Title
Supramolecular chemical biology for applications in protein modification and nanomedicine

Author
Finbloom, Joel Abraham

Publication Date
2018

Peer reviewed|Thesis/dissertation

Supramolecular chemical biology for applications in protein modification and nanomedicine

By

Joel Abraham Finbloom

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 Matthew B. Francis, Chair
Professor Evan W. Miller
Professor David V. Schaffer

Summer 2018

Supramolecular chemical biology for applications in protein modification and nanomedicine

By: Joel Abraham Finbloom

Abstract

Supramolecular chemical biology for applications in protein modification and nanomedicine

By

Joel Abraham Finbloom

Doctor of Philosophy in Chemistry
University of California, Berkeley
Professor Matthew B. Francis, Chair

Supramolecular chemistry is a diverse field that encompasses host-guest interactions, mechanical bond formation, metal coordination, and other noncovalent chemistries. The design strategies of supramolecular chemistry have been implemented in a diverse set of applications, but has seen limited use in the field of chemical biology. By combining supramolecular chemistry with chemical biology, new discoveries and advances could be made that would not be possible without a supramolecular approach. This work therefore sought to combine supramolecular strategies and chemical biology for applications in protein modification and nanomedicine. By using cucurbituril host-guest chemistry, a supramolecular-promoted copper-free click chemistry reaction was developed for protein modification. Cucurbituril host-guest chemistry was additionally explored as a selective protecting group for either the N terminus in the case of cucurbit[7]uril (CB7), or as a lysine protecting group in the case of cucurbit[6]uril (CB6). Supramolecular interactions were further utilized in the design of new turn-on diagnostic probes for ^{129}Xe NMR/MRI detection of cancer, via the rotaxane-mediated suppression, and subsequent activation of Xe-CB6 host-guest interactions. Lastly, self-assembled protein-based nanomaterials were explored for applications in chemotherapeutic drug delivery to glioblastomas, with a special focus on how the morphology of protein-based nanomaterials can affect their biodistributions and drug delivery efficacies. Throughout this work, supramolecular interactions are interwoven with applications and techniques in chemical biology to achieve new advances that would not be possible otherwise.

In memory of my father Dr. David S. Finbloom ז״ל
who first showed me what it means to be a scientist

Table of Contents

Acknowledgments

There are many people without whom I would have been unable to achieve my Ph.D. I am lucky enough to have a group of supportive coworkers, friends, and family who helped me not only survive grad school, but also helped me truly enjoy my time at Berkeley. In the grand tradition of Francis lab theses, I will now acknowledge a bunch of people for their help with my graduate work. While the acknowledgments may be lengthy, it is a testament to the vast support network that is required for someone to make it through grad school. I could not have done this without all of their help.

I'd like to start by acknowledging my advisor Matt Francis, for his help throughout my graduate school experience. Matt has created a lab environment that provided me with the resources, support, and freedom to pursue different avenues of research that I found compelling. Matt was always willing to let me try new things in lab and carve my own research path within the group. It was also nice to nerd out with him about the upcoming Star Wars movies, and bond over our shared love of the lab's fancy espresso machine.

The Francis lab has been a wonderful place for me to work the past 5 years. I've met an incredible group of scientists who have all influenced how I approach my research and scientific aims. The Francis lab maintains a collaborative environment where members frequently work together to achieve their goals. In that vein, I've had the distinct pleasure of working closely with several labmates during my time at Berkeley. Ioana Aanei and Jenna Bernard are the two former grad students that I worked with the most. They are both incredibly dedicated and hard-working individuals who, with me, made up the original "Nanoforma" drug delivery team. While parts of that project frustrated me, Ioana and Jenna always approached it with a positive attitude that made that collaboration much better. The newest addition to the team was Sarah Klass, a talented scientist who is taking certain aspects of that project and implementing them in new and innovative ways. Sarah not only works with me on the drug delivery projects, but also works in the field of hyperCEST detection with cucurbituril rotaxanes. She is my successor to both of those projects and I'm sure that she'll take them in important directions that I would never have thought of. I've also had the pleasure of working with Carson Bruns, a fomer postdoc in the lab who first introduced me to the world of rotaxanes and the mechanical bond. Although we knew each other in the Stupp lab at Northwestern, where Carson was a grad student and I was an undergrad, we only started to work together when Carson joined the Francis lab during my second year. I am very grateful to Carson for all of his help with the cucurbituril rotaxane project and for starting me on the path of supramolecular chemistry. Ariel Furst is another postdoc with whom I had a brief collaboration. Although we were never able to get that project fully up and running, working with Ariel during that brief time showed me just how thoughtful and hard-working she is.

