cells and incubated for 48 hours humidified incubator containing 5% CO_2 before tdTomato expression levels were assayed using flow cytometry (*Attune NxT, Invitrogen*).

As shown in Figure 45A, PEG-pEosin+LLO successfully delivered Cre recombinase to Ai9 cells, with 12.3% and 20.5% tdTomato-positive cells for 143nM and 214nM of LLO, respectively. In addition, no tdTomato expression was detected for Cre or PEG-pEosin individually. Importantly, a physical mixture of Cre recombinase and LLO did not yield significant tdTomato expression. This illustrated that PEG-pEosin complexation is necessary for successful protein delivery. Toxicity of the Peg-pEosin+LLO/Cre complex was then measured using a standard resazurin assay. As seen in Figure 45B, PEG-pEosin complexation causes drastic increase in cell viability compared to free LLO+Cre; however, toxicity of 40-50% was observed for 143nM and 214nM LLO-containing complex, which could be due to LLO leakage.

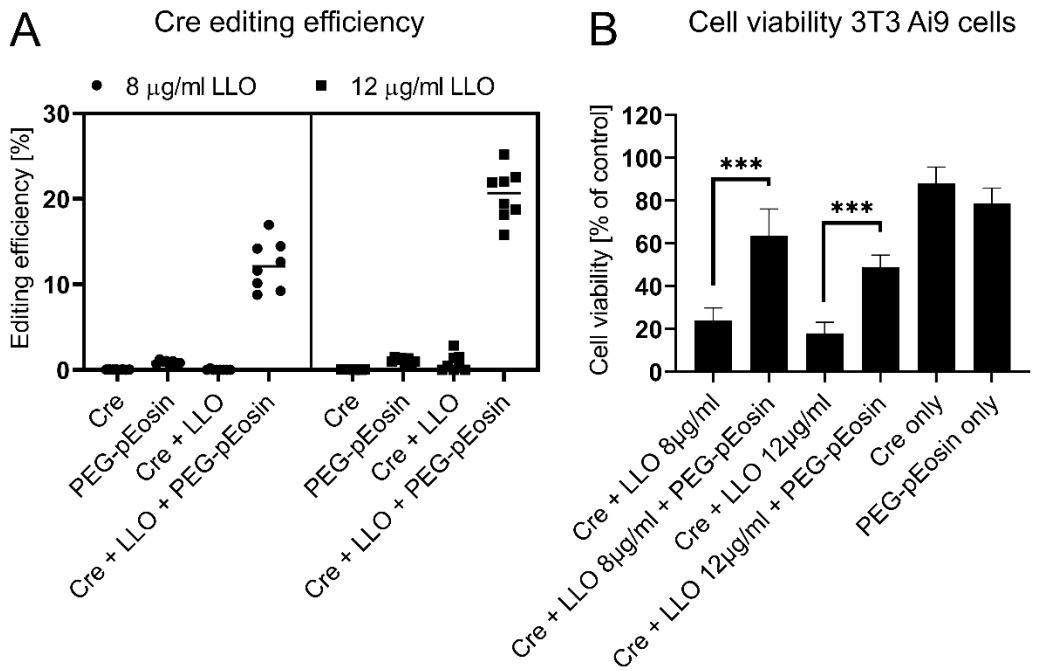


Figure 45: ***PEG-pEosin-formulated Cre can efficiently edit cells****. (A) Gene editing efficiency of Cre recombinase co-formulated with two different concentrations of LLO and PEG-pEosin. Cre+LLO+PEG-pEosin can edit 3T3 Ai9 cells significantly better than Cre only, PEG-pEosin and Cre + free LLO. Gene editing was determined by flow cytometry and analyzed the TdTomato-positive cells. (B) Cell viability of 3T3 Ai9 cells after incubation with Cre recombinase co-formulations with two different concentrations of*

*LLO and PEG-pEosin, Cre only, PEG-pEosin and Cre + free LLO. ***: p≤0.001. Figure reprinted from Reinhard et al[51].*

2.4 Conclusion

In this chapter we presented two new vehicles for the delivery of proteins, **FEDS** and **PEG-pEosin**. **FEDS** is able to associate with Cas9/Lipofectamine vesicles to increase cytosolic delivery and allow for fluorescent sorting to enrich for transfected cells. **PEG-pEosin** is able to reversibly bind to a range of proteins, including BSA, LLO, and Cre recombinase, releasing its cargo in endosomal conditions. This was illustrated by delivery Cre recombinase to Ai9 cells using a PEG-pEosin+LLO/Cre complex, releasing both LLO and Cre in the endosome, followed by LLO-mediated endosomal disruption and cytosolic Cre delivery.

Appendix B1 – Chapter 2.2

B1.1 Methods and Materials

All commercially available compounds were purchased from Sigma Aldrich or Alfa Aesar and all solvents were purchased dry from Sigma Aldrich with a Sure seal system. NMR spectra were recorded on a Bruker Avance 400 spectrometer. The chemical shifts (δ) are given in parts per million relative to CDCl3 (7.26 ppm for 1H) CDCl3 (77.16 ppm for 13C). Flash column chromatography was performed on silica gel (particle size 200-300 mesh) or using a Biotage Isolera Spektra system. Particle size and zeta potential were measured with a Zetasizer Nano-ZS (Malvern Instruments Ltd., UK). Flow cytometry was performed on an Attune NxT flow cytometer (Invitrogen). Cell sorting was performed on an Influx cell sorter (BD). The fluorescence of FEDS was measured by a Tecan Infinite M200 microplate reader. Fetal Bovine Serum (FBS), Opti-MEM, Ham's F-12 Nutrient Mix and Dulbecco's Modified Eagle Medium (DMEM) were purchased from Gibco. Accutase was purchased from Invitrogen. NucSpot® Live 488 Nuclear Stains was purchased from Biotium. The Genomic extraction kit and PCR product purification kit were purchased from Qiagen. T7 Endonuclease 1 (T7E1) was purchased from NEB. Rabbit red blood cells were purchased from Hemostat Laboratories. Monoclonal HEK-RT3-4 reporter cells expressing GFP under the control of the Ptet promoter, here referred to as “RT HEK”, was a generous gift from GenEdit 1 . Primary Ai9 mouse myoblast cells, were also a generous gift from GenEdit1 . Cas9-NLS and Cas9-mCherry fusion proteins were purchased from the UC Berkeley QB3 MacroLab core facility. CrRNA and tracrRNA were purchased from IDT. The protospacer sequences of crRNA were described as below:

GFP: GCTGAAGCACTGCACGCCAT.

Ai9: AAGTAAAACCTCTACAAATG.

B1.2 Protocols

DLS measurements

DLS experiments were performed with a Zetasizer Nano instrument (Malvern Instruments Ltd., UK) equipped with a 10-mW helium-neon laser (λ = 632.8 nm) and thermoelectric temperature controller at 25 °C. Measurements were taken at a 90° scattering angle. The Cas9 RNP and FEDEX were dissolved in a pH 7.2 10 mM HEPES buffer with a Cas9 RNP concentration of 0.05 mg/mL.

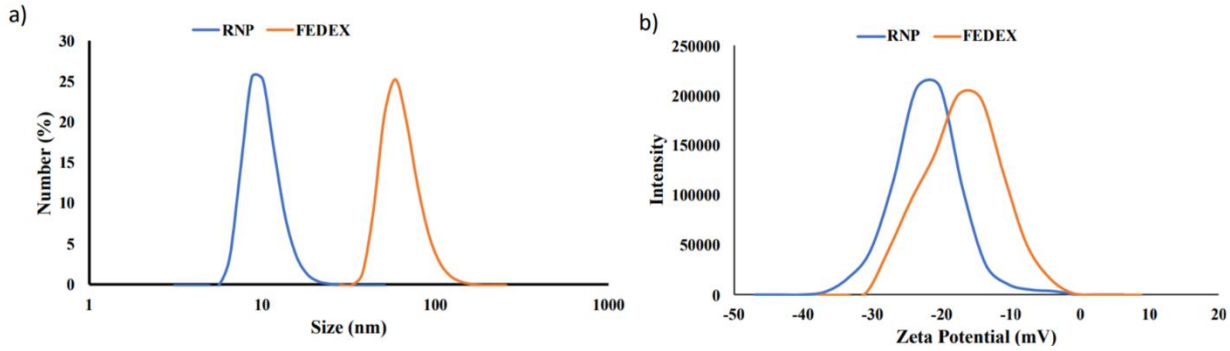


Figure 46: Size and Zeta Potential of the Cas9 RNP and FEDEX.

Hemolysis assay

Fresh rabbit red blood cells (RBC) were washed three times with 150 mM PBS. The RBCs were then suspended in 1 ml of 0.1 M sodium phosphate buffer at the appropriate pH. FEDS and reduced FEDS were dissolved in 0.1 M sodium phosphate buffer, at different concentrations. The hemolysis assay was performed by adding FEDS to 108 RBCs suspended in 1 ml of the appropriate pH phosphate buffer. The RBCs were inverted several times for mixing, and incubated in a 37°C water bath for 60 min. The cells were centrifuged at 1000 g for 5 min and the absorbance of the supernatant was measured at 541 nm. To determine a 100% hemolysis, 108 RBCs were lysed by suspending them in distilled water. The control was 108 RBCs in buffer. All hemolysis experiments were done in triplicate.

Cell culture

HeLa cells and RT HEK cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (Gibco), and 1% penicillin−streptomycin (Invitrogen) and maintained at 37 °C in a humidified incubator (Thermo Electron Corporation) containing 5% CO2. When the cells reached 80% confluency, they were passaged using an Accutase solution (Invitrogen); media was changed every 2 − 3 days. Ai9 myoblast cells were cultured in Ham's F-12 Nutrient Mix (Gibco) supplemented with 20% fetal bovine serum (Gibco), and 1% penicillin−streptomycin (Invitrogen) and 10 ng/mL of bFGF (Sigma Aldrich) (myoblast growth medium), and maintained at 37 °C in a humidified incubator (Thermo Electron Corporation) containing 5% CO2. When the cells reached 80% confluency, they were passaged

using an Accutase solution (Invitrogen) and seeded to a Matrigel coated flask. Media was changed every 2 − 3 days.

Flow cytometry

RT HEK cells were seeded in 48 well plates at a density of 20K per well. After 24 hours, 150 µL Opti-MEM (Gibco) containing Cas9 RNP (5 µg Cas9) mixed with 1µL of lipofectamine 2000 and 100 µg of FEDS was added to the RT HEK cells. The cells were transfected for 48 hrs at 37 °C in a humidified incubator containing 5% CO2 and medium was replaced with fresh DMEM medium supplemented with 10% FBS, 1% penicillin-streptomycin and 1 µg/mL doxycycline. RT HEK cells were detached by Accutase (Invitrogen) and the GFP signals were analysed by Attune NxT flow cytometer (Invitrogen). Gene editing efficiency was determined by the percentage of GFP negative cells. Primary Ai9 mouse myoblast cells were seeded in 48 well plates at a density of 20K per well. After 24 hours, 150 µL of Opti-MEM (Gibco) containing Cas9 RNP that targeted the Ai9 STOP cassette (5 µg Cas9) mixed with 1µL of lipofectamine 2000 and 100 µg of FEDS, was added to the Ai9 myoblast cells. The cells were transfected for 24 hrs at 37 °C in a humidified incubator containing 5% CO2, and the media was replaced with myoblast growth medium and the cells were grown in myoblast growth medium for another 24 hrs. The myoblasts were detached with Accutase (Invitrogen) and the TdTomato signal was analysed with a Attune NxT flow cytometer (Invitrogen). The gene editing efficiency was determined by identifying the percentage of TdTomato positive cells.

Cell Viability

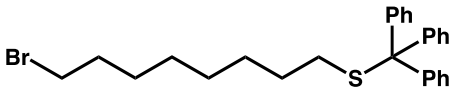
HEK 293T cells were seeded in 96 well plates at a density of 5K per well and incubated at 37 °C in a humidified incubator containing 5% CO2 for 24 hrs. 1 µL of lipofectamine 2000, 100 µg of FEDS or 1 µL lipofectamine + 100 µg FEDS were added to 100 µL of DMEM medium supplemented with 10% FBS and 1% penicillinstreptomycin and the cells were incubated with the solution for 24 hours. The cells were further washed with PBS and incubated with 100 µL fresh medium containing resazurin (Alamar Blue) (0.01 mg/mL) for another 4 h. The cell viability was then determined by measuring the fluorescence intensity of each well (Ex = 550 nm, Em = 595 nm) using a microplate reader. Untreated cells and fresh medium were used as controls for a 100% and 0% cell proliferation, respectively. Double concentration (2X) of lipofectamine 2000, FEDS and lipofectamine 2000 + FEDS were also added to the cells for the viability assay.

Cell Sorting

RT HEK cells were seeded in 48 well plates at a density of 20K per well. After 24 hours, 150 µL Opti-MEM (Gibco) containing Cas9 RNP (5 µg Cas9) mixed with 1µL of lipofectamine 2000 and 100 µg of FEDS was added to RT HEK cells. The cells were transfected for 24 hrs at 37 °C in a humidified incubator containing 5% CO2 and the medium was replaced with fresh DMEM medium supplemented with 10% FBS, 1% penicillin-streptomycin and 1 µg/mL doxycycline. RT HEK cells were detached by Accutase (Invitrogen) and sorted with an Influx cell sorter (BD)

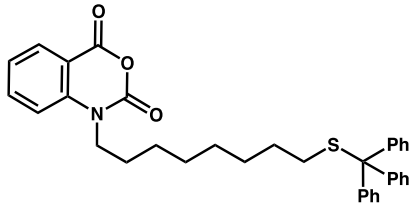
according to the intensity of FEDS. Both FEDS positive and negative cells were collected and cultured for further flow cytometry analysis.

B1.3 Synthesis



1

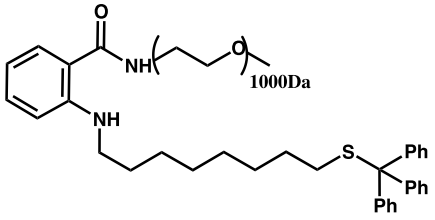
1-Tritylthiol-8-bromo-octane (1). 1,8-dibromooctane (2.76g, 10mmol), triphenylmethanethiol (6.8g, 25mmol) and potassium carbonate (11 g, 80 mmol) were added into acetonitrile (50 mL) in a 250 mL round-bottom flask and stirred at room temperature for 20 h. The solvent was then removed under vacuum. The resulting mixture was then re-dissolved in 150 mL CH2Cl2 and washed with water and brine, dried over anhydrous Na2SO4 and concentrated under reduced pressure. The crude product was purified by silica column chromatography (hexane/CH2Cl2, 9/1) to obtain 1-Tritylthiol-8-bromo-octane (1.9 g, 41%). 1H NMR (400 MHz, CDCl3) δ 7.42 (d, J = 7.9 Hz, 6H), 7.30-7.26 (m, 6H), 7.23-7.19 (m, 3H), 3.39 (t, J = 6.8 Hz, 2H), 2.14 (t, J = 7.3 Hz, 2H), 1.82 (p, J = 7.0 Hz, 2H), 1.38 (m, 4H), 1.29-1.09 (m, 6H). 13C NMR (101 MHz, CDCl3) δ 145.1, 129.6, 127.8, 126.5, 77.4, 77.1, 76.7, 66.4, 34.1, 32.8, 32.0, 29.0, 28.9, 28.6, 28.1. HRMS (EI) m/z calculated for: [C27H31BrS]+: 466.1330. Found: 466.1299



2

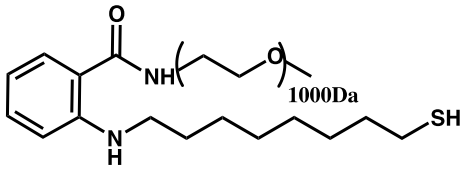
N-(1-tritylthio-)octylisatoic anhydride (**2**). NaH (90 mg, 2.1 mmol) was slowly added into a DMF solution (10 mL) of isatoic anhydride (0.3 g, 1.9 mmol) at 0°C and stirred for 10 min. To this solution, 1-Tritylthiol-8-bromo-octane (1) (0.9 g, 1.9 mmol) was then added dropwise and the resulting mixture was stirred for 24 h at room temperature. The reaction mixture was diluted with 150 mL EtOAc, washed with saturated ammonium chloride solution and brine, dried over anhydrous Na2SO4, and filtered. The filtrate was concentrated under reduced pressure and was purified by silica column chromatography (hexane/EtOAc, 8/2) to obtain N-(1-tritylthio-)octylisatoic anhydride as white powder (0.5 g, 48%). 1H NMR (400 MHz, Chloroform-d) δ 8.21 (dd, J = 7.8, 1.6 Hz, 1H), 7.79 (ddd, J = 8.7, 7.3, 1.7 Hz, 1H), 7.51 – 7.38 (m, 6H), 7.36

– 7.28 (m, 7H), 7.28 – 7.21 (m, 3H), 7.18 (d, J = 8.5 Hz, 1H), 4.11 – 4.01 (m, 2H), 2.18 (t, J = 7.3 Hz, 2H), 1.83 – 1.68 (m, 2H), 1.42 (m, 4H), 1.37 – 1.15 (m, 6H). 13C NMR (101 MHz, CDCl3) δ 158.7, 147.8, 145.2, 141.5, 137.3, 131.2, 129.7, 127.9, 126.7, 124.0, 114.0, 112.0, 77.5, 77.2, 76.8, 66.5, 45.1, 32.1, 29.1, 29.0, 28.7, 27.0, 26.7. HRMS (ESI) m/z calculated for: [C35H35NO3SNa]+: 572.2230. Found: 572.2224



3

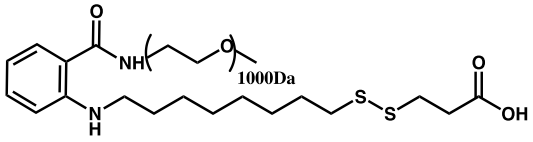
Compound **3**. Amino-PEG1000-OMe (200 mg, 0.2 mmol) and triethylamine (90 μL, 0.64 mmol) are dissolved in DMF (3 mL) and stirred for 10 min at room temperature. N-(1-tritylthio-)octylisatoic anhydride (**2**) (110 mg, 0.2 mmol) was then added into the solution and the reaction mixture was stirred at room temperature overnight. The solvent was then removed under reduced pressure and the resulting crude was purified by silica column chromatography (CH2Cl2/MeOH, 9/0.5, v/v) to obtain 3 as viscous colorless oil (290 mg, 93%). 1H NMR (400 MHz, CDCl3) δ 7.38 – 7.34 (m, 6H), 7.25 – 7.21 (m, 6H), 7.18 – 7.14 (m, 3H), 6.71 (s, 1H), 6.63 (d, J = 8.4 Hz, 1H), 6.52 (t, J = 7.6 Hz, 1H), 3.62 – 3.57 (m, 88H), 3.51 (dd, J = 5.8, 3.6 Hz, 2H), 3.34 (s, 3H), 3.06 (t, J = 7.1 Hz, 2H), 2.09 (t, J = 7.3 Hz, 2H), 1.59 (p, J = 7.2 Hz, 2H), 1.35 – 1.30 (m, 4H), 1.22 – 1.12 (m, 6H). 13C NMR (101 MHz, CDCl3) δ 145.0, 132.7, 129.6, 127.8, 127.6, 126.5, 77.5, 77.2, 76.8, 71.9, 70.6, 70.5, 70.3, 69.8, 66.3, 59.0, 39.3, 32.0, 29.2, 29.0, 28.9, 28.6, 27.1. HRMS (ESI) m/z calculated for n=17: [C69H109N2O18S]+: 1285.7391. Found: 1285.7427



4

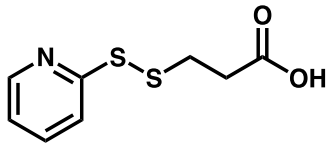
Compound **4**. Compound **3** (290 mg, 0.2 mmol) was dissolved in CH2Cl2 (5 mL). TFA (0.8 mL) and Et3SiH (0.8 mL) were added and the solution was stirred for 2 h at room temperature. The reaction mixture was concentrated under reduced pressure and the crude product was purified by silica column chromatography (CH2Cl2/MeOH, 9/0.6, v/v) to obtain compound 4 as viscous colorless oil (180 mg, 74%). 1H NMR (400 MHz, CDCl3) δ 7.56 (d, J = 7.8 Hz, 1H), 7.37 (t, J = 7.7 Hz, 1H), 7.02 (d, J = 8.3 Hz, 1H), 6.83 (t, J = 7.4 Hz, 1H), 3.66 – 3.60 (m, 84H), 3.53 (dd, J = 5.8, 3.4 Hz, 2H), 3.36 (s, 3H), 3.17 (t, J = 7.3 Hz, 2H), 2.53 – 2.47 (m, 2H), 1.71 – 1.65 (m, 2H),

1.60 – 1.53 (m, 2H), 1.43 – 1.28 (m, 9H), 1.24 (d, J = 6.7 Hz, 3H). 13C NMR (101 MHz, CDCl3) δ 169.3, 144.7, 136.1, 133.1, 128.0, 100.5, 77.5, 77.4, 77.2, 76.8, 71.9, 70.5, 70.4, 70.2, 69.8, 59.1, 39.6, 34.1, 29.3, 29.0, 28.4, 27.0, 24.7, 23.6. HRMS (ESI) m/z calculated for n=17: [C50H95N2O18S]+: 1043.6295. Found: 1043.6321



FEDS

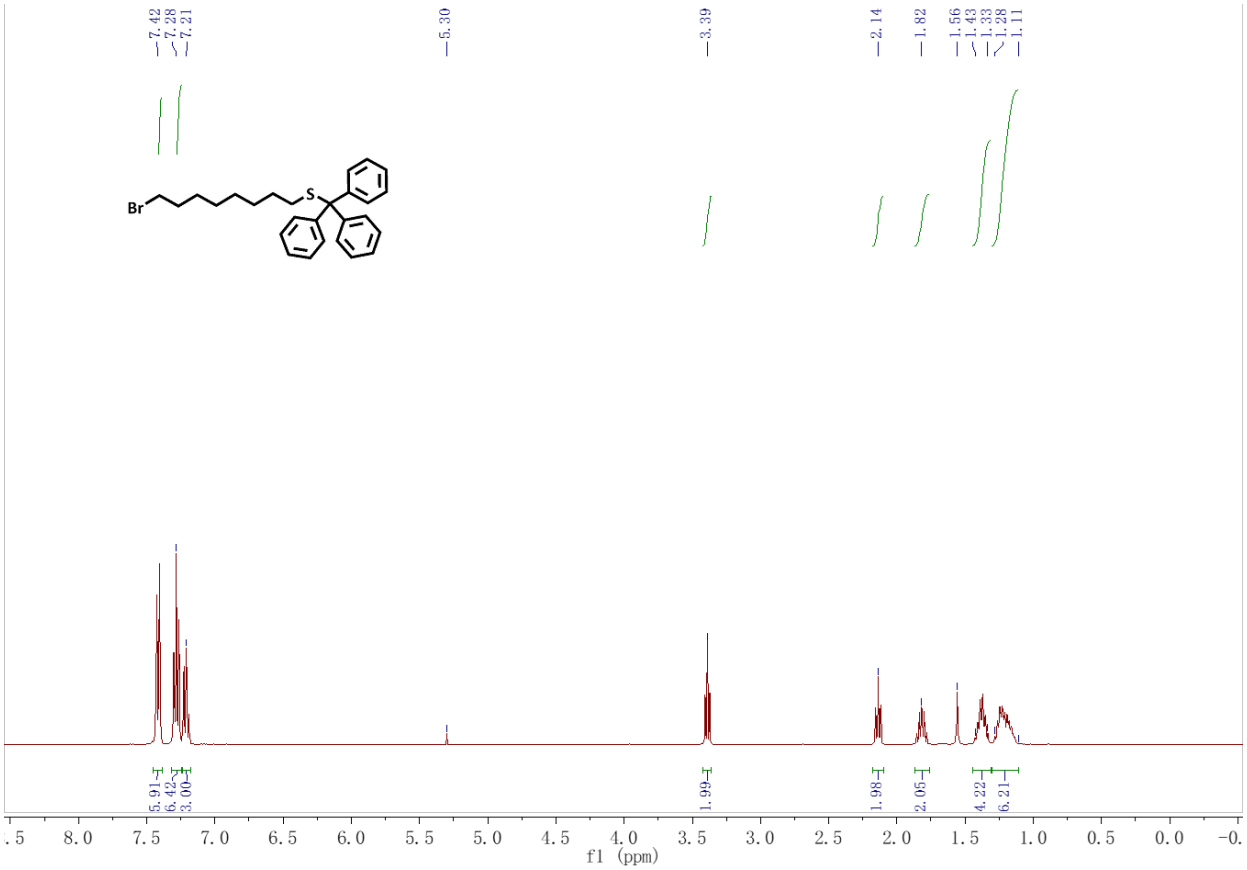
3-((8-((2-((2-(2methoxyethoxy)ethyl)carbamoyl)phenyl)amino)octyl)disulfaneyl)propanoic acid (FEDS). Compound 4 (200.9 mg, 0.16 mmol, 1eq) was added into the solution of 3-(pyridin-2-yldisulfaneyl)propanoic acid (70.6 mg, 0.33mmol, 2.1 eq) and AcOH (317 μL, 0.33g, 5.5mmol, 34 eq) in CH2Cl2 (8 mL) and stirred overnight. After the co-evaporation with toluene under reduced pressure to remoce acetic acid, the crude product was purified by silica column chromatography (DCM/EtOAc 9/1 to DCM/MeOH, 9/1, to DCM/MeOH, 8/2, v/v) to obtain FEDS as a viscous colorless oil (194.7mg, 0.14mmol, 87% yield). 1H NMR (300 MHz, CD3OD) δ 7.49 (d, J = 7.9 Hz, 1H), 7.28 (t, J = 9.6, 1H), 6.71 (d, J = 8.4 Hz, 1H), 6.58 (t, J = 7.5 Hz, 1H), 3.66 – 3.51 (m, 86H), 3.14 (t, J = 6.7 Hz, 2H), 2.92 (s, 2H), 2.74 – 2.69 (m, 2H), 1.71 – 1.64 (m, 3H), 1.45 – 1.29 (m, 9H). 13C NMR (101 MHz, CDCl3) δ 170.2, 160.8, 149.9, 149.7, 137.2, 132.8, 128.1, 127.9, 120.6, 119.7, 115.0, 114.5, 111.5, 71.8, 70.2, 59.2, 53.6, 43.1, 39.5, 39.2, 29.8, 29.4, 29.3, 29.0, 28.6, 28.5, 28.2, 27.3, 27.1, 23.6, 14.2. HRMS (ESI) m/z calculated for: n=17, [C53H97N2O20S2]-: 1145.6082. Found: 1145.6067

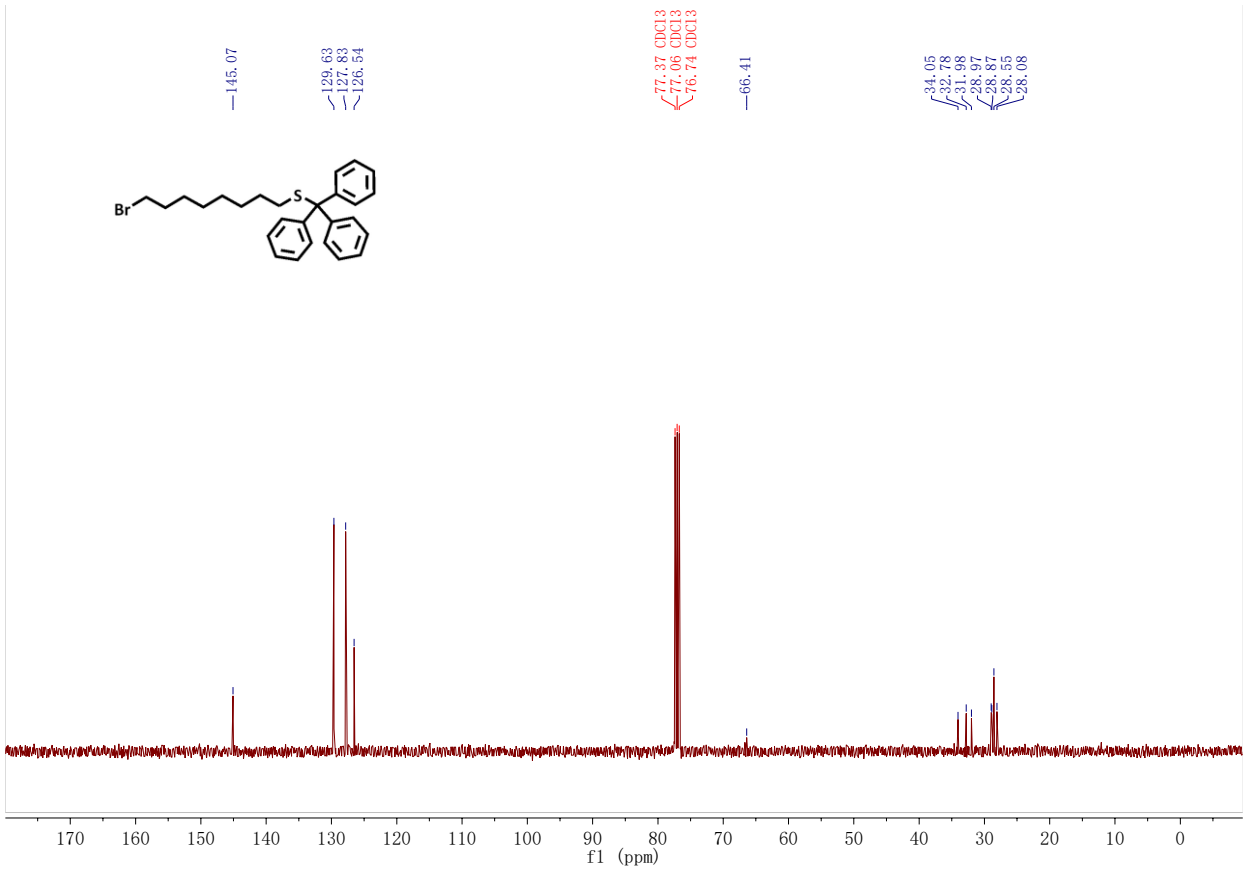


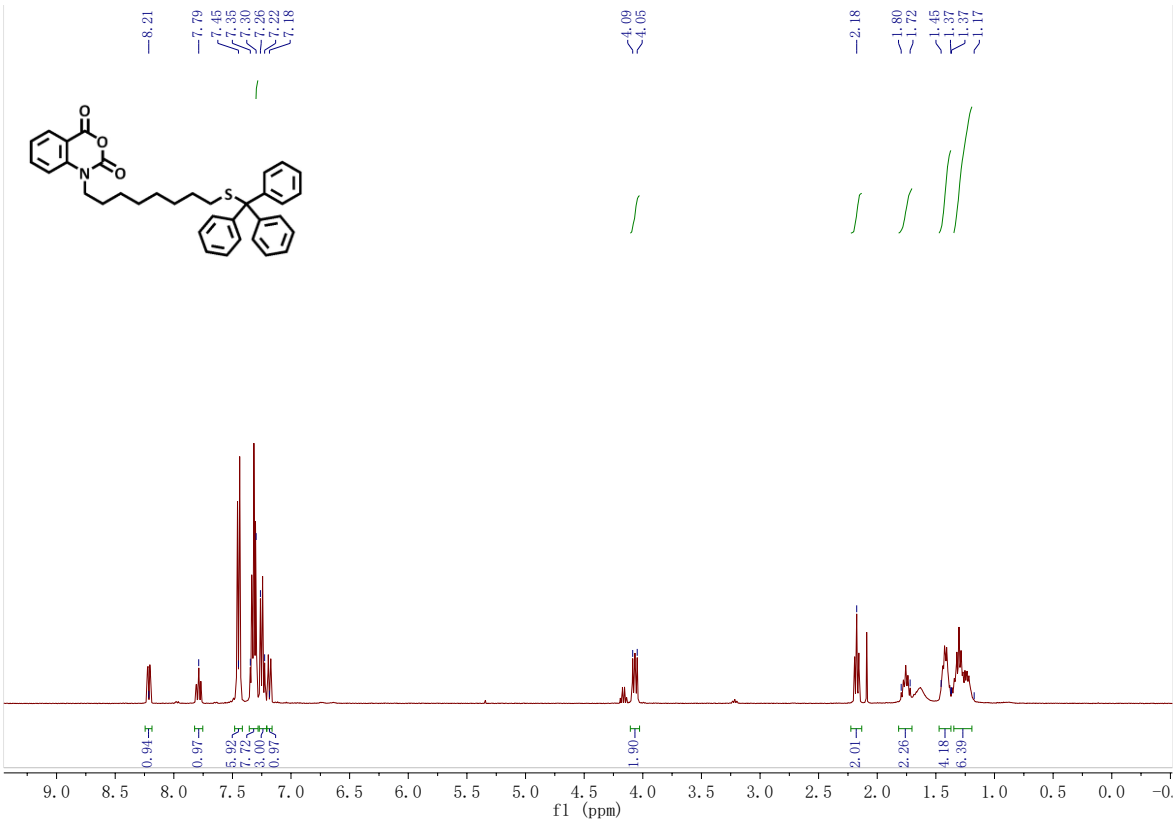
5

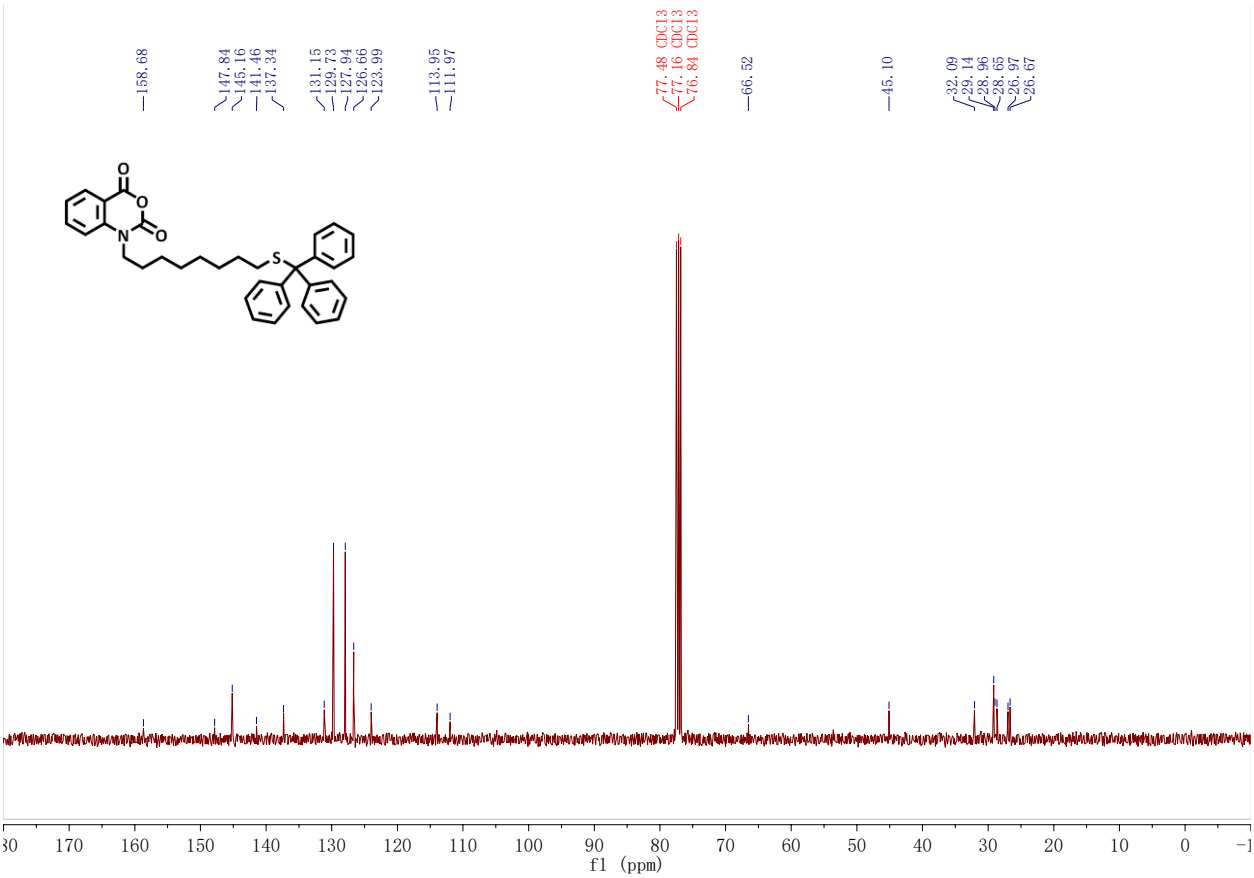
3-(pyridin-2-yldisulfaneyl)propanoic acid (5). 3-mercaptopropanoic acid (1.5 g, 14 mmol) in 10 mL EtOH was added dropwise into a solution of 1,2-di(pyridin-2-yl)disulfane (6.2 g, 28 mmol) and AcOH (0.8 mL) in EtOH (15 mL) and stirred overnight at room temperature. The reaction mixture was concentrated under reduced pressure and purified by silica column chromatography (hexane/EtOAc/AcOH, 50/50/0.5, v/v/v) to obtain 5 as a white powder (1.6 g, 53%). 1H NMR (300 MHz, CDCl3) δ 11.60 (s, 1H), 8.47 (d, J = 4.8 Hz, 1H), 7.96 – 7.48 (m, 2H), 7.16 – 7.10 (m, 1H), 3.03 (t, J = 6.9 Hz, 2H), 2.78 (t, J = 6.9 Hz, 2H). HRMS (ESI) m/z calculated for: [C8H10NO2S2]+: 216.0147. Found: 216.0144