Outside of the Francis lab, I have collaborated with a group of talented scientists both at Berkeley, and at UCSF. My closest collaboration was with the Pines lab, and particularly with Ashley Truxal and Clancy Slack. Clancy and I worked very closely on all of the cucurbituril rotaxane work, as Clancy provided the much-needed NMR expertise for the project. While we were both occasionally cynical about grad school, we worked very well together to create a fruitful and fun collaboration. Ashley started working with me when Clancy graduated and I am grateful for all of her help and advice with this work.

There are quite a few members of the lab who had a particularly strong influence on me, even though we never had a chance to collaborate on a project together. Many of them have gone hiking, camping, climbing, and traveling with me, and they made the Francis lab the best possible place

for me to be. I would like to give a special shout-out to my 733 Latimer buddies, with whom I shared a lab space for the past 5 years. Abby Knight, who first welcomed me into the lab and perhaps more than anyone in lab influenced my approach to science and grad school. Kanwal Palla, who made 733 an incredibly fun room to be in, and who decorated my desk with pink snowflakes when I first joined the lab. Ioana Aanei, for always being an example of a hard-working labmate and for always maintaining a positive attitude. While they have all graduated, I will be forever grateful to them for their friendliness and help when I was just starting in lab. My current 733 roommates are some of my closest friends in lab and help foster a fun work environment every day: Tyler Hurlburt, Kristin Wucherer, Daniel Brauer (our newest addition), and our honorary 733 roommate Matt Smith. I'll miss our 733 shenanigans more than I'd care to admit. I'd like to especially acknowledge Kenneth Han, an incredibly talented undergraduate student whom I mentored in lab. Kenneth is an intelligent, hard-working, fun, and quirky person who helped me immensely in every aspect of my work. I know he will go far in his scientific career.

Lastly, I'd like to acknowledge my family for their tremendous love and support over the past 5 years. My mom and my stepdad Sam are wonderful parents who have supported me in every way imaginable. Their frequent visits to the Bay are something that I always look forward to, as is visiting them at home for the holidays. I want to especially thank my mom for raising me through good times and bad to become the person I am today. My brother Aaron has always nurtured my creative side, and even though we are very different people, he has always inspired me to approach my scientific pursuits with greater creativity and outside-the-box thinking. I look forward to us both getting our Ph.D.'s and becoming Drs. Finbloom. I would also like to acknowledge the menagerie of pets that I've had throughout grad school: Cornelius, Cecelia, Winston, Humphrey, and now Shasta have all helped me calm down after a tough day at the lab, and have helped me maintain my mental wellbeing throughout grad school. Finally, I want to thank my wife Rebecca, for her endless love and support during grad school. Becca moved to the Bay with me even though she didn't have a job lined up, and for that I am eternally grateful. She has continually inspired me to be a better person and a better scientist. Since moving here, she has embraced the Bay Area lifestyle. She's a talented and successful product manager at a San Fran tech company and has become a real weekend warrior. She constantly encourages me to go on adventures with her, and I am forever grateful to her for making me a more adventurous and fun person. Rather than spending my weekends in lab, I have had the great pleasure of exploring the Bay and the world with Becca through our many travels. I can't wait for our next big adventure together.

Chapter 1

Supramolecular interactions in chemical biology and nanomedicine

ABSTRACT: Supramolecular interactions present opportunities to advance a wide range of chemical biology applications, including in proteomics and drug delivery. For many of these applications, covalent strategies are traditionally used. However, supramolecular chemistry can offer certain advantages over covalent strategies, and offers unique opportunities to create nanomaterials that would not be accessible otherwise. Recent studies have used supramolecular interactions, particularly host-guest chemistry, to control protein modification and immobilization. Supramolecular strategies for protein-conjugate purification and capture have been utilized, as well as for protein modification via host-guest interactions and metal coordination. Supramolecular interactions have also seen use in the fields of drug delivery and diagnostics, leading to improved drug bioavailability and the development of turn-on probes. Lastly, naturally occurring and *de novo* designed supramolecular interactions in protein systems lead to self-assembling protein-based nanomaterials, which have been used in drug delivery and diagnostic nanomedicine.

This chapter is based on the following publication:
JA Finbloom, MB Francis. Supramolecular strategies for protein immobilization and modification. *Curr. Opin. Chem. Biol.* 2018 [in press]

1.1 Introduction

Protein modification and immobilization are important techniques that are used in a variety of fields, including drug delivery, environmental science, and proteomics.[1–4] For these applications, protein modification reactions should offer chemical orthogonality to other functional groups that are present in complex biological mixtures or on protein surfaces. It is also frequently required that proteins only be modified at a single site of interest, resulting in a homogeneous product, rather than in a heterogeneous mixture that can display analogous heterogeneous properties.[1] While covalent strategies have been the standard approach for protein modification and immobilization, new advances in supramolecular chemistry could afford alternative approaches to controllable bioconjugation reactions. Supramolecular chemistry, and particularly host-guest chemistry, allow for highly controllable interactions with tailorable binding affinities.[5–8] Additionally, many supramolecular interactions with synthetic macrocyclic hosts are highly specific in nature, and therefore mitigate side reactions that interfere with downstream applications. By taking advantage of these properties, researchers have expanded the scope of bioconjugation reactions, and have developed protein conjugates that would not have been possible with traditional covalent approaches. Supramolecular bioconjugation also offers the potential for new developments in protein-based materials, as these modification strategies can afford unique classes of materials that combine the function and precision of proteins with the responsive behavior of supramolecular assemblies.