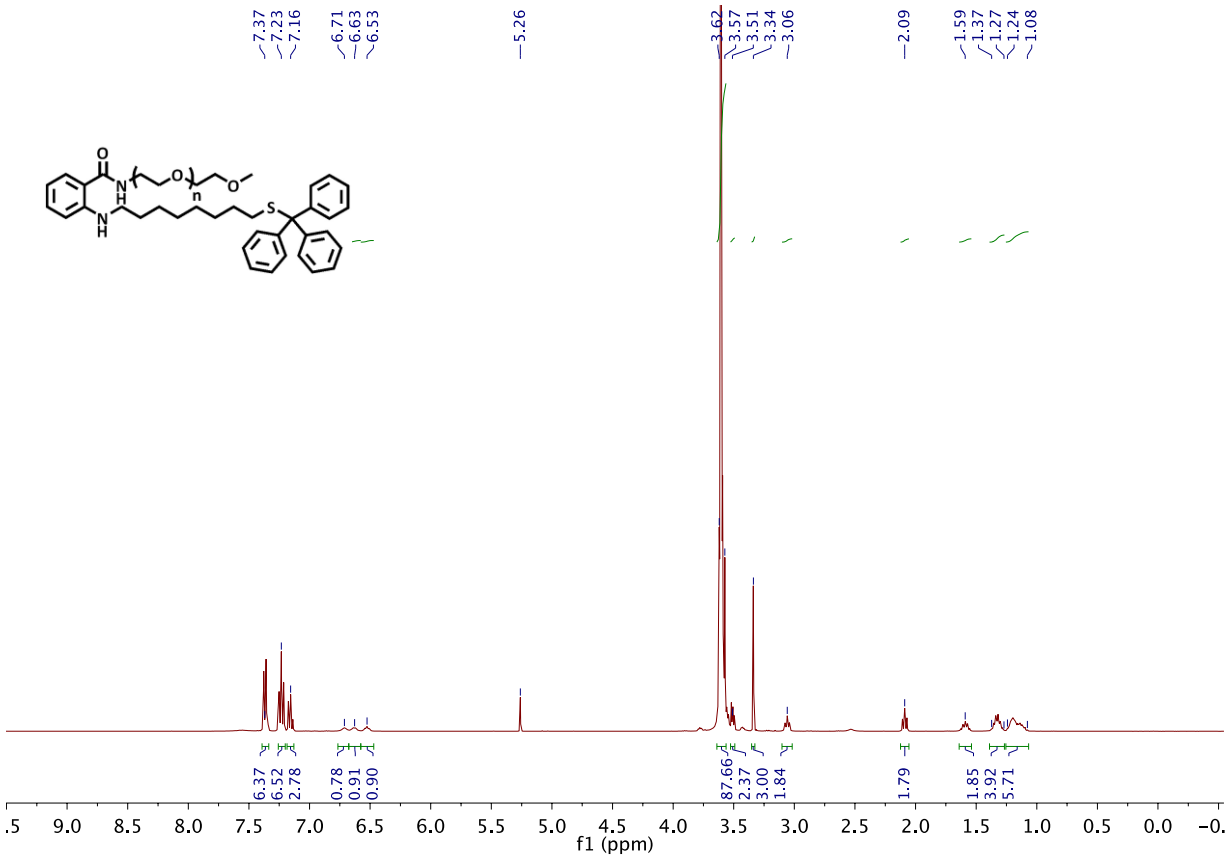
B1.4 NMR Spectra

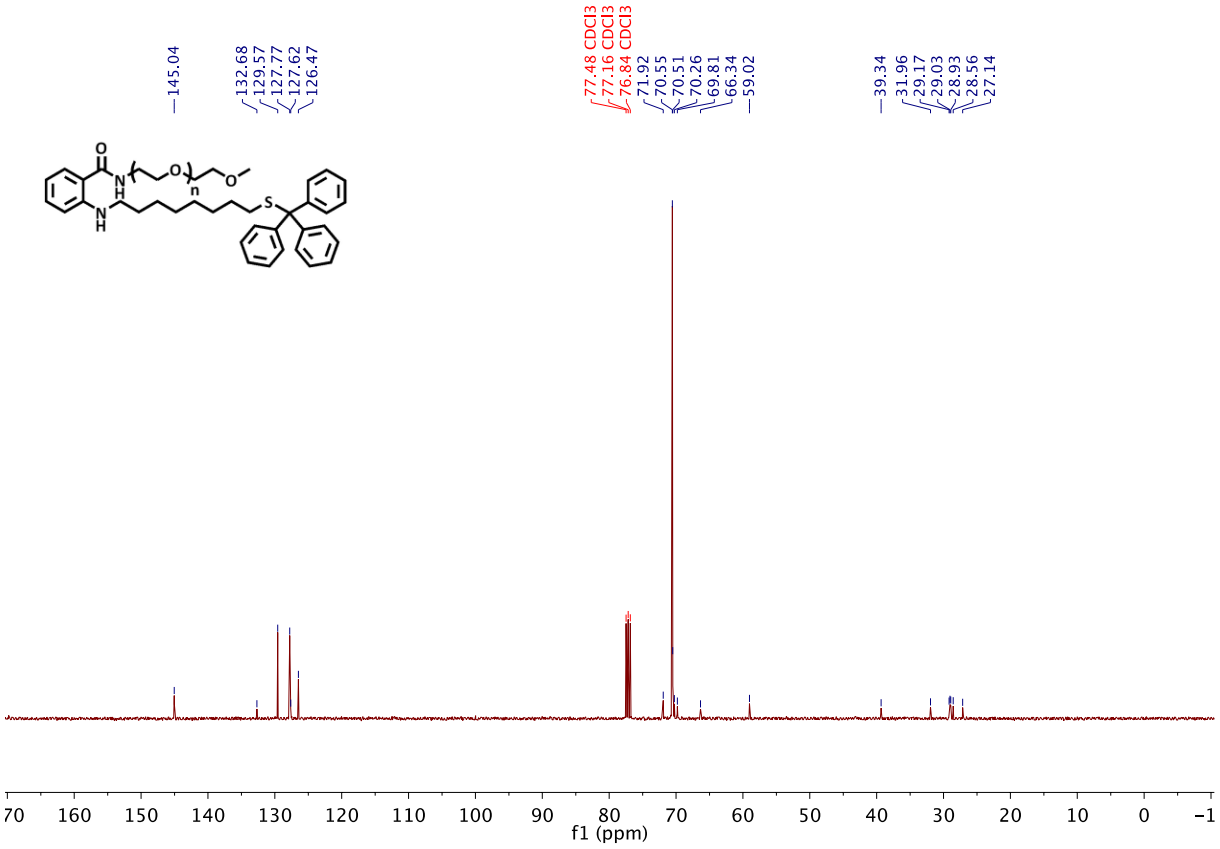


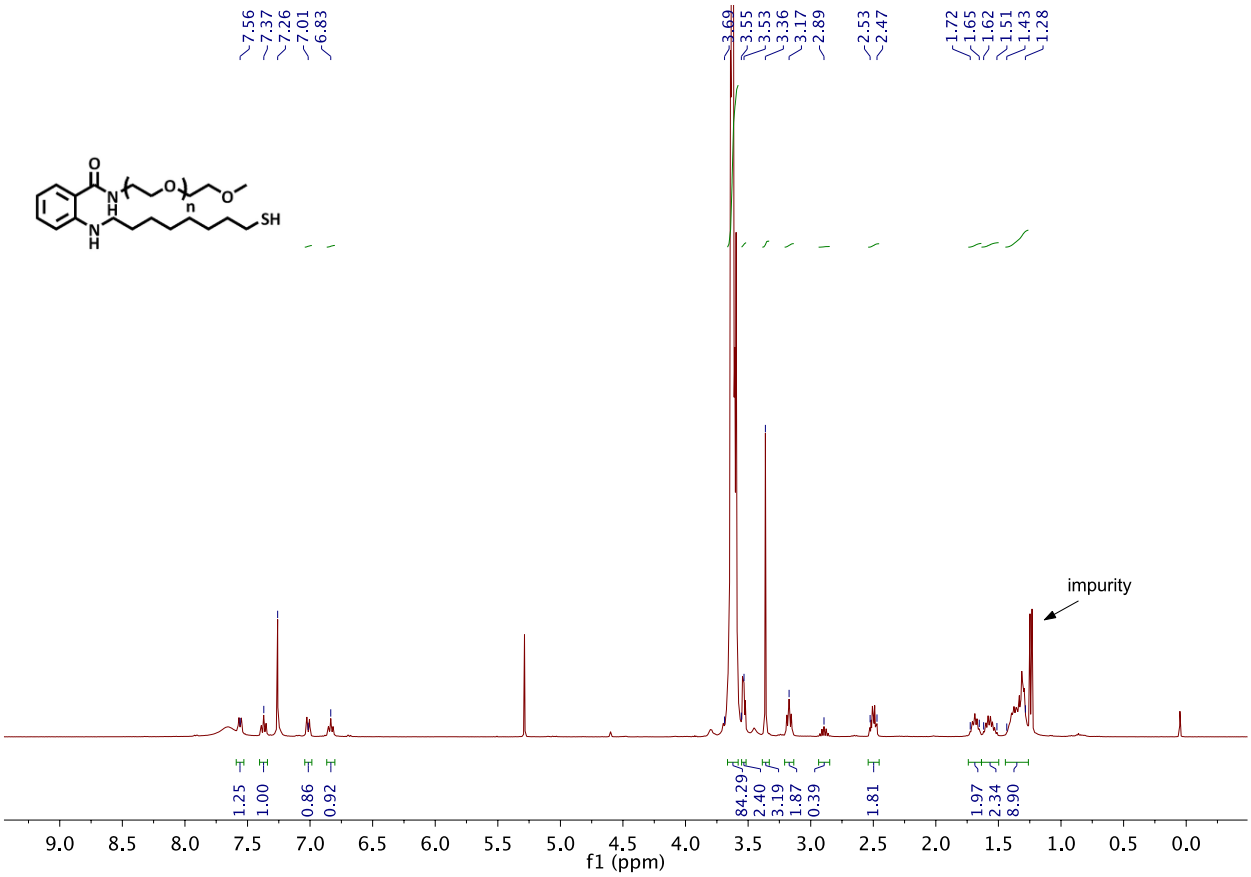

145.07
129.63
127.83
126.54
77.37 CDCl3
77.06 CDCl3
76.74 CDCl3
66.41
34.05
32.78
31.98
28.97
28.87
28.55
28.08
Br
S
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

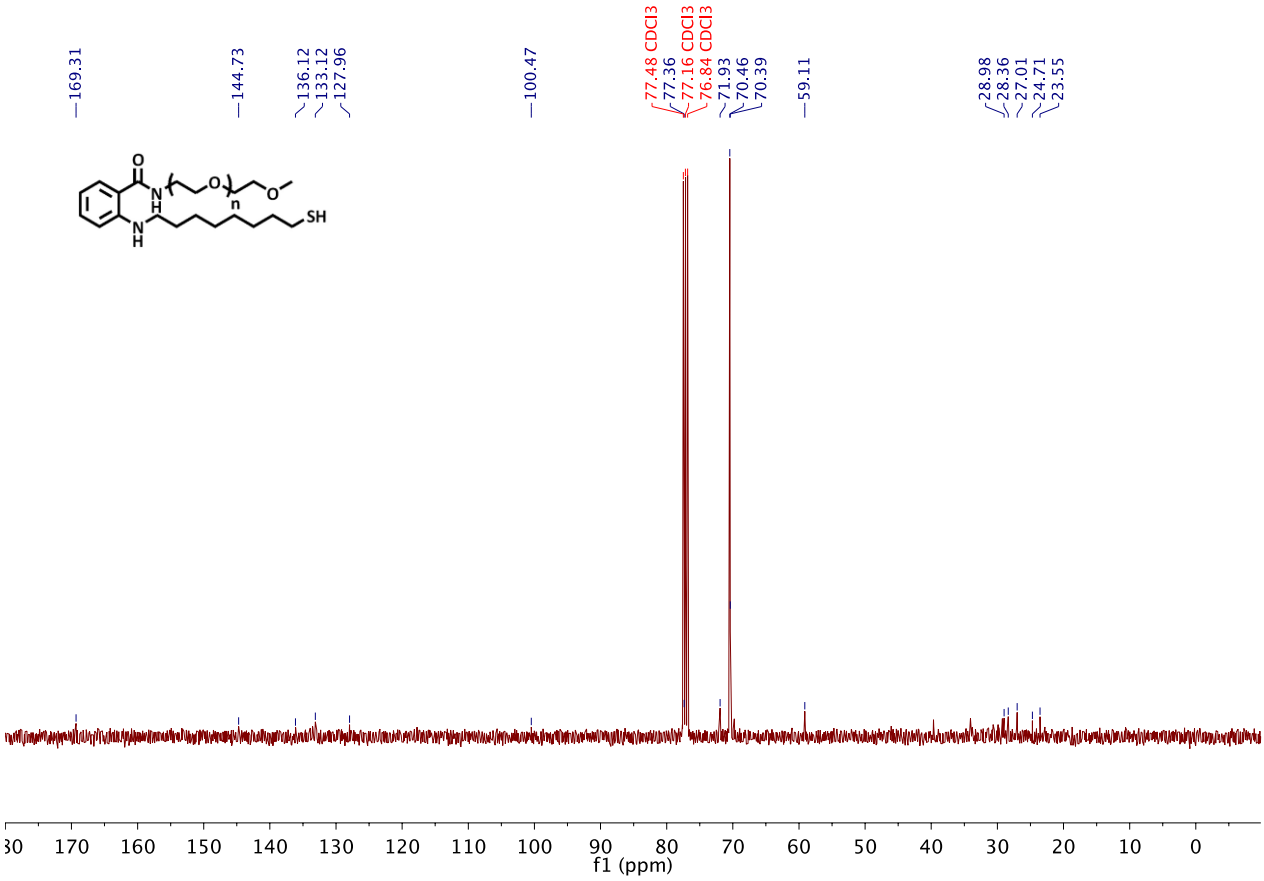

8.21
7.79
7.45
7.35
7.30
7.26
7.22
7.18
4.09
4.05
2.18
1.80
1.72
1.45
1.37
1.37
1.17
O
O
N
O
S
0.94
0.97
5.92
7.22
3.00
0.97
1.90
2.01
2.26
4.18
6.39
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

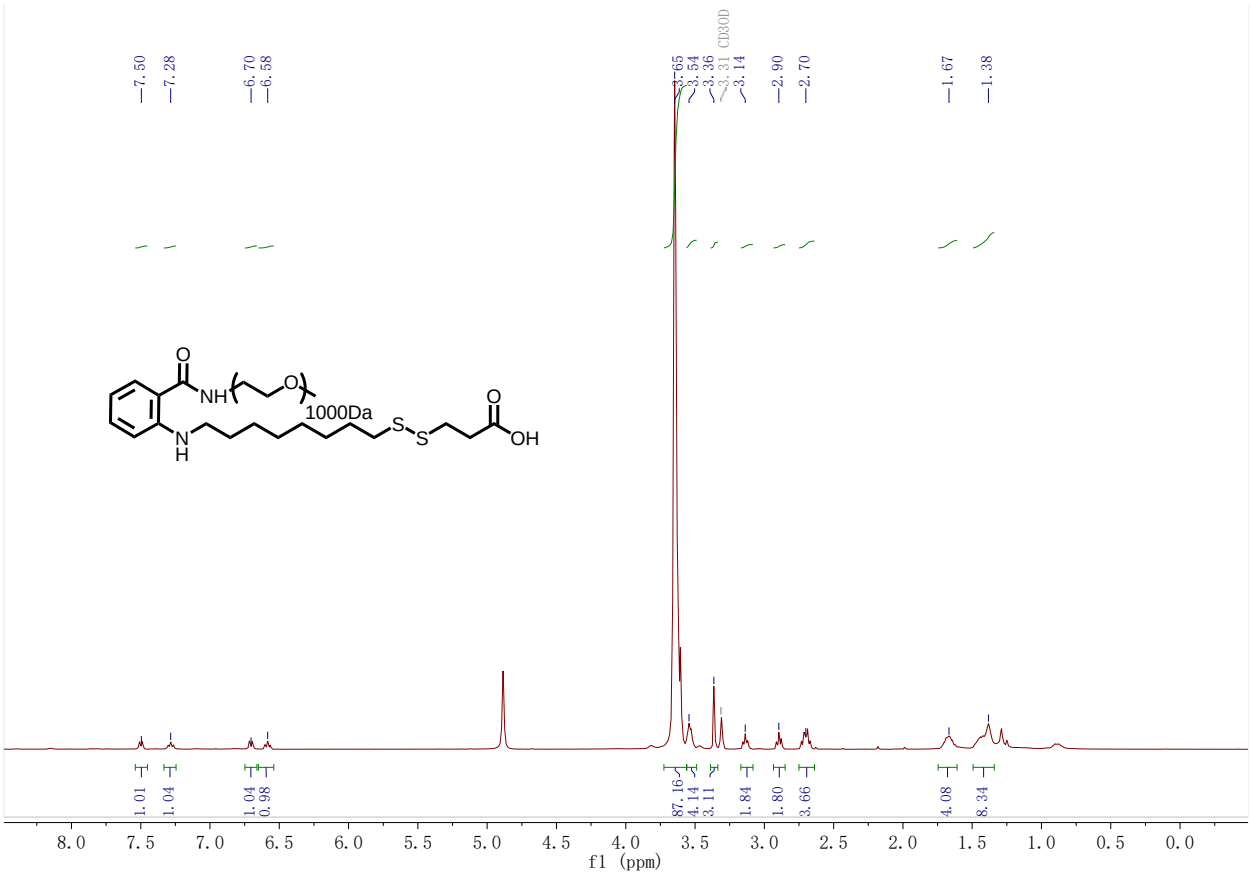

158.68
147.84
145.16
141.46
137.34
131.15
129.73
127.94
126.66
123.99
113.95
111.97
77.48 CDCl3
77.16 CDCl3
76.84 CDCl3
66.52
45.10
32.09
29.14
28.96
28.65
26.97
26.67
O
O
O
N
S
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

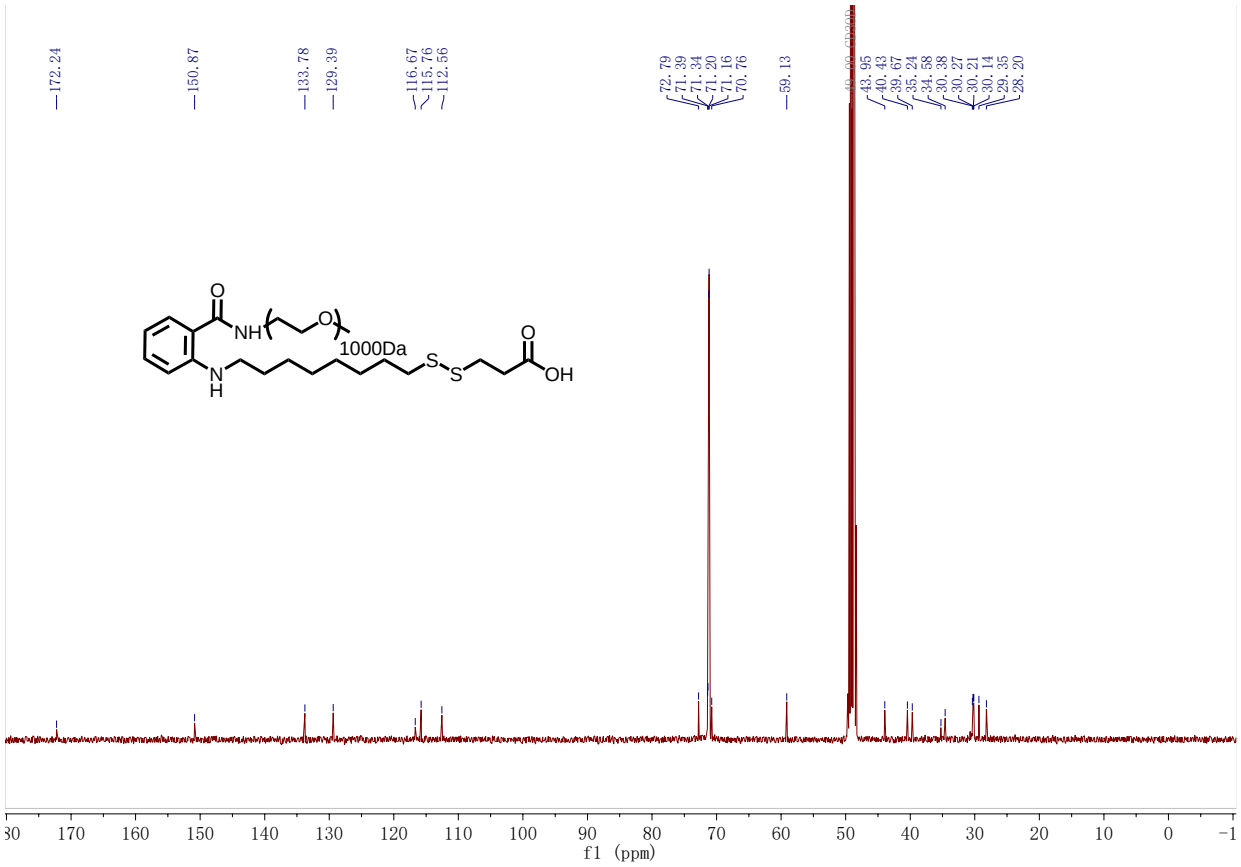

7.37
7.23
7.16
6.71
6.63
6.53
5.26
3.62
3.57
3.51
3.34
3.06
2.09
1.59
1.37
1.27
1.24
1.08
6.37
6.52
2.78
0.78
0.91
0.90
87.66
2.37
3.00
1.84
1.79
1.85
3.92
5.71
f1 (ppm)

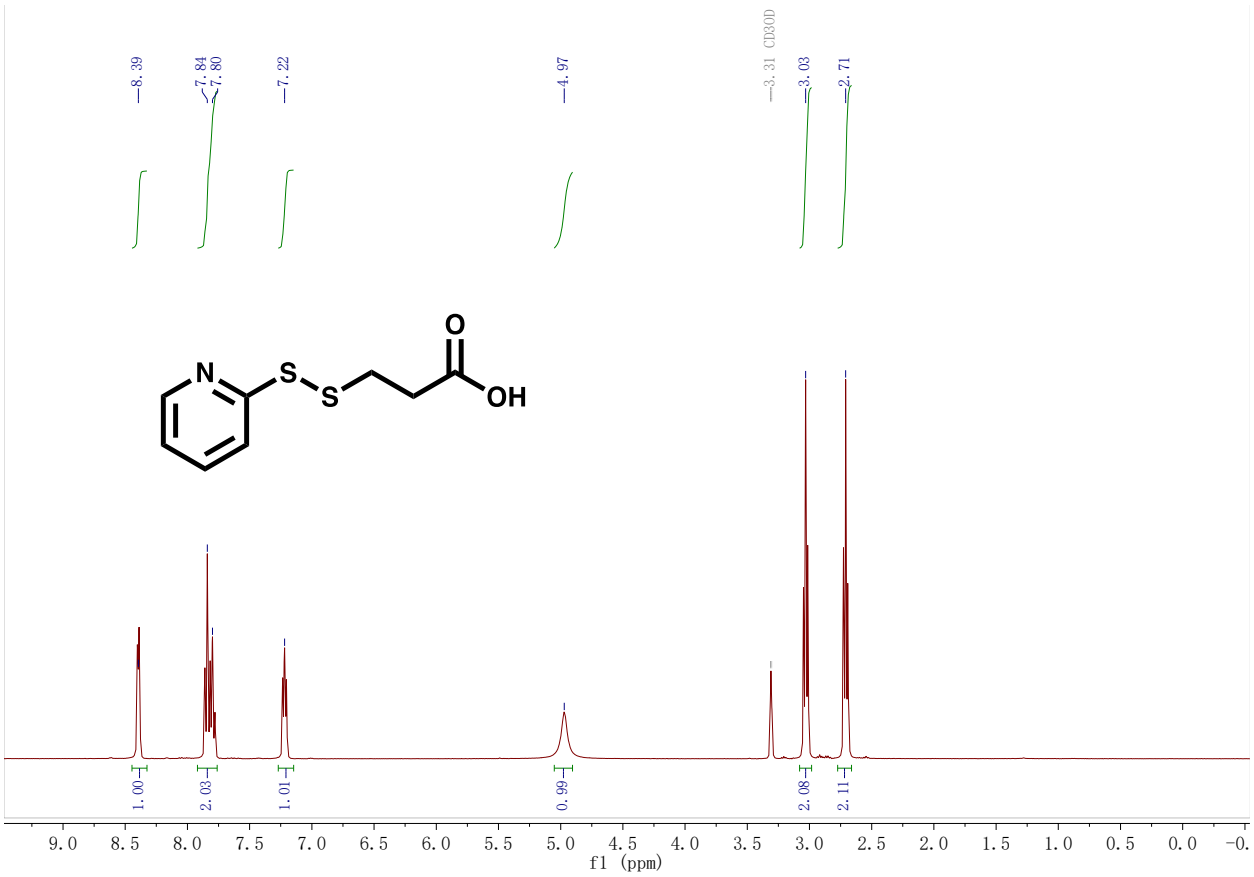

145.04
132.68
129.57
127.77
127.62
126.47
77.48 CDCl3
77.16 CDCl3
76.84 CDCl3
71.92
70.55
70.51
70.26
69.81
66.34
59.02
39.34
31.96
29.17
29.03
28.93
28.56
27.14
O
N
H
N
H
O
n
O
S
70
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
−1
f1 (ppm)

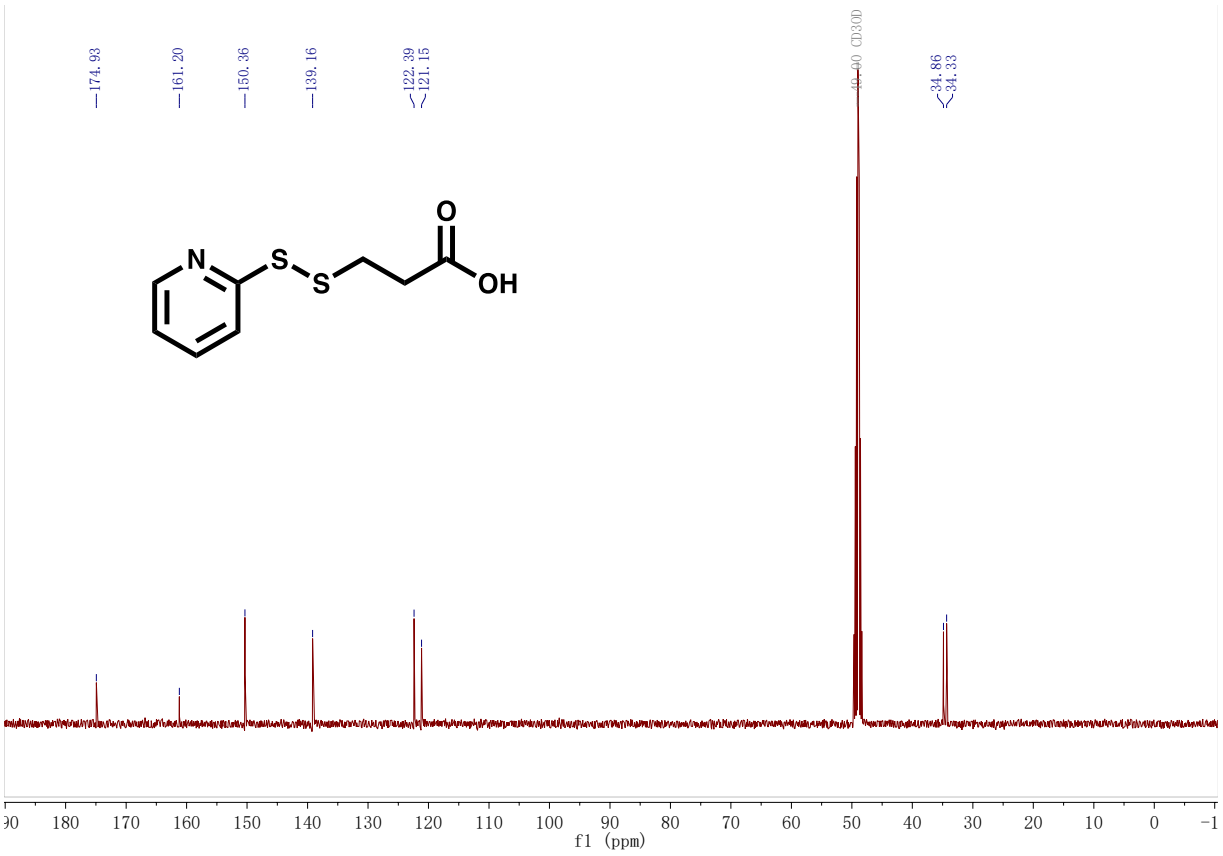

7.56
7.37
7.26
7.01
6.83
3.69
3.55
3.53
3.36
3.17
2.89
2.53
2.47
1.72
1.65
1.62
1.51
1.43
1.28
impurity
1.25
1.00
0.86
0.92
84.29
2.40
3.19
1.87
0.39
1.81
1.97
2.34
8.90
f1 (ppm)


169.31
144.73
136.12
133.12
127.96
100.47
77.48 CDCl3
77.36
77.16 CDCl3
76.84 CDCl3
71.93
70.46
70.39
59.11
28.98
28.36
27.01
24.71
23.55
O
N
H
N
H
O
n
O
SH
f1 (ppm)


7.50
7.28
6.70
6.58
3.65
3.54
3.36
3.31 CD3OD
3.14
2.90
2.70
1.67
1.38
O
NH
O
1000Da
N
H
S
S
O
OH
1.01
1.04
1.04
0.98
87.16
4.14
3.11
1.84
1.80
3.66
4.08
8.34
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


172.24
150.87
133.78
129.39
116.67
115.76
112.56
72.79
71.39
71.34
71.20
71.16
70.76
59.13
43.95
40.43
39.67
35.24
34.58
30.38
30.27
30.21
30.14
29.35
28.20
O
NH
O
1000Da
N
H
S
S
O
OH
f1 (ppm)


8.39
7.84
7.80
7.22
4.97
3.31 CD3OD
3.03
2.71
N
S
S
O
OH
1.00
2.03
1.01
0.99
2.08
2.11
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.
f1 (ppm)


174.93
161.20
150.36
139.16
122.39
121.15
49.00 CD3OD
34.86
34.33
N
S
S
O
OH
f1 (ppm)

Appendix B2 – Chapter 2.3

B2.1 Characterization and Cell Work Protocols

Preparation of protein formulations with PEG-pEosin

PEG-pEosin and proteins were diluted separately in equal volumes of phosphate-buffered saline (PBS) and mixed at indicated ratios by rapid pipetting. The formulations were incubated for 5 min at room temperature.

Native polyacrylamide gel electrophoresis

Bovine serum albumin (BSA) formulations with PEG-pEosin were prepared at indicated ratios. In the case of binding studies, the formulations were loaded immediately into the gels. In the case of release studies, the formulations were incubated for 4 hours at room temperature in PBS at the indicated pH before loading into the gels. BSA formulations contained 1.5 μg protein in a 15 μL volume and were pipetted into the well after mixing with 5 μL of loading buffer (prepared from 4 mL of glycerol, 6 mL running buffer and 0.02 g of bromophenol blue). Native 4–20% Mini-PROTEAN® TGX™ Precast Protein Gels (Biorad) were used. The running buffer contained 3.0 g of Tris base and 14.4 g of glycine in 1000 mL of H2O. The pH of the buffer was 8.3. Native PAGE was performed at 70 V until samples were stacked and subsequently at 140 V. Gels were stained with Bio-Safe™ Coomassie Stain (Biorad).

Particle size and zeta potential

For dynamic light scattering (DLS) measurements the protein formulations were measured in a folded capillary cell (DTS 1070) using a Zetasizer Nano ZS with backscatter detection (Malvern Instruments). Proteins were mixed with a 6-fold molar excess of PEG-pEosin in a total volume of 50 μL containing 10 μg protein. Samples were measured in triplicates. For zeta potential measurements, the samples were diluted with 750 μL PBS. The influence of acid hydrolysis on the particle distribution was studied by measuring DLS after 10 min incubation in PBS at pH 5.5.

Hemolysis assay

Fresh rabbit blood (900 μL) was washed 3 times with PBS. The erythrocyte suspension was centrifuged at 1500 RCF for 0.5 min and the pellet was resuspended in 2 mL PBS. The suspension was diluted 1:20 in PBS at pH 7.4, 6.8 and 5.5. Volumes of 400 μL of erythrocyte suspension and 400 μL of LLO solutions (with or without PEG-pEosin, previously formulated with PBS at the respective pH to achieve a 0.5 μg/mL final concentration of LLO) were incubated for 2 hours at 37°C. At indicated time points, aliquots of 100 μL were centrifuged and 60 μL of the supernatant was analyzed for hemoglobin release at 405 nm wavelength using a TECAN infinity 200 microplate reader. Samples were analyzed in triplicates. PBS at the respective pH was used as a

negative control (0% lysis) and 1% solutions of Triton-X in PBS was used as a positive control (100% lysis).