1.2 Protein purification through host-guest interactions

The use of biotin-avidin interactions could be considered the first supramolecular strategy widely used in protein chemistry.[9,10] However, the exceptionally high binding affinity of biotin for avidin, estimated at 10^{15} M^{-1}, poses challenges in separating heterogeneous mixtures of biotinylated proteins.[10] Several examples of more tunable purification strategies have been reported, which rely on weaker host-guest interactions between synthetic molecules conjugated to proteins and resins modified with macrocyclic hosts.

In 2006, Francis and coworkers reported a strategy for purification of fluorescently-labeled biomolecules using cyclodextrin host-guest chemistry.[11] Proteins were modified with fluorescent dyes and purified with cyclodextrin-modified beads, which selectively bound to the protein conjugates. Since that initial report, other examples of cyclodextrin-mediated protein purification have been reported,[12] including a general approach reported by Francis and coworkers in 2015.[13] Proteins were modified heterogeneously with an azobenzene derivative that has a high binding affinity for β-cyclodextrin. Chromatographic purification of the proteins with a cyclodextrin column afforded separation of unmodified, singly, and doubly modified proteins. After purification, the azobenzene was cleaved with sodium dithionite, exposing an aniline on the surface that can undergo an orthogonal reaction for site-selective modification.

1.3 Supramolecular protein immobilization and detection

The interactions between synthetic hosts and native protein residues have been used for protein recognition and analysis. The binding interaction (K_a = 1.5x10^6 M^{-1}) between the N-terminal

phenylalanine of insulin and cucurbit[7]uril (CB7) macrocycles was reported by Urbach and coworkers in 2011.[14] Since that report, other synthetic hosts have been used as receptors for protein binding and recognition. These synthetic protein receptors have seen use in proteomic analysis. Urbach, Kim, and coworkers used N-terminal Phe-CB7 interactions as an MS/MS enhancement technique for the detection of proteins and protein digests.[15] Other synthetic receptors have been developed for selective binding to post-translational modifications on protein surfaces.[16,17]