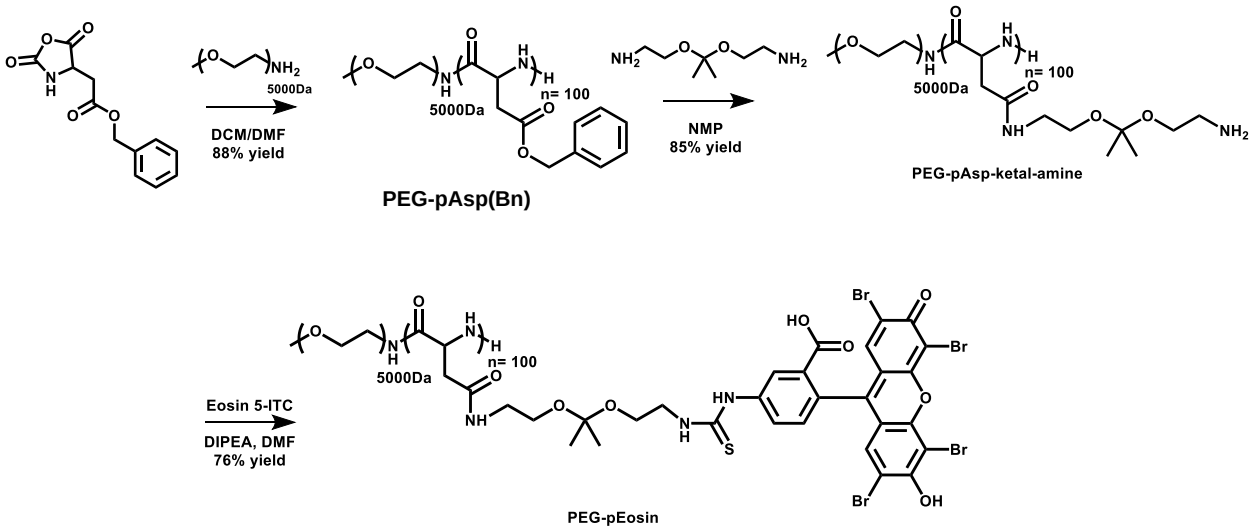
Cell viability assay

Cell viability of HeLa cells was analyzed after incubation with free and PEG-pEosin-formulated LLO at 2.5 μg/mL concentration at pH 7.4 and 6.8 for 4 hours in serum-free media and 24 hours at different concentrations in serum-containing media. Cell viability of 3T3 Ai9 cells was analyzed after 24 hours incubation with Cre recombinase co-formulations with 8 and 12 μg/mL concentrations of LLO and 6 molar equivalents of PEG-pEosin in comparison to controls with Cre only, PEG-pEosin and Cre + free LLO in serum-containing media. The cells were washed with PBS and incubated with 100 μL fresh medium containing resazurin (10 μL 0.1mg/mL in PBS) for 3 h. The cell viability was then determined by measuring the fluorescence intensity of each well (Ex = 550 nm, Em = 595 nm), using a TECAN infinity 200 microplate reader. Untreated cells and fresh medium were used as controls for 100% and 0% cell proliferation, respectively.

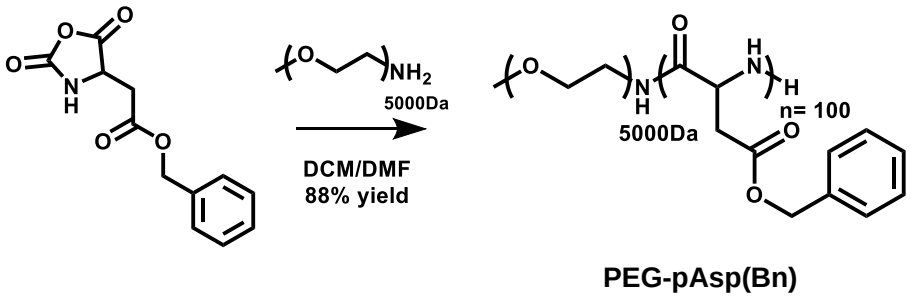
Cre recombinase assay

Gene editing efficiency of Cre recombinase co-formulations was analyzed in 3T3 Ai9 cells, using 8 μg/mL (143 nM) and 12 μg/mL (214 nM) concentrations of LLO, 3.8 μg/mL (100 nM) Cre and 6 molar equivalents of PEG-pEosin (1.46 μM and 1.88 μM respectively). The controls were composed of Cre only, PEG-pEosin, and Cre + free LLO. The formulations were incubated with 3T3 Ai9 cells for 48 hrs at 37 °C in a humidified incubator containing 5% CO2. The 3T3 Ai9 cells were detached with Accutase (Invitrogen) and the TdTomato signal was analysed with a Attune NxT flow cytometer (Invitrogen). The gene editing efficiency was determined by identifying the percentage of TdTomato positive cells.

B2.2 Synthetic Protocols

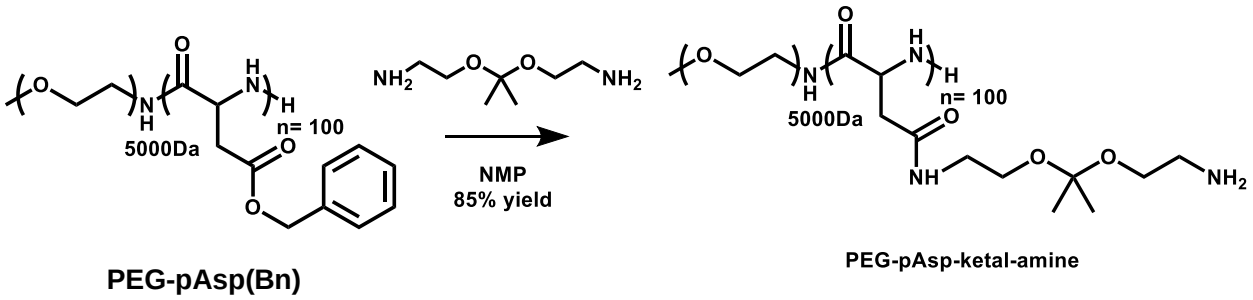


Scheme 14: Scheme for the synthesis of PEG-pEosin.



To a solution of benzyl 2-(2,5-dioxooxazolidin-4-yl)acetate (1.0 g, 4 mmol, 100eq) in 10:1 of DCM:DMF (4.5mL) was added a solution of MeO-PEG5000-NH2 (0.2 g, 0.04 mmol, 1eq) in 10:1 of DCM:DMF (7mL). The solution was heated to 40oC and allowed to stir for 48 hours. The mixture was then cooled to room temperature and added to cold ethyl ether (125 mL). The resulting precipitate was filtered and washed with additional ethyl ether (100mL). This yielded the product PEG-pAsp(Bz) as a white solid (898.7mg, 88% yield). NMR analysis indicated that the degree of polymerization was approximately 100.

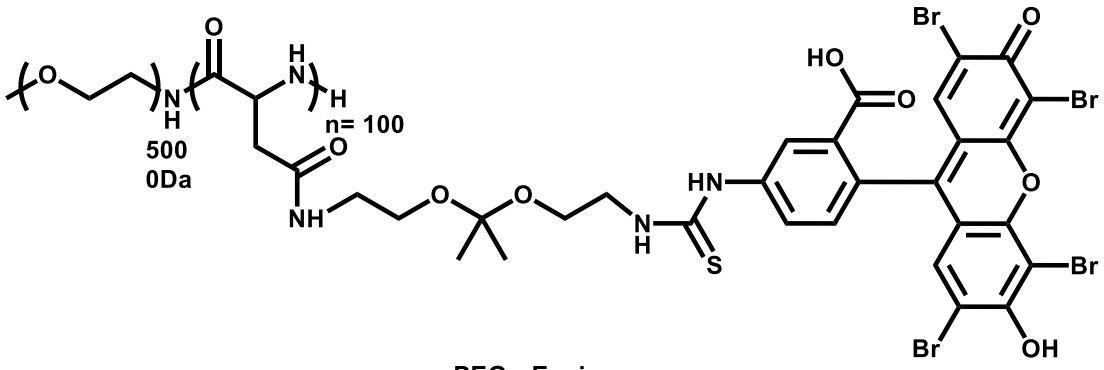
^{1}H NMR (400 MHz, DMSO) δ 8.27 – 8.06 (m, 87H), 7.36 – 7.21 (m, 502H), 5.10 – 4.90 (m, 196H), 4.71 – 4.50 (m, 98H), 3.57 – 3.40 (m, 455H), 2.92 – 2.71 (m, 104H), 2.61 (s, 95H).



PEG-pAsp(Bn) **PEG-pAsp-ketal-amine**

To 2,2-Bis(aminoethoxy)propane (827 mg, 5.1 mmol, 3000eq) was added dropwise over 10 mins a solution of PEGpAsp(Bn) (29.4 mg, 1.7 umol, 1eq) in NMP (1 mL). The solution was allowed to stir at room temperature for 16 hours. The mixture was added to cold ethyl ether (5 mL). The resulting precipitate was filtered and washed 3 times with cold ethyl ether (5 mL). This yielded the product PEG-pAsp-ketal-amine as a white solid (29.7 mg, 85% yield).

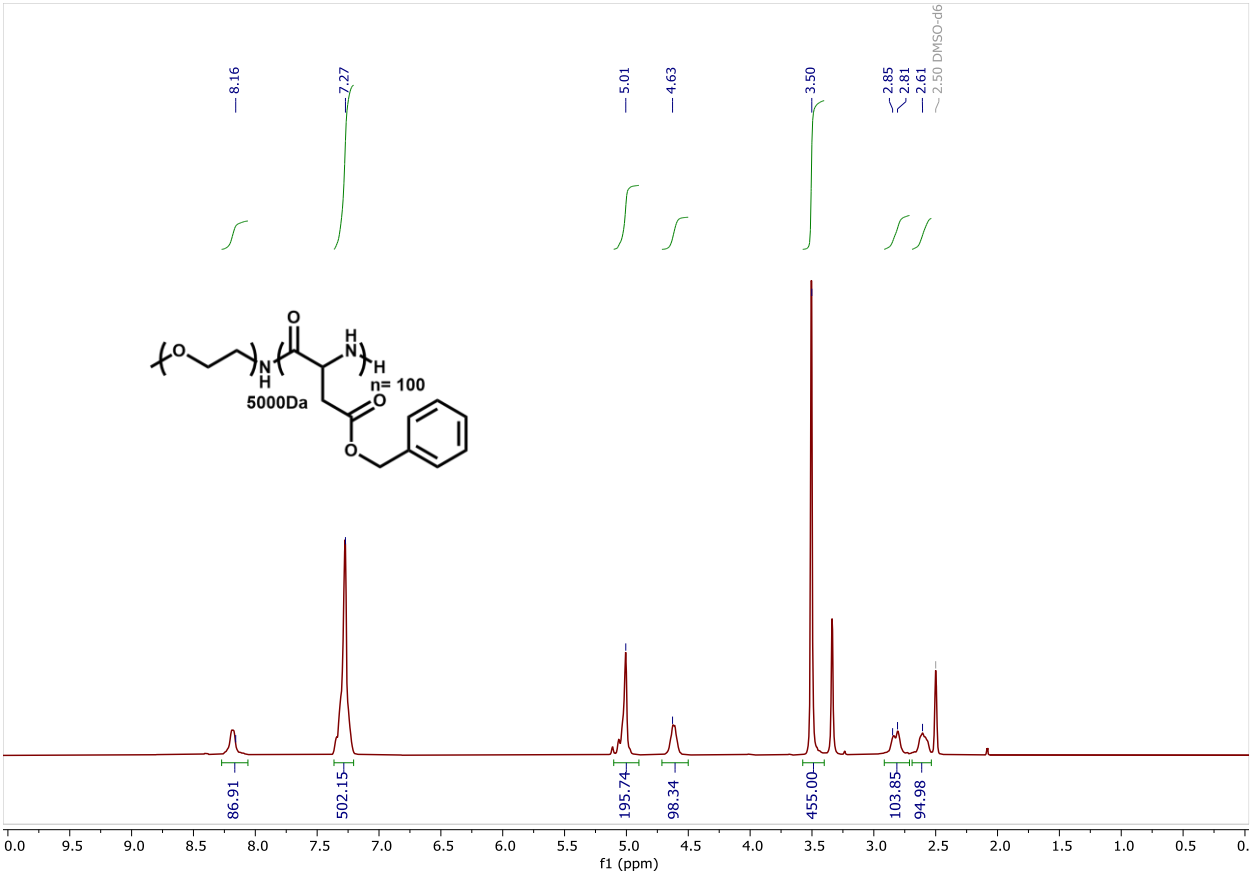
^{1}H NMR (400 MHz, DMSO) δ 4.62 – 4.35 (m, 95H), 3.54 – 2.94 (m, 1110H), 2.66 (s, 255H), 1.22 (s, 574H).