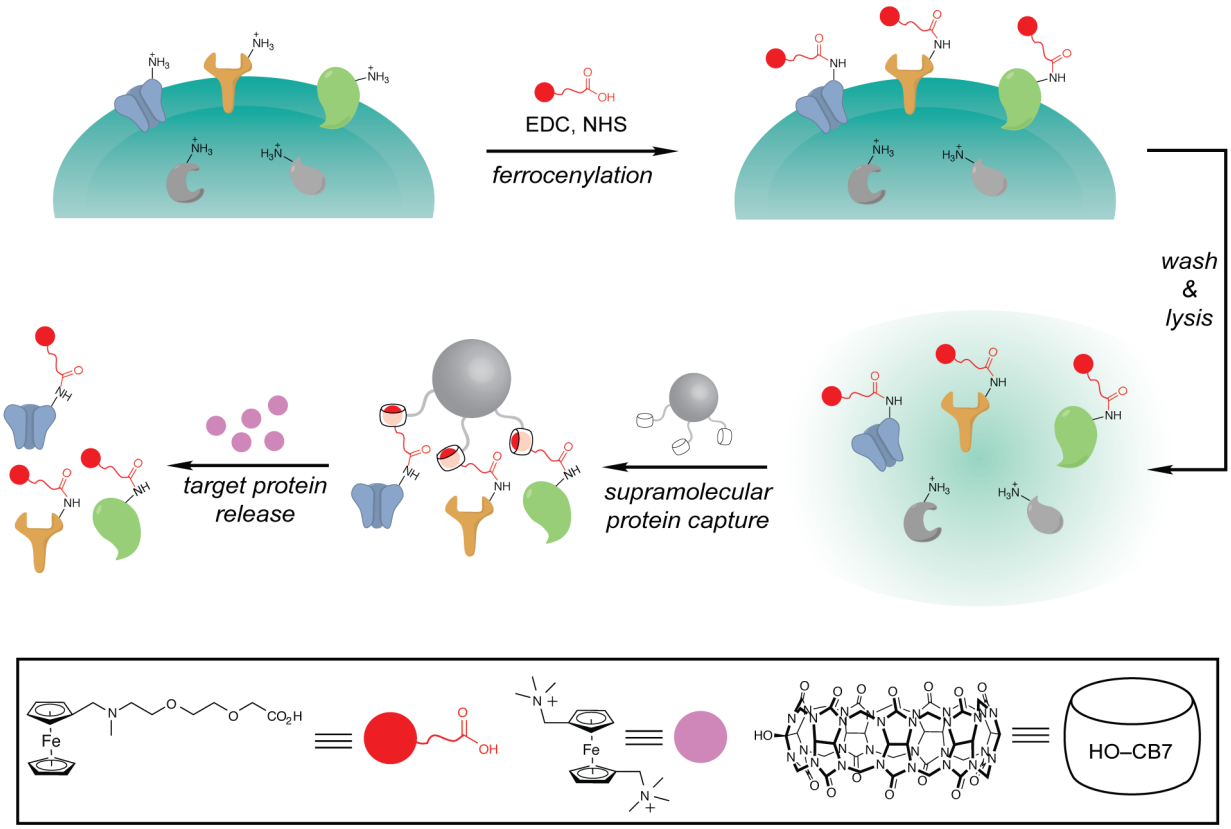


Figure 1.1. Supramolecular protein capture in complex cellular environments mediated by ferrocene-CB7 host-guest chemistry. Figure adapted from reference 18.

In 2011, Kim and coworkers reported a strategy for the enrichment and analysis of synthetically modified proteins in cell culture using ferrocene-CB7 interactions (Figure 1).[18] Membrane-bound proteins were modified with a ferrocene NHS-ester derivative with a CB7 binding affinity of ~10^{12} M^{-1}. The modified proteins were then enriched with CB7-functionalized beads, and were released for analysis by competition against a small molecule ferrocene derivative. This technique overcame several challenges associated with biotin-mediated enrichment, namely cross-contamination from naturally biotinylated proteins, and avidin contamination during the enrichment process. In a follow-up study, the same research group used a bifunctional probe for the proteomic analysis of intracellular proteins. In this work, one segment of the reagent labeled the proteins of interest, while the other side contained a ferrocene moiety for CB7-binding.[19] This strategy can be applied to a variety of applications, with a ferrocene tag substituted for the traditional biotin handle.

In addition to protein capture for proteomics, protein immobilization on surfaces has been used in applications ranging from regenerative medicine to the study of signaling pathways. While nonspecific adsorption of proteins on surfaces and covalent modification strategies are traditionally used, supramolecular means of protein immobilization offer opportunities for modulated and reversible binding.[20–23] In one example, Francis and coworkers used host-guest interactions to enable enzyme recycling.[24] Sortase A was modified site-specifically at the N terminus with a lithocholic acid (LA) derivative. LA has a reported binding affinity of 10^6 M^{-1} for β-cyclodextrin (βCD). This mid-range affinity allowed for binding of sortase A to βCD-modified resin, followed by competition-mediated release. The enzyme retained its activity after LA modification and was easily removed from solution by bead capture for repeated use.

1.4 Supramolecular protein modification

Coordination chemistry offers a powerful approach to modify proteins with synthetic substrates. The most well-known example of synthetic coordination in the realm of protein chemistry is the affinity of hexahistidine peptide sequences for nickel.[25] This affinity has been used extensively in biochemistry for the purification of proteins containing a 6xHis tag at either the N or C terminus via Ni-NTA column chromatography. Recent work has taken advantage of this interaction for applications in protein modification.[26,27] In one example, Lee and coworkers expressed hexahistidine-tagged proteins and site-specifically modified them with Ni-NTA functionalized polyethylene glycol (PEG) polymers.[28] These PEGylated proteins displayed improved half-lives and *in vivo* efficacy over unmodified proteins. In another example, Gothelf and coworkers used the 6xHis to Ni-NTA interaction to direct covalent modification of proteins with DNA-NHS esters.[29] Typically, NHS ester modification results in a heterogeneous mixture of protein conjugates, with reactions occurring at the N terminus and any surface lysine residues. The researchers expressed GFP with a 6xHis tag at the N terminus in order to bind a Ni-NTA-modified strand of DNA. The Ni-NTA-DNA was then hybridized to its complementary DNA strand, which contained an NHS ester moiety. Because of the hybridization, the NHS ester only reacted with the lysine residue in closest proximity to the 6xHis tag. Thus, a combination of supramolecular interactions, namely metal coordination and Watson-Crick-Franklin base pairing, led to site-specific covalent lysine modification.

Synthetic host-guest pairs are oftentimes completely distinct from other functionalities that are found on protein surfaces, providing highly specific binding events, and therefore opportunities for site-specific protein functionalization through non-covalent interactions. Two basic approaches have been explored in this regard: (1) the protein (modified or native) forms a binary complex with a respective partner bound to a synthetic substrate,[30–34] or (2) the protein and substrate have each been modified with synthetic guests that are brought together in a ternary complex with an additional macrocyclic host.[35–38] These strategies have been used to modify proteins with both synthetic cargo and biomolecules.