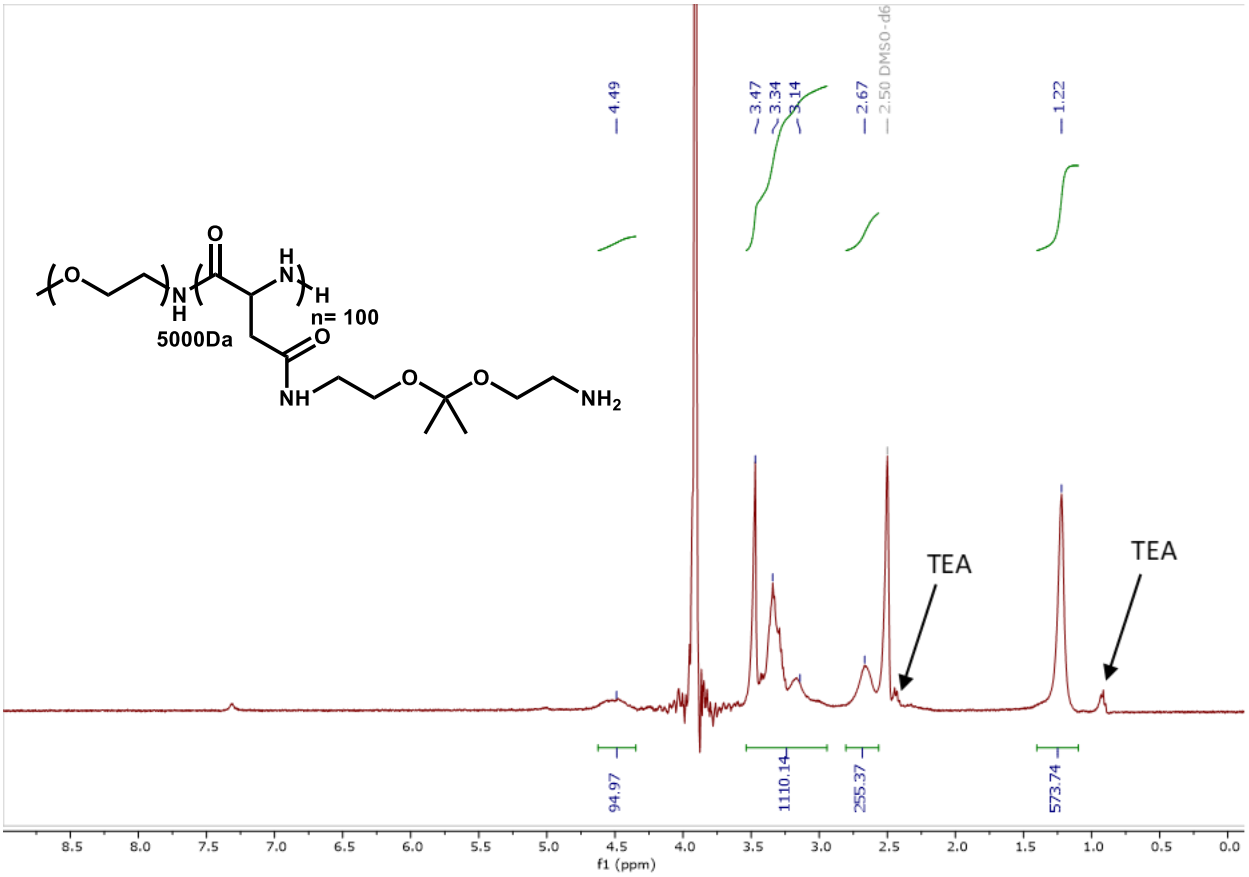


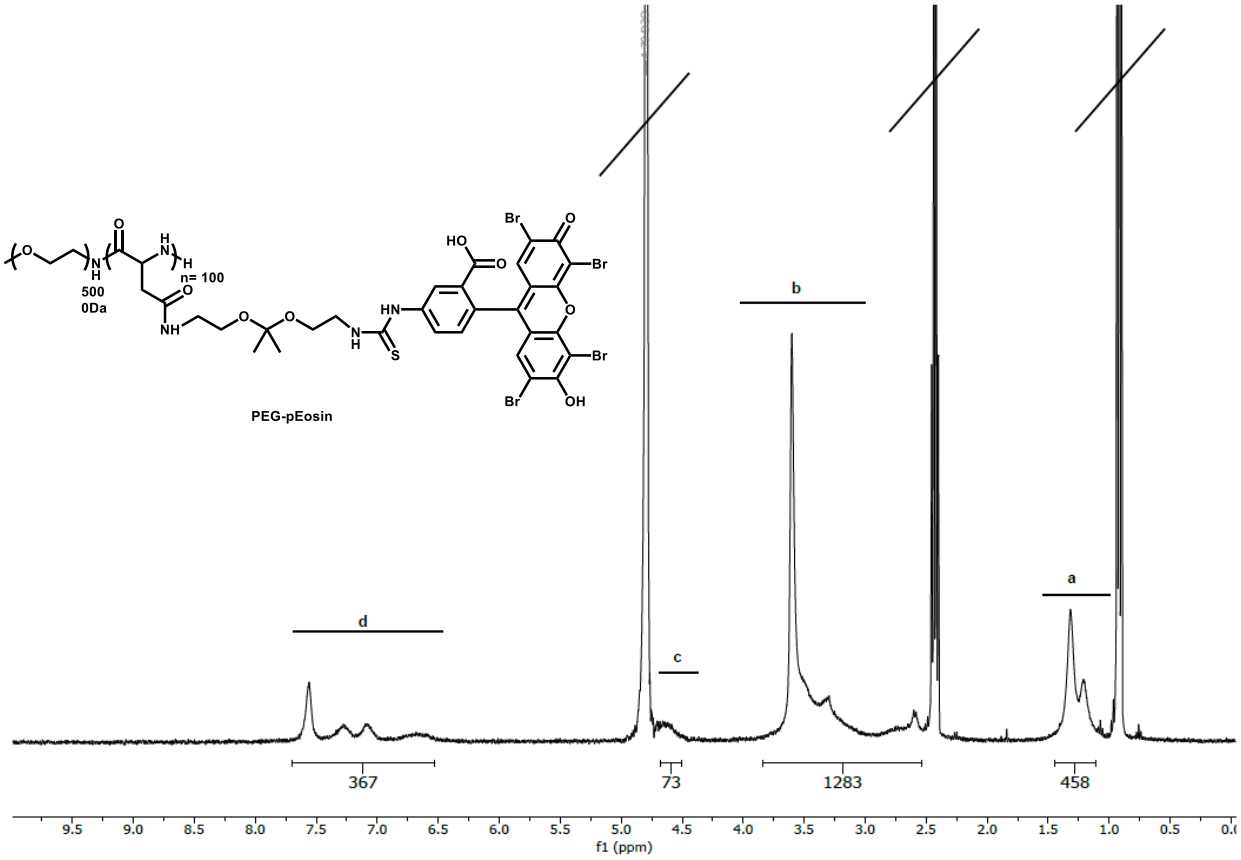
PEG-pEosin

PEG-pAsp(ketal) (10 mg, 0.476 μmol) was dissolved in 0.5 mL of anhydrous DMF. To this solution were added DIPEA (9.23 mg, 13.2 μL, 71.4 μmol) and Eosin 5-isothiocyanate (49.8 mg, 70.6 μmol, dissolved in 1 mL anhyd. DMF), both in 2.5-fold molar excess per free amine of PEG-pAsp(ketal). The reaction mixture was stirred at room temperature for 120 min. The reaction solution was added dropwise to 40 mL of cold n-hexane/ether 50/50. The red precipitate was collected and then redissolved in 0.2 mL of DMF, and precipitated with cold n-hexane/ether 50/50 3 times, and dried under vacuum. The red solid was dissolved in 2 mL 0.1% EtN3 in water and further purified using a Zeba™ Spin Desalting Column, 7K MWCO, 10 mL to afford PEG-pEosin as a red powder after lyophilization. The yield was 22.6 mg, 76%. PEG-pEosin was readily soluble in water at a concentration of 10 mg/mL (161 μM).

B2.3 NMR Spectra




4.49
3.47
3.34
3.14
2.67
2.50 DMSO-d6
1.22
n= 100
5000Da
NH2
TEA
TEA
94.97
1110.14
255.37
573.74
f1 (ppm)


n= 100
500
0Da
PEG-pEosin
d
c
b
a
367
73
1283
458
f1 (ppm)

Chapter 3 – Delivery of Fluorescent Markers to Bacteria Using Maltohexaose Sugars

Infectious endocarditis is a life-threatening disease, and diagnostics are urgently needed to accurately diagnose this disease especially in the case of prosthetic valve endocarditis. We show here that maltohexaose conjugated to indocyanine green (MH-ICG) can detect Staphylococcus aureus (S. aureus) infection in a rat model of infective endocarditis. The affinity of MH-ICG to S. aureus was determined and had a Km and Vmax of 5.4 μM and 3.0 X 10^{-6} μmol/minutes/10^{8} CFU, respectively. MH-ICG had no detectable toxicity to mammalian cells at concentrations as high as 100 μM. The *in vivo* efficiency of MH-ICG in rats was evaluated using a right heart endocarditis model, and the accumulation of MH-ICG in the bacterial vegetations was 2.5 ± 0.2 times higher than that in the control left ventricular wall. The biological half-life of MH-ICG in healthy rats was 14.0 ± 1.3 minutes, and approximately 50% of injected MH-ICG was excreted into the feces after 24 hours. These data demonstrate that MH-ICG was internalized by bacteria with high specificity and that MH-ICG specifically accumulated in bacterial vegetations in a rat model of endocarditis. These results demonstrate the potential efficacy of this agent in the detection of infective endocarditis.

Chapter 3 is based on the following published research study:

Takemiya K, **Røise JJ**, He M, Taing C, Rodriguez AG, Murthy N, Goodman MM & Taylor WR (2021) *Maltohexaose-indocyanine green (MH-ICG) for near infrared imaging of endocarditis*. PLoS ONE 16(3): e0247673

3.1 Introduction

Despite advances in the field of bacterial diagnostics and treatment, bacterial infections remain a major challenge in the medical field due to the increasing emergence of antibiotic resistance, ageing populations, and through the use of organ transplants and implants. In particular, implant infections can develop months to years after being incorporated, and can be challenging to diagnose[52]. Early diagnosis is important to give treatment before the formation of excessive biofilms, which complicates the treatment of infections.

The selective delivery of imaging agents to bacteria has been studied as an option for early bacterial detection. This study will focus on the maltose transport system, which is found specifically on bacterial membranes. Maltose molecules cross the outer bacterial membrane by the help of maltoporin (LamB) before it is recognized by the periplasmic maltose binding protein (MBP) and internalized by the MBP-dependent ATP binding cassette (ABS) transporter[53,54]. This maltose transportation system is found in a wide range of bacteria, but not present on mammalian membranes[55]. MBP recognizes maltose chains up to maltohexaose[53], with maltose and maltotrioses also being used for bacterial targeting[56–58]. The maltose transport system also tolerates modifications on the reducing end of the sugars. This opens up for a wide range of targeted imaging and therapeutic applications. In this chapter, the development of using maltohexaose-based imaging agents for bacterial imaging will be reviewed. Then, a maltohexaose-indocyanine green (MH-ICG) imaging agent will be introduced and assessed for in-vivo imaging of infective endocarditis (IE).

3.2 Maltohexaose-based Imaging

Maltohexaose-based delivery was shown in 2011 by Ning et al.[59], when they synthesized perylene- and IR-786-conjugates of maltohexaose. Since then, several PET-based and NIR imaging probes have been developed for *in vivo* imaging of bacterial infections. This section will cover current maltohexaose-based bacterial imaging probes and their applications (*see Table 1*).

Table 1: Summary of the current studies on maltohexaose-based bacterial imaging materials.

Imaging Moiety	***Application***	***Reference***
Perylene	*Initial studies that showed bacterial delivery specificity*	Ning et al.[59]
IR-786	*In vivo imaging of E. coli-induced muscular infection*	Ning et al.[59]
	NIR Imaging of cardiac device infections	Takemiya et al.[60]
Cy-7	*Fluorescence and photoacoustics-based in vivo imaging of E. coli and S. aureus wound infections*	Zlitni et al.[58]
^{18}F	*PET imaging of in vivo E. coli muscular infections*	Ning et al.[61]
	PET imaging of cardiac device infections	Takemiya et al.[60]
Indocyanine Green (ICG)	*NIR imaging of infective endocarditis*	*Takemiya et al.*[62] *(this study)*

Fluorescence-based Probes

The earliest maltohexaose-based imaging agents were those made by conjugation of fluorescent dyes to maltohexaose (Figure 47). In 2011, Ning et al.[59] synthesized perylene (MH-perylene) and IR-786 (MH-IR786) based maltohexaose probes. Perylene is a fluorescent dye with λ_{em} and λ_{ex} of 400nm and 445nm respectively, and MH-Perylene was subsequently used for *in vitro* characterization studies. In the same study, MH-IR786 was used for *in vivo* imaging of E. coli-induced muscular infections. IR-786 is a NIR dye, which due to its longer excitation and emission wavelengths (λ_{em} = 710nm/ λ_{ex} = 790nm) is more viable for *in vivo* imaging. Using MH-IR786, Ning et al.[59] were able to image down to 10^5 CFUs of *E. coli* following an intramuscular injection. Takemiya et al.[60] later used MH-IR786 for *in vivo* detection of implant infections in cardiac devices. Cardiac devices are prone to developing persistent infections that are difficult to diagnose. These infections can in some cases require removal of cardiac devices. Early detection using maltohexaose-based imaging agents has the potential to allow for early treatment of these infections. In 2020, Zlitni et al.[58] reported a Cy-7-conjugate with maltohexaose (MH-Cy7) that was used to detect *E. coli*-induced muscular infections in mice.

^{18}F PET-based Probes

Positron emission tomography (PET)-based imaging is a highly sensitive imaging technique based on the release of gamma-radiation from radioisotopes (typically ^{18}F), which is detected to create a 3D-image of accumulated ^{18}F-probe. When ^{18}F is incorporated into maltohexaose (MH-^{18}F, Figure 48), PET imaging can be used to detect the accumulation of ^{18}F in areas with bacterial growth. This can be used to image small amounts of bacteria to allow for early detection of infections, for example Ning et al.[61] used MH-^{18}F to image down to 10^5 CFUs using PET imaging. In addition, they also found a significant decrease in ^{18}F signal after the use of antibiotics, allowing for rapid screening of antibiotic resistance. Takemiya et al.[60] later used the same molecule to detect implant-induced cardiac infections.

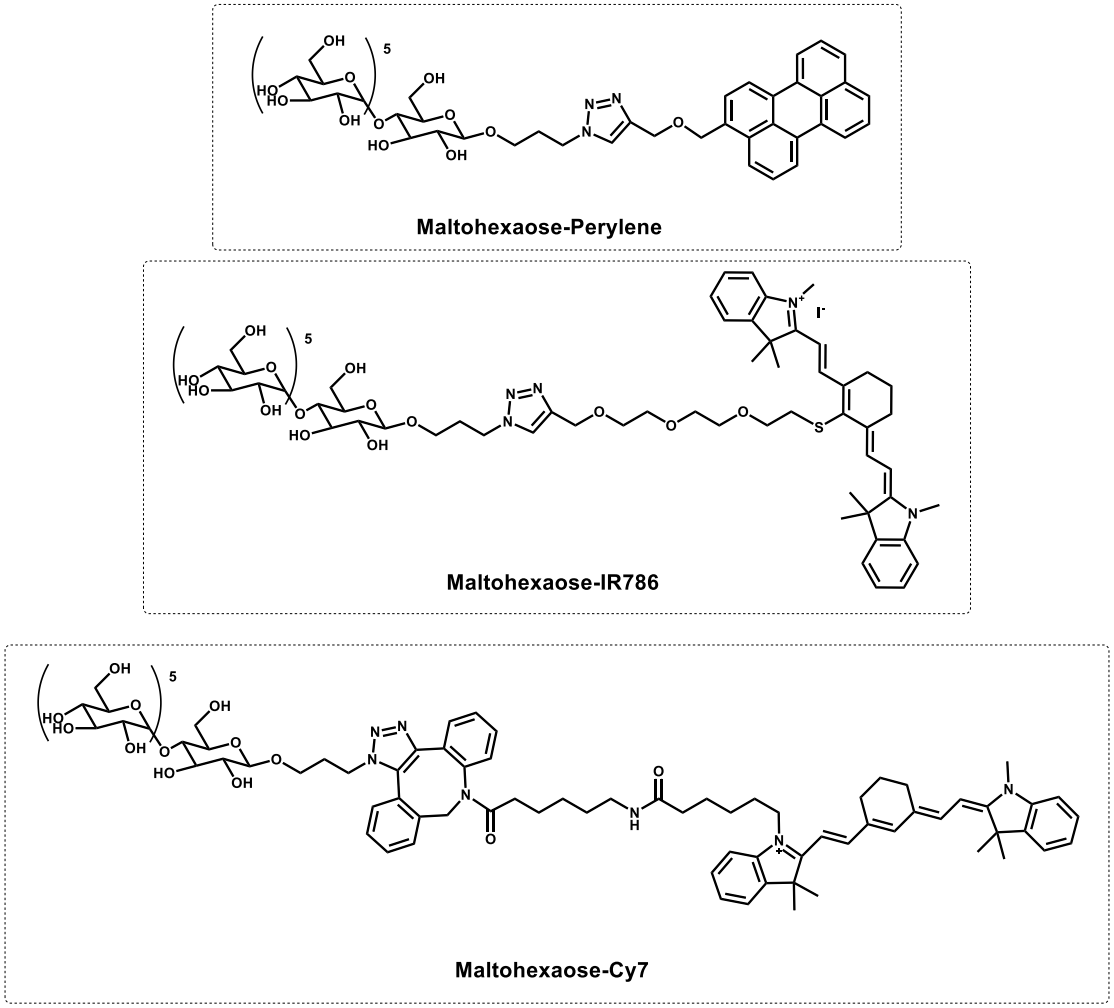


Figure 47: Fluorescence-based maltohexaose probes for detection of bacteria.

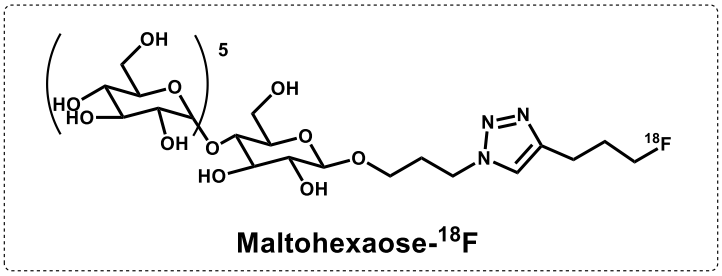
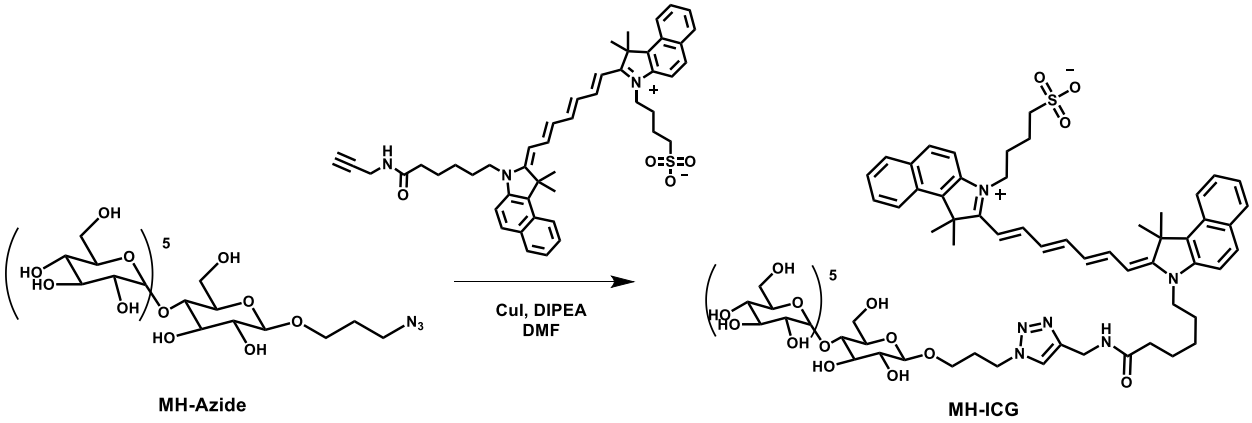


Figure 48: PET-based maltohexaose probes for detection of bacteria.

3.3 Synthesis and Characteristics of MH-ICG

MH-ICG Synthesis

MH-ICG was synthesized through a copper-catalyzed click reaction with MH-azide (25 mg, 0.023 mmol, 1 eq), which was synthesized according to published protocols[59], and commercially available ICG-alkyne (21.5 mg, 0.028 mmol, 1.2 eq, obtained from Iris biotech GMBH), in the presence of DIPEA (1 mg, 7.6 µmol, 0.33 eq) and CuI in catalytic amounts (Scheme 15). After stirring in DMF (0.01 M to ICG-alkyne) at room temperature for 48 hours, the crude was diluted to 0.1X in DI H_2O and filtered before purification using HPLC on a XBridge Prep C8 5 um OBD, 19 x 150 mm semi-prep column (Flow rate: 10 ml/min. Solvent system: MeCN/H2O + 0.1 % TFA, 0-20 min: 37 % MeCN, 20-50 min: gradient to 100 % MeCN. Product eluted at 27 min.). The resulting MH-ICG was analyzed using mass spectrometry (ESI+: found: 1863.7391m/z, calc: 1863.7408m/z for $[C_{87}H_{120}O_{35}N_6SNa+]$, (see Appendix), and the purity was assessed by analytic HPLC (see Appendix). The spectrofluorimetric properties of MH-ICG was assessed, and the absorption and emission maxima were observed at 782 nm and 814 nm respectively, with emission being measured following excitation at 760 nm. (Figure 49). In this study, 760nm and 830 nm was used as the excitation and emission wavelength, respectively.