As an example of the first approach, in 2016, Langer, Anderson, and coworkers reported the supramolecular PEGylation of biopharmaceuticals. The researchers site-specifically PEGylated proteins at the N terminus using cucurbit[7]uril host-guest chemistry (Figure 2a).[30] CB7-functionalized PEG derivatives were synthesized and used to modify insulin. PEGylation of insulin with CB7-PEG at the N-terminal phenylalanine led to marked enhancement in stability and

therapeutic properties. Unlike covalent attachment strategies, this method did not alter the native structure of insulin, demonstrating a potential advantage with supramolecular PEGylation.

In 2011, Scherman and coworkers reported a ternary approach to protein PEGylation, whereby a ternary complex was formed within a cucurbit[8]uril (CB8) cavity between methylviologen (MV)-modified proteins and naphthalene-modified PEG (Figure 2b).[36] MV and CB8 form a 1:1 binding complex exclusively, but upon addition of an electron-rich substrate, such as naphthalene, a 1:1:1 complex is formed between MV and naphthalene within the CB8 cavity. Interestingly, electron-rich substrates such as naphthalene and anthracene do not bind to CB8 alone, and only bind in the presence of the MV guest.[7] Thus, this strategy allows for a well-controlled complexation between two distinct substrates without the occurrence of other binding pathways. Ternary complexation with CB8 has been used to induce protein-protein dimerization and higher order assemblies,[37–41] and this topic has been highlighted in recent reviews.[42,43]

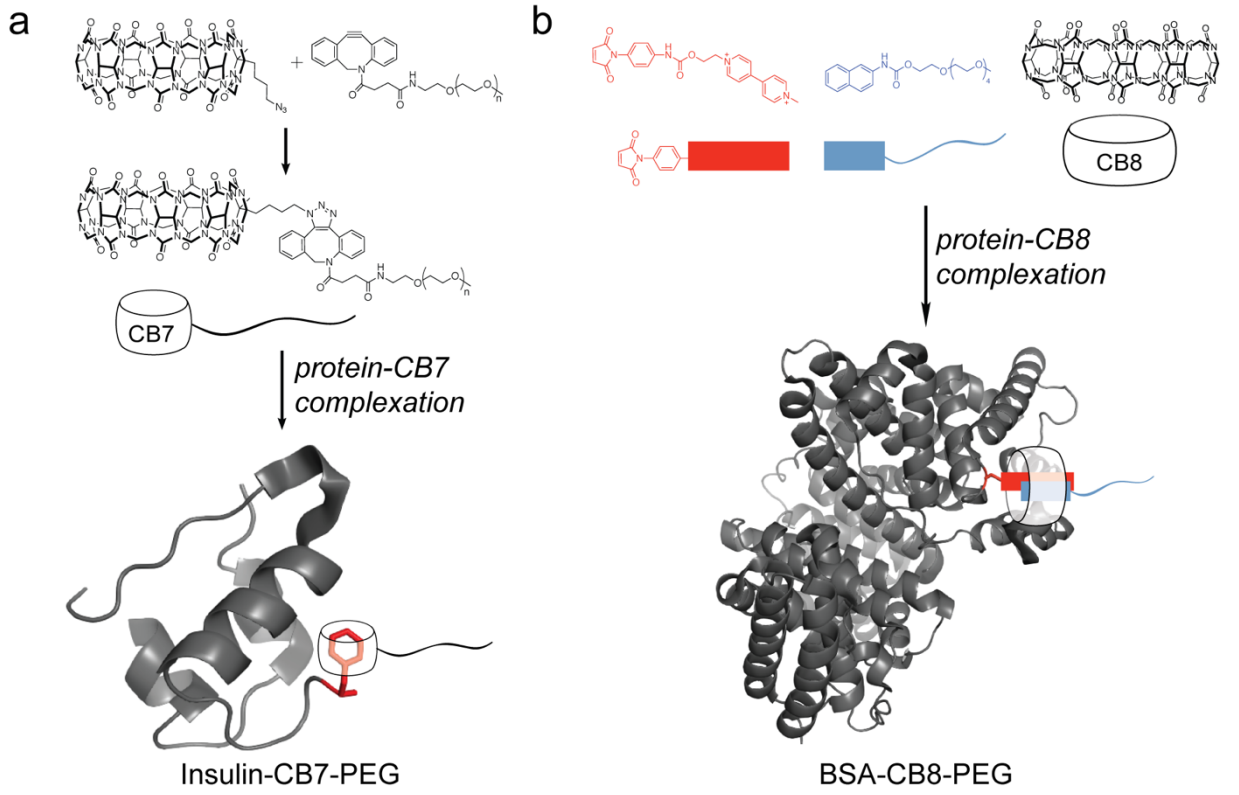


Figure 1.2. PEGylation of proteins using cucurbituril host-guest interactions through (a) binary Phe-CB7 or (b) ternary MV-Napthalene-CB8 complexation. Figure adapted from references 30,36.