*Scheme 15: **Synthesis of MH-ICG.** MH-Azide was conjugated with ICG-alkyne through a copper-catalyzed click reaction. respectively. The emission spectrum was generated with an excitation wavelength of 760 nm.*

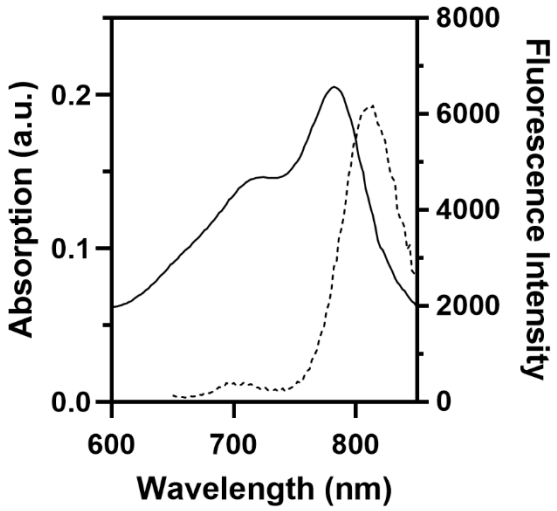


Figure 49: ***Fluorescence spectrum of MH-ICG****. Excitation and emission intensities are indicated as solid and dashed lines, respectively. The emission spectrum was generated with an excitation wavelength of 760 nm. Absorption and excitation maxima were found to be 782nm and 814nm, respectively. Figure reprinted from Takemiya et al.*[62]

In vitro characteristics of MH-ICG

Next, MH-ICG was characterized for its uptake kinetics by S. aureus (ATCC 25923, clinical isolate) and E. coli (ATCC 33456, water sample isolate) by the Takemiya/Taylor lab. Both bacteria were measured at 10^8 CFUs, and the data is shown in Figure 50. The clinical isolate S. aureus was used since it is one of the most commonly observed pathogens in patients with IE[63]. Clinical samples of S. aureus are known to efficiently transport maltose[64]. Briefly, for S. aureus internalization velocity reached a plateau at 20μM, and K_m and V_{max} values (found by Lineweaver-Burk plot analysis) were found to be 5.4μM and $3.0x10^{-6}$ μmol/minutes/10^8CFUs, respectively. For E. coli, these values were found to be 6.9μM and $6.3x10^{-7}$μmol/minutes/10^8CFUs, respectively; and uptake velocity plateaued at a similar concentration. Further, the uptake specificity was confirmed by testing MH-ICG uptake by LamB-mutant E. coli (JW3996-1 E. coli), which did not show Michaelis-Menten kinetics. This illustrates that uptake of MH-ICG is dependent on the bacteria-specific maltose transport system. Cytotoxicity of MH-ICG towards CHO-K1 cells was measured at 24hr and 72hr, and showed no significant toxicity up to 100μM, the highest measured concentration.

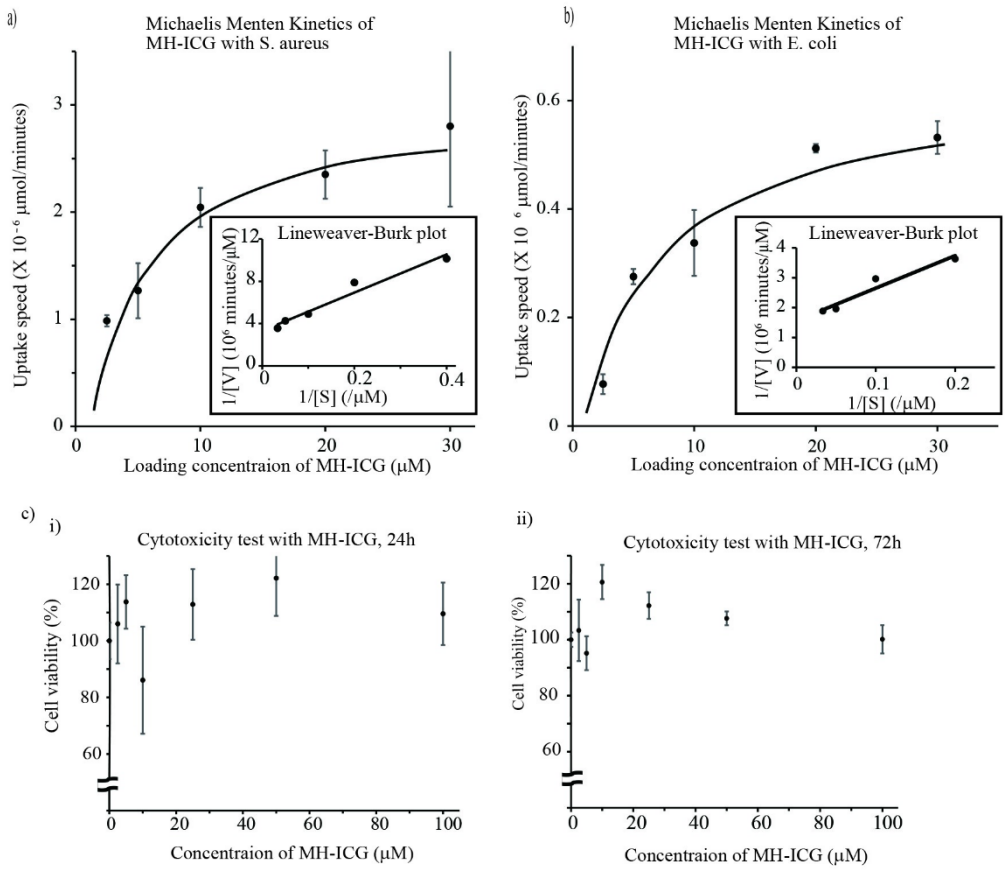


Figure 50: Basic characteristics of MH-ICG in vitro. To evaluate the uptake of MH-ICG in bacteria, S. aureus, E. coli and LamB mutant E. coli were cultured with MH-ICG at various concentrations (n=3 /each concentration). The internalized amount of MH-ICG by S. aureus (a) and E. coli (b) followed Michaelis-Menten kinetics. Km and Vmax calculated with Lineweaver-Burk plot were 5.4 µM and 3.0 X 10^{-6} µmol/minutes/10^8 CFU for S aureus and 6.9 µM and 6.3 X 10^{-7} µmol/minutes/10^8 CFU for E. coli, respectively. LamB mutant E. coli internalized small amount of MH-ICG which at the level of detection. c) Cytotoxicity of MH-ICG was evaluated in CHO-K1 cells. The cell viability was evaluated 24 hours (i) and 72 hours (ii) after loading of MH-ICG. The cells were incubated with 0 to 100µM of MH-ICG (n=4 /each concentration), and no reduction in cell viability was observed at up to 100 µM of MH-ICG at both time points. Figure reprinted from Takemiya et al.[62]

3.4 Imaging of Infective Endocarditis by MH-ICG

MH-ICG accumulates in bacterial vegetations

Finally, the ability for MH-IGC to accumulate *in vivo* in infected areas was investigated at the Takemiya/Taylor labs. An IE rat model was used for the following experiments (see Appendix). The rats were injected with $1x10^8$CFUs/0.1mL of S. aureus. The rats developed right heart endocarditis two days after injection (Figure 51A). The infection was confirmed by a gram staining evaluation (Figure 51B). Total bacterial count in the myocardium was found to be $(1.3 \pm 0.5)x10^8$ CFUs (n = 3).

Accumulation of MH-ICG in IE rats was assessed after an injection of 0.25mL of MH-ICG (1mM) on post-operative day 3. Four hours after the injection, the rats were sacrificed, and the hearts were extracted. The hearts were split into three segments and imaged using an *in vivo*

fluorescence imager (In-Vivo Extreme, Bruker). The fluorescent intensity in the vegetation and that in the left ventricle free wall were measured in each segment, and the mean intensity ratio defined as [mean intensity of vegetation]/ [mean intensity of left ventricle free wall] was calculated for each segment. The average of the intensity ratio in 3 slices was used as the intensity ratio of the sample. In the control samples, the mean intensity in the right ventricle free wall was used instead of that of the vegetation. Figure 51C shows the selective accumulation of MH-ICG in bacterial vegetations after a fluorescent scan. This was further illustrated in the form of fluorescent intensity ratios in Figure 51D, which shows a 2.5-fold increase in fluorescent intensity in the infected area, compared to the left ventricle free wall.

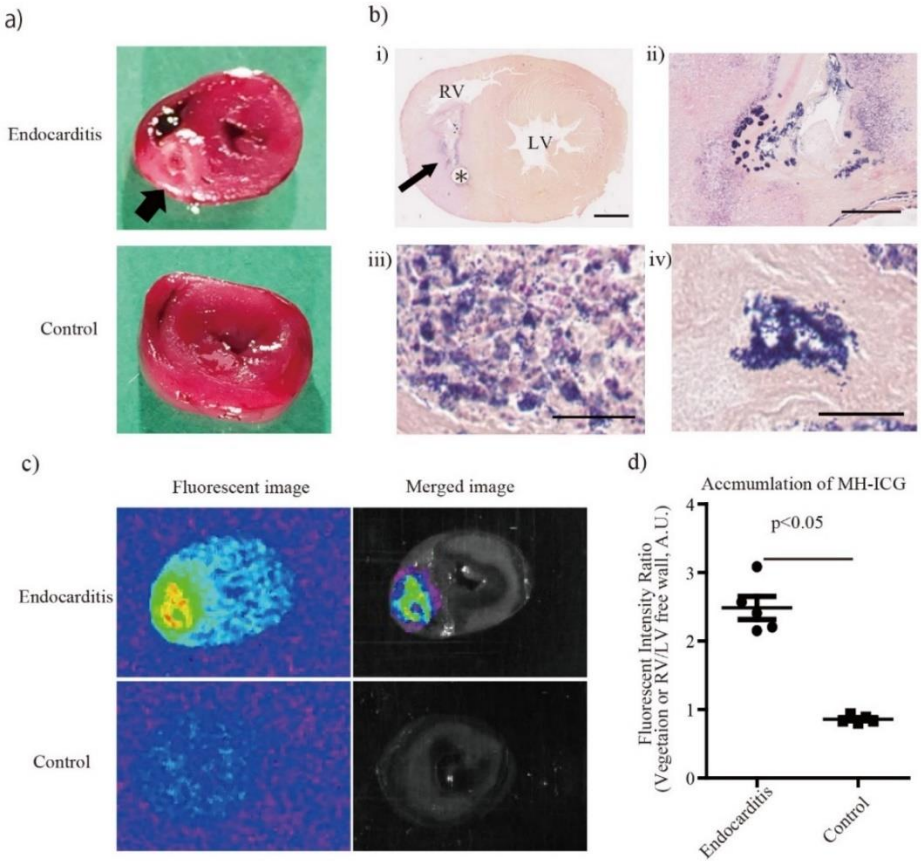


*Figure 51: Accumulation of MH-ICG in vegetation. a) Right heart endocarditis was established by catheterizing in the right ventricle followed by injection of S. aureus. The vegetation was found around the catheter in the right ventricle (the arrow). b) i) Whole image of the specimen. Arrow: vegetation, *: the catheter in right ventricle, RV: right ventricle, LV: left ventricle. Bar: 2mm. ii) Low power field of vegetation around the catheter. In vegetation, spreading and accumulation of Gram-positive bacteria are observed in the fibrous tissue. Bar: 250µm. iii) High power field of vegetation. Gram-positive cocci are observed among large inflammatory cell infiltrate. Bar: 25µm. vi) High power field of vegetation with accumulation of bacteria. The deep purple spot is composed of Gram-positive cocci. Bar: 25 µm. c) Near infrared imaging with MH-ICG in the right heart endocarditis model in rats. Accumulation of MH-ICG was observed only in the vegetation. Hearts from the control rat had very low fluorescent signal. d) Accumulation of MH-ICG in vegetation was quantified as intensity ratio which is defined as [fluorescent intensity in vegetation or the right ventricle]/[fluorescent intensity in the left ventricle free wall]. The*

intensity ratio in IE rats was 2.5 ± 0.2, which was significantly increased when compared to the control rats (n=5 /group). Figure reprinted from Takemiya et al.[62]

3.5 Conclusion

In this chapter, the last ten years of maltohexaose-based bacterial imaging agent development was briefly presented, and this technology was utilized to image infective endocarditis in rats. MH-ICG is an FDA-approved dye with a relatively high wavelength for emission and excitation. This allows for deeper tissue penetration, and we foresee that MH-ICG is a potential tool for *in vivo* IE imaging through transesophageal fluorescence detection. In addition, MH-ICG shows low toxicity in mice. Due to these properties, we foresee that MH-ICG is a promising diagnostic tool for infective bacterial endocarditis.

Appendix C – Chapter 3

C.1 – MH-ICG Characterization

Preparatory HPLC

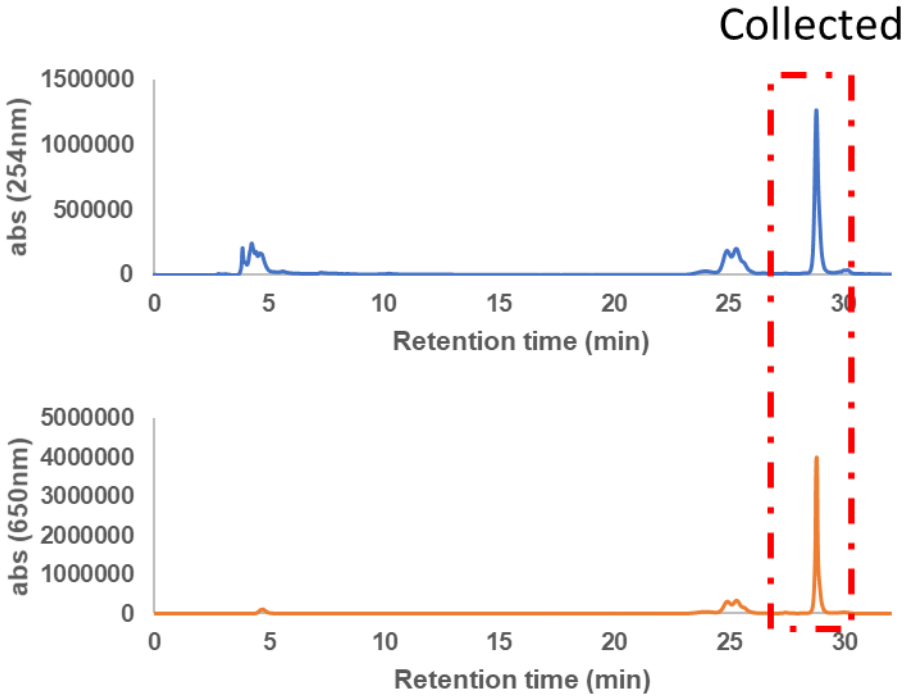


Figure 52: Preparatory HPLC of crude MH-ICG at 254 (blue) and 650nm (orange).The compound which eluded at 27minutes was collected. System: XBridge Prep C8 5 um OBD, 19 x 150 mm semi-prep column (Flow rate: 10 ml/min. Solvent system: MeCN/H_2O + 0.1 % TFA, 0-20 min: 37 % MeCN, 20-50 min: gradient to 100 % MeCN.

Analytical HPLC

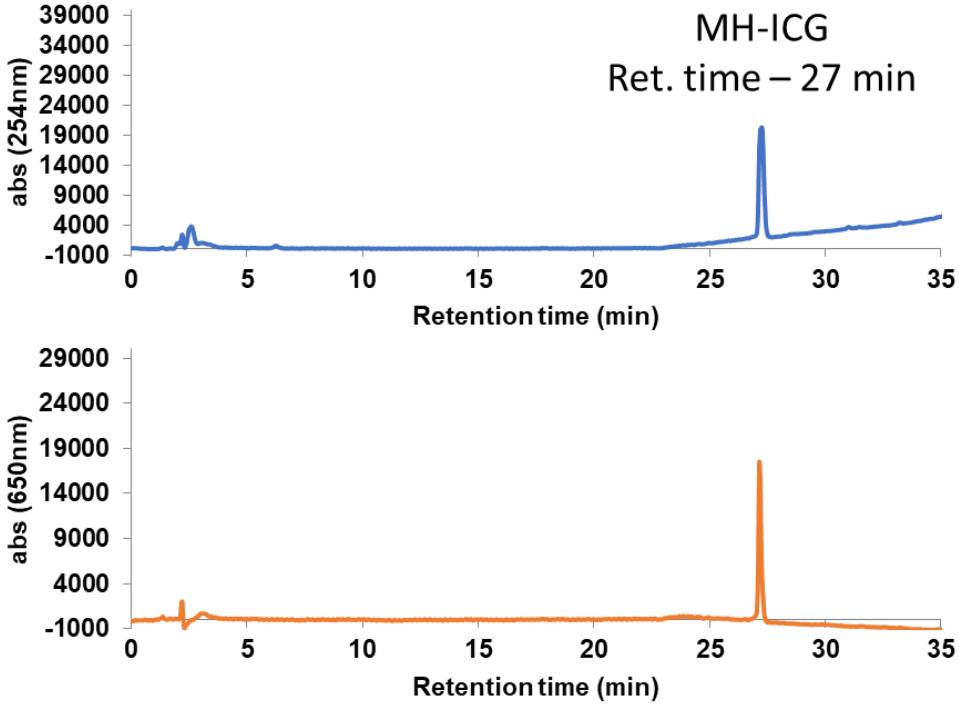
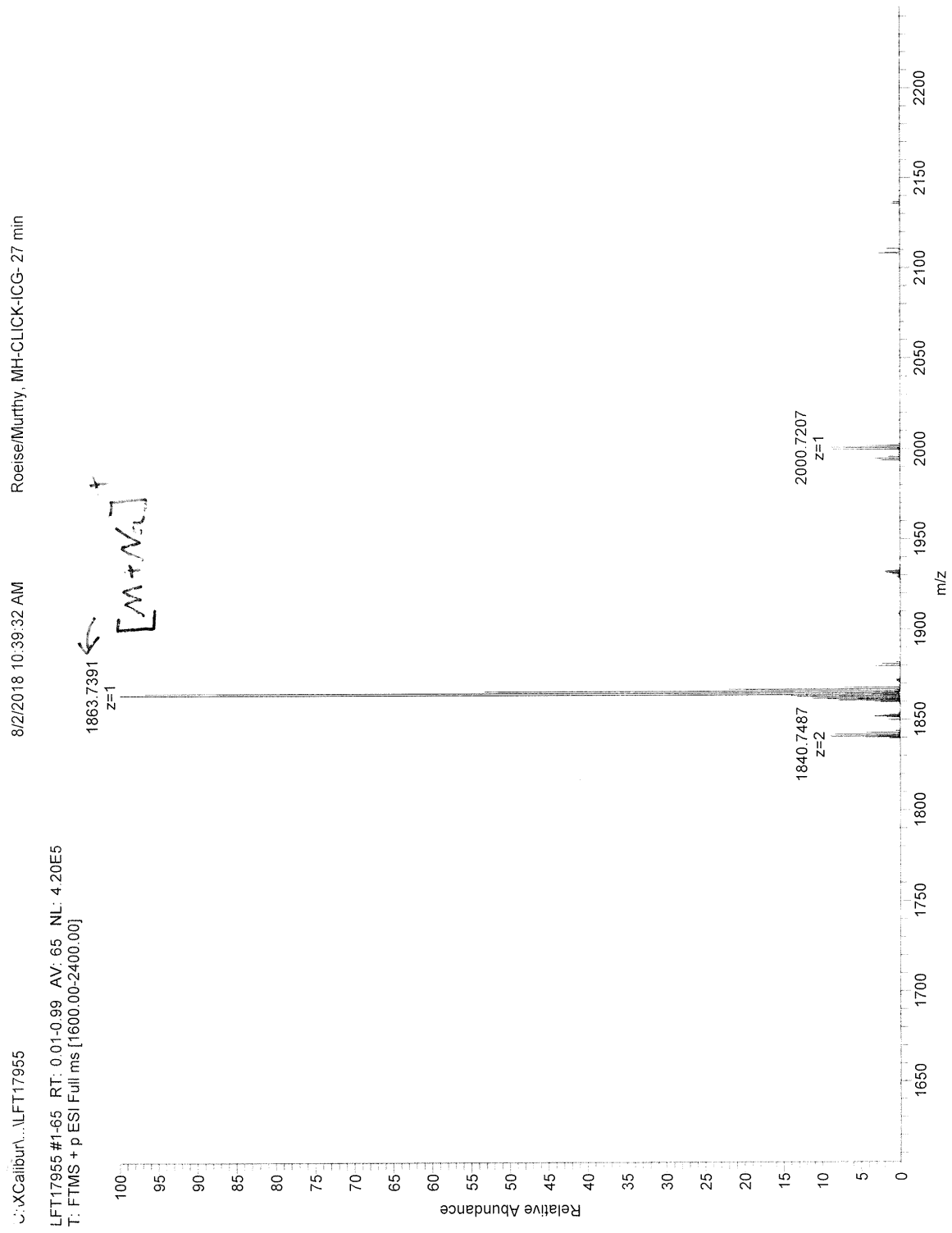