1.5 Controlling covalent protein modification through supramolecular chemistry

Supramolecular chemistry offers a unique toolkit to control chemical reactions in biological environments. Several examples of host-guest controlled proteins have been reported,[44,45] as well as the first report of a protein-rotaxane conjugate,[46] which may facilitate future stimulus-responsive control over enzymatic activity. In the realm of protein modification, supramolecular interactions have seen recent use as a means of controlling chemical reactivity.[29] In 2016, Pentelute and coworkers reported a supramolecular means of controlling cysteine arylation.[47] When cysteine

is located within an FCPF peptide sequence, a π-clamp is formed with a perfluoroaromatic compound, placing the small molecule reactive handle in close proximity to the thiol, promoting the reaction at that specific site (Figure 3). The researchers used this technique to site-selectively modify several proteins, including a reduced antibody that contained over 20 cysteines on its surface. In a subsequent report, the researchers determined how salt effects can be used to tune the reactivity of the π-clamp cysteines.[48]

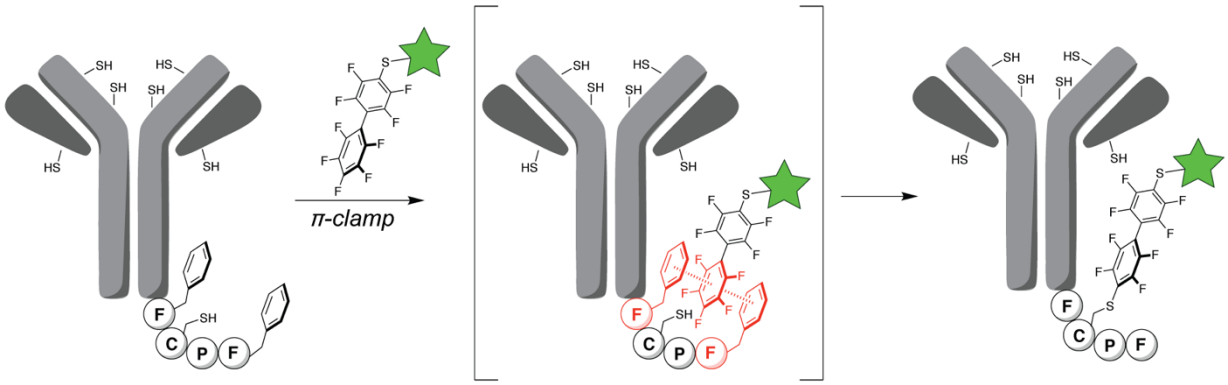


Figure 1.3. Supramolecular π-clamp directed cysteine arylation led to the selective modification of antibodies with biotin and drug cargos. Figure adapted from reference 47.

1.6 Supramolecular chemistry for drug delivery and diagnostics

Supramolecular approaches have seen use in drug delivery and diagnostic medicine. Supramolecular strategies of drug encapsulation have been established for decades and are used in FDA-approved pharmaceuticals.[8] Commonly, cyclodextrins are the macrocycle of choice for drug encapsulation, as hydrophobic molecules bind tightly into the cyclodextrin cavity. By complexing hydrophobic drugs into macrocycles, the complex gains increased water solubility, better bioavailability, and a greater *in vivo* half-life.[8,49] While cyclodextrin-drug complexes are the only ones currently available on the market, cucurbiturils have also been used for similar applications and have seen success in animal models.[50,51] Currently, supramolecular complexes have primarily seen use in drug delivery applications; however, supramolecular strategies could offer opportunities in responsive diagnostic complexes, whereby a turn-on probe can be created through triggered host-guest interactions.

Supramolecular nanomaterials have also been explored for drug delivery applications. These approaches involve either the encapsulation of a drug cargo into a macrocycle that is incorporated into the nanomaterial backbone,[8] or rely on supramolecular interactions to trigger the self-assembly of a nanomaterial (Figure 1.4a).[52] Both strategies have yielded success in improving the bioavailability of drug molecules and diagnostic cargo.[53] A particularly promising opportunity for supramolecular materials is the incorporation of responsive host-guest interactions to release drug cargo on demand. While most stimuli-responsive nanomaterials rely on covalent bond breakages to release cargo or disassemble the nanomaterial, supramolecular host-guest interactions could provide a modular approach for triggered nanomaterial disassembly and subsequent cargo release.

A less explored strategy for triggered drug delivery utilizes supramolecular interactions as gate-keepers for drug release. Stoddart and coworkers developed a mesoporous silica-nanoparticle system whereby a drug or imaging cargo is loaded into the nanoparticles and the external surface is modified with cucurbituril pseudorotaxanes.[54] These pseudorotaxanes prevent any cargo from releasing until an external stimulus such as pH change or UV light triggers the release of the macrocycle from the supramolecular complex.[54] An alternative approach was taken by Rotello and coworkers with cytotoxic gold nanoparticles. The gold nanoparticles were rendered nontoxic in cell culture by capping gold nanoparticles with amine-cucurbituril complexes.[55] These supramolecular complexes caused the gold nanoparticles to accumulate into endosomes where they were unable to harm the cell. Upon addition of adamantylamine to the cell culture, the cucurbituril caps were competed off and the newly toxic gold nanoparticles were released from the endosome and effectively killed the cancer cells (Figure 1.4b). Supramolecular strategies offer many exciting opportunities in the fields of drug delivery and diagnostics, and this topic is covered in several recent reviews.[8,53]

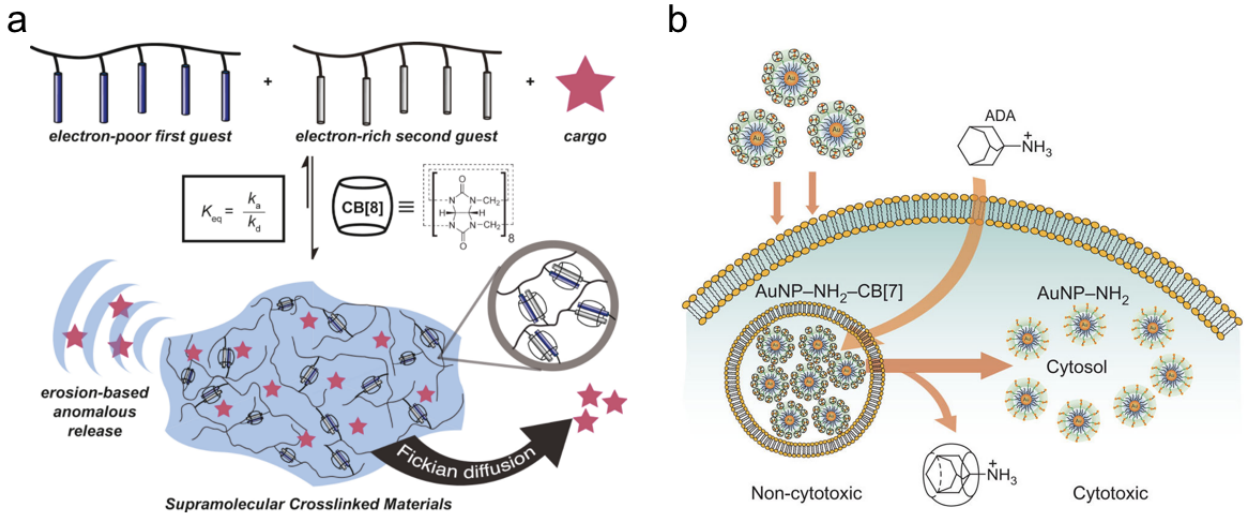


Figure 1.4. Supramolecular materials for drug delivery applications. (a) A supramolecular hydrogel network for drug encapsulation and subsequent release. (b) Supramolecular complexation onto gold nanoparticles leads to competition-triggered nanoparticle cytotoxicity. Figures from references 52 and 55.

1.7 Self-assembling protein nanomaterials

Protein-based nanomaterials are a particularly promising class of nanocarrier for drug delivery and diagnostic applications. These materials rely on the supramolecular self-assembly between protein monomers into larger nanoscale scaffolds of varying size and shape (Figure 1.5). Protein-based nanomaterials, and specifically virus-like particles (VLPs) based on naturally occurring viruses, have demonstrated effectiveness in drug delivery and imaging applications.[2,56] Unlike other synthetic delivery vehicles, such as polymeric micelles and liposomes, VLPs are homogenous in their size distribution and are produced by inexpensive recombinant expression.[57,58] Further, protein-based nanomaterials are degradable in the body and have demonstrated few toxicity issues.[56,58] Protein-based nanomaterials also allow for site-selective modification through amino acid mutagenesis of natural or noncanonical amino acids into the

protein backbone.[2] This site-selective conjugation allows for greater control over the location and amount of cargo loaded onto the capsids, which can have a significant effect on cancer targeting and imaging efficiencies.

While viral capsids offer ease of control and homogeneity in the nanoscale assemblies of the protein monomers, non-naturally-occurring protein assemblies can also be designed for use in varying applications. Synthetic supramolecular interactions offer possibilities in designing reversible protein assemblies through host-guest chemistry. Cucurbit[8]uril has been used in the design of supramolecular protein dimers and higher order assemblies.[42] While these protein-based materials have not yet seen use in diagnostic or drug delivery applications, they may be beneficial for applications in which reversible protein assembly is desired. An alternative approach to develop non-natural protein assemblies was taken by Baker and coworkers. Through Rosetta computational modeling, new higher assembly protein structures can be designed *de novo*.[59] Expression of the designed proteins led to new shapes and morphologies of protein-based nanomaterials, as well as virus-like nanomaterials with properties that had not been observed previously.[60–62]

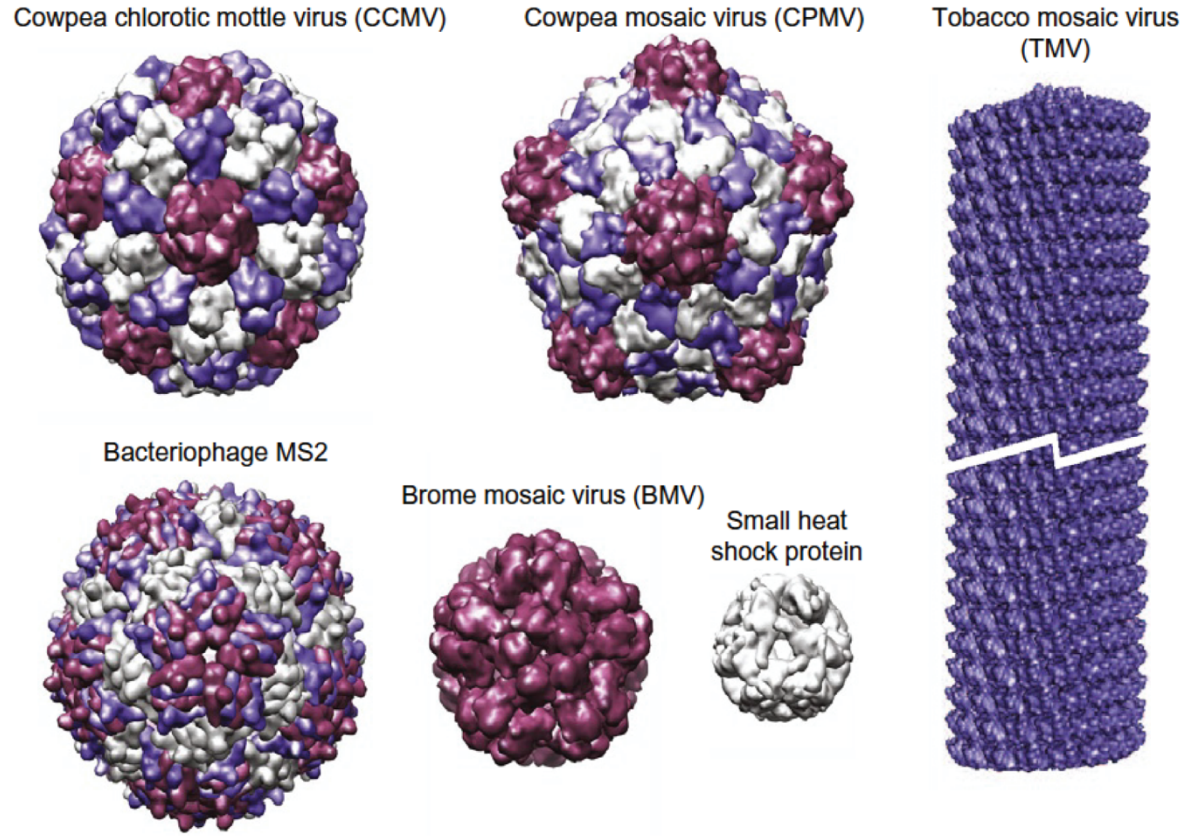


Figure 1.5. Viral capsid protein-based nanomaterials of varying shapes and sizes have been exploited for drug delivery and imaging applications. Figure from reference 57.

An underexplored aspect of supramolecular protein assemblies is stimuli-responsive protein capture and release. Host-guest chemistry offers opportunities for responsive behaviors, such as those used in supramolecular materials and molecular machines. Stimuli such as light, pH, and redox conditions can be used to dissociate and re-associate host-guest pairs for drug delivery and

diagnostics.[8] Additionally, the incorporation of mechanical bonds into proteins could lead to new biomaterials. Responsive rotaxane linkages within a protein network could create protein-based artificial muscles,[63] and the installation of catenanes into protein backbones could lead to enhanced stability.[64] Lastly, host-guest chemistry could allow for greater control over protein function. By installing molecular machines near enzyme active sites, stimuli-responsive shuttling could be used to activate or deactivate the protein. This process could similarly be used to direct binding events of proteins to their receptors for use in diagnostics. By combining developments in responsive supramolecular materials with protein-based assemblies, new advances in chemical biology can be realized.

1.8 Conclusions and thesis