Figure 53: Analytical HPLC of purified MH-ICG at 254 (blue) and 650nm (orange).

Mass Spectroscopy of MH-ICG



C.2 Protocols

Cytotoxicity

CHO-K1 cells (ATCC CCL-61) cultured in F-12K medium with 10% FBS were used to evaluate the cytotoxicity of MH-ICG. On day-1, the cells were seeded in 96 well plates at density of 1 X 10^3 cells /well. On day 0, the cells were loaded with 0, 2.5, 10, 25, 50 or 100 μM of MH-ICG, and the cells were cultured for 24 and 72 hours. On day 1 and 3, the medium with MH-ICG was changed to 100 μl /well of fresh F-12K medium, and then, 20 μl /well of a tetrazolium compound-based reagent (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS, CellTiter 96 AQueous One Solution Cell Proliferation Assay, Promega) was added to evaluate the cell viability. The cells were incubated at 37 °C for 60 minutes, and the formazan generated was measured at 490 nm with a microplate reader (Varioskan LUX, Thermo Fisher Scientific). The viability of MH-ICG treated cells was calculated as [absorbance of MH-ICG loaded cells]/[absorbance of control cells] (n = 4 /each concentration).

Bacterial Uptake Study

In this study, we used Staphylococcus aureus (S aureus, ATCC 25923), a Gram-positive bacteria clinical isolate, as S. aureus is one of the most commonly observed pathogens in patients with IE[63]. To evaluate the affinity of MH-ICG for S. aureus, the bacteria at a density of 1 X 10^8 colony forming units (CFU) /ml were cultured with MH-ICG at concentrations of 0, 2.5, 5, 10, 20 and 30 μM for 1 hour at 37 °C. The bacteria in 1 ml of culture were then washed in phosphate buffered saline (PBS) 3 times and re-suspended in 100 μl of PBS. The amount of MH-ICG internalized by bacteria was evaluated with a fluorescent microplate reader (Varioskan LUX, Thermo Fisher Scientific) (n = 3 /each concentration).

Right heart IE model

Male Sprague Dawley rats (SD rats) weighing 250 to 275g were obtained from Charles River Laboratories. The animals were anesthetized with 1–2% isoflurane, and plastic catheters were inserted into the right ventricle via the right jugular vein. The location of the catheter was determined by the backflow of blood. The proximal end of the catheter was connected to a vascular access port (Braintree Scientific, Inc, U.S.A.) implanted on the back, and the incision was closed. On post-operative day 1 (POD 1), the rats were injected with 1 X 10^8 CFU/0.1ml of S. aureus via the tail vein. On POD 3, two days after inoculation of S. aureus, when a vegetation was formed around the catheter in the right ventricle, the rats were used for experiments. Control rats received no catheterization nor inoculation of bacteria. (n=5/each group) After the experiments, the rats were euthanized with carbon dioxide.

Quantification of bacteria in the right ventricle

On POD 3, the right heart infectious endocarditis model rats (IE rats) were sacrificed with carbon dioxide, and the heart was extracted aseptically. The heart was cut in pieces and placed in Luria-Bertani (LB) broth. The bacteria in the heart were extracted and suspended in LB broth using a TissueLyser II (Qiagen). Serial dilutions of the suspensions were cultured on LB agar plates to determine the number of bacteria in the hearts.

Accumulation of MH-ICG in vivo

On POD 3, the IE rats were injected with 0.25 ml of 1 mM MH-ICG via the tail vein. Rats without catheterization and inoculation were used as controls. The rats were sacrificed 4 hours after injection of MH-ICG, and the hearts were sliced in 3 segments and were scanned with an *in vivo* imaging device (In-Vivo Xtreme, Bruker). The fluorescent intensity in the vegetation and that in the left ventricle free wall were measured in each slice, and the mean intensity ratio defined as [the mean intensity of the vegetation]/ [the mean intensity of the left ventricle free wall] was calculated for each slice. The average of the intensity ratio in 3 slices was used as the intensity ratio of the sample. In the control samples, the mean intensity in the right ventricle free wall was used instead of that of the vegetation.

Statistical Analysis

Analysis was performed with Prism statistical software (GraphPad Software). For comparison of two groups, the Student's T test was used, and $P<0.05$ was regarded as significant. All data are shown as mean ± standard error.

References

1 M. P. Stewart, R. Langer and K. F. Jensen, *Chem. Rev.*, 2018, **118**, 7409–7531.
2 A. K. Varkouhi, M. Scholte, G. Storm and H. J. Haisma, *J. Control. Release*, 2011, **151**, 220–228.
3 R. N. Cohen, M. A. E. M. van der Aa, N. Macaraeg, A. P. Lee and F. C. Szoka, *J. Control. Release*, 2009, **135**, 166–174.
4 I. Mellman, R. Fuchs and A. Helenius, *Annu. Rev. Biochem*, 1986, **55**, 663–700.
5 J. Li, J. J. Røise, J. Zhang, J. Yang, D. L. Kerr, H. Han and N. Murthy, *Chem. Commun.*, 2019, **55**, 4562–4565.
6 N. Murthy, J. Campbell, N. Fausto, A. S. Hoffman and P. S. Stayton, *Bioconjugate Chem.*, 2003, **14**, 412–419.
7 M. Akishiba, T. Takeuchi, Y. Kawaguchi, K. Sakamoto, H.-H. Yu, I. Nakase, T. Takatani-Nakase, F. Madani, A. Gräslund and S. Futaki, *Nat. Chem.*, 2017, **9**, 751–761.
8 C. Kusonwiriyawong, P. van de Wetering, J. A. Hubbell, H. P. Merkle and E. Walter, *Eur. J. Pharm. Biopharm*, 2003, **56**, 237–246.
9 I. Nakase, S. Kobayashi and S. Futaki, *Peptide Science*, 2010, **94**, 763–770.
10 H. Hatakeyama, E. Ito, H. Akita, M. Oishi, Y. Nagasaki, S. Futaki and H. Harashima, *J. Control. Release*, 2009, **139**, 127–132.
11 T. Kakudo, S. Chaki, S. Futaki, I. Nakase, K. Akaji, T. Kawakami, K. Maruyama, H. Kamiya and H. Harashima, *Biochemistry*, 2004, **43**, 5618–5628.
12 B. Lou, S. De Koker, C. Y. J. Lau, W. E. Hennink and E. Mastrobattista, *Bioconjugate Chem.*, 2019, **30**, 461–475.
13 N. Murthy, J. R. Robichaud, D. A. Tirrell, P. S. Stayton and A. S. Hoffman, *J. Control. Release*, 1999, **61**, 137–143.
14 N. Murthy, J. Campbell, N. Fausto, A. S. Hoffman and P. S. Stayton, *J. Control. Release*, 2003, **89**, 365–374.
15 Y. Nie, M. Günther, Z. Gu and E. Wagner, *Biomaterials*, 2011, **32**, 858–869.
16 J. Zhu, M. Qiao, Q. Wang, Y. Ye, S. Ba, J. Ma, H. Hu, X. Zhao and D. Chen, *Biomaterials*, 2018, **162**, 47–59.
17 C.-L. Chan, R. N. Majzoub, R. S. Shirazi, K. K. Ewert, Y.-J. Chen, K. S. Liang and C. R. Safinya, *Biomaterials*, 2012, **33**, 4928–4935.
18 L. Collins, G. J. Sawyer, X. Zhang, K. Gustafsson and J. W. Fabre, *Transplantation*, 2000, **69**, 1168–1176.
19 X. Zhang, G. J. Sawyer, X. Dong, Y. Qiu, L. Collins and J. W. Fabre, *J. Genet. Med.*, 2003, **5**, 209–218.
20 C. D. Ritchie and W. F. Sager, in *Progress in Physical Organic Chemistry*, 1964, vol. 2, pp. 323–456.
21 J. Thyberg, U. Hedin and K. Stenseth, *Cell Tissue Res.*, 1985, **241**, 299–303.
22 K. Müller, T. Nahde, A. Fahr, R. Müller and S. Brüsselbach, *Cancer Gene Ther.*, 2001, **8**, 107–117.
23 C. K. Goldman, L. Soroceanu, N. Smith, G. Y. Gillespie, W. Shaw, S. Burgess, G. Bilbao and D. T. Curiel, *Nat. Biotechnol.*, 1997, **15**, 462–466.
24 F. Cardarelli, L. Digiacomo, C. Marchini, A. Amici, F. Salomone, G. Fiume, A. Rossetta, E. Gratton, D. Pozzi and G. Caracciolo, *Sci. Rep.*, 2016, **6**, 1–8.
25 Z. ur Rehman, D. Hoekstra and I. S. Zuhorn, *ACS Nano*, 2013, **7**, 3767–3777.

26 M. A. Hunt, M. J. Currie, B. A. Robinson and G. U. Dachs, *J Biomol Tech*, 2010, **21**, 66–72.
27 B. P. Bowen and N. W. Woodbury, *Photochem. Photobiol*, 2003, **78**, 582–586.
28 J. A. Poliskey, S. T. Crowley, R. Ramanathan, C. W. White, B. Mathew and K. G. Rice, *Gene Ther*, 2018, **25**, 473–484.
29 S. T. Crowley, J. A. Poliskey, N. J. Baumhover and K. G. Rice, *Gene Ther*, 2015, **22**, 993–999.
30 K. Kizzire, S. Khargharia and K. G. Rice, *Gene Ther*, 2013, **20**, 407–416.
31 D. Vercauteren, R. E. Vandenbroucke, A. T. Jones, J. Rejman, J. Demeester, S. C. De Smedt, N. N. Sanders and K. Braeckmans, *Molecular Therapy*, 2010, **18**, 561–569.
32 G. J. Doherty and H. T. McMahon, *Annual Review of Biochemistry*, 2009, **78**, 857–902.
33 S. Cui, S. Zhang, H. Chen, B. Wang, Y. Zhao and D. Zhi, *ENG*, 2012, **04**, 172–175.
34 P. Lajoie and I. R. Nabi, *Journal of Cellular and Molecular Medicine*, 2007, **11**, 644–653.
35 D. L. Boger, B. E. Fink, S. R. Brunette, W. C. Tse and M. P. Hedrick, *J. Am. Chem. Soc.*, 2001, **123**, 5878–5891.
36 J. ter Maat, R. Regeling, C. J. Ingham, C. A. G. M. Weijers, M. Giesbers, W. M. de Vos and H. Zuilhof, *Langmuir*, 2011, **27**, 13606–13617.
37 J. Li, J. J. Røise, M. He, R. Das and N. Murthy, *Advanced Drug Delivery Reviews*, , DOI:10.1016/j.addr.2020.09.004.
38 D. Wilbie, J. Walther and E. Mastrobattista, *Acc. Chem. Res.*, 2019, **52**, 1555–1564.
39 S. Zhang, J. Shen, D. Li and Y. Cheng, *Theranostics*, 2021, **11**, 614–648.
40 S. S. Usmani, G. Bedi, J. S. Samuel, S. Singh, S. Kalra, P. Kumar, A. A. Ahuja, M. Sharma, A. Gautam and G. P. S. Raghava, *PLoS One*, 2017, **12**, e0181748.
41 Y. Zhang, J. J. Røise, K. Lee, J. Li and N. Murthy, *Current Opinion in Biotechnology*, 2018, **52**, 25–31.
42 A. Khurana, P. Allawadhi, I. Khurana, S. Allwadhi, R. Weiskirchen, A. K. Banothu, D. Chhabra, K. Joshi and K. K. Bharani, *Nano Today*, 2021, **38**, 101142.
43 I. Sarkar, I. Hauber, J. Hauber and F. Buchholz, *Science*, 2007, **316**, 1912–1915.
44 M. A. Hamon, D. Ribet, F. Stavru and P. Cossart, *Trends in Microbiology*, 2012, **20**, 360–368.
45 N. J. Yang, D. V. Liu, D. Sklaviadis, D. Y. Gui, M. G. Vander Heiden and K. D. Wittrup, *Mol. Pharmaceutics*, 2015, **12**, 1992–2000.
46 C. M. Walton, C. H. Wu and G. Y. Wu, *World Journal of Gastroenterology*, 1999, **5**, 465–469.
47 P. Schnupf and D. A. Portnoy, *Microbes and Infection*, 2007, **9**, 1176–1187.
48 A. J. Huisman, L. R. Hartsell, B. P. Krueger and M. J. Pikaart, *J. Chem. Educ.*, 2010, **87**, 299–302.
49 J. C. Skou and M. Esmann, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1983, **727**, 101–107.
50 A. A. Waheed, K. S. Rao and P. D. Gupta, *Analytical Biochemistry*, 2000, **287**, 73–79.
51 S. Reinhard, H. Han, J. Tuma, J. J. Røise, I.-C. Li, J. Li, H. Y. Lee and N. Murthy, *Chem. Commun.*, , DOI:10.1039/D0CC05165A.
52 C. R. Arciola, D. Campoccia and L. Montanaro, *Nat Rev Microbiol*, 2018, **16**, 397–409.
53 W. Boos and H. Shuman, *Microbiol. Mol. Biol. Rev.*
54 M. L. Oldham, S. Chen and J. Chen, *PNAS*, 2013, **110**, 18132–18137.

55 S. Gopal, D. Berg, N. Hagen, E.-M. Schriefer, R. Stoll, W. Goebel and J. Kreft, *PLOS ONE*, 2010, **5**, e10349.
56 X. Wang, C. A. Borges, X. Ning, M. Rafi, J. Zhang, B. Park, K. Takemiya, C. Lo Sterzo, W. R. Taylor, L. Riley and N. Murthy, *Bioconjugate Chem.*, 2018, **29**, 1729–1735.
57 M. M. Welling, A. W. Hensbergen, A. Bunschoten, A. H. Velders, M. Roestenberg and F. W. B. van Leeuwen, *Clin Transl Imaging*, 2019, **7**, 105–124.
58 A. Zlitni, G. Gowrishankar, I. Steinberg, T. Haywood and S. Sam Gambhir, *Nat Commun*, 2020, **11**, 1250.
59 X. Ning, S. Lee, Z. Wang, D. Kim, B. Stubblefield, E. Gilbert and N. Murthy, *Nature Mater*, 2011, **10**, 602–607.
60 K. Takemiya, X. Ning, W. Seo, X. Wang, R. Mohammad, G. Joseph, J. S. Titterington, C. S. Kraft, J. A. Nye, N. Murthy, M. M. Goodman and W. R. Taylor, *JACC: Cardiovascular Imaging*, 2019, **12**, 875–886.
61 X. Ning, W. Seo, S. Lee, K. Takemiya, M. Rafi, X. Feng, D. Weiss, X. Wang, L. Williams, V. M. Camp, M. Eugene, W. R. Taylor, M. Goodman and N. Murthy, *Angewandte Chemie International Edition*, 2014, **53**, 14096–14101.
62 K. Takemiya, J. J. Røise, M. He, C. Taing, A. G. Rodriguez, N. Murthy, M. M. Goodman and W. R. Taylor, *PLOS ONE*, 2021, **16**, e0247673.
63 T. J. Cahill and B. D. Prendergast, *The Lancet*, 2016, **387**, 882–893.
64 J. B. Egan and M. L. Morse, *Biochimica et Biophysica Acta (BBA) - Biophysics including Photosynthesis*, 1966, **112**, 63–73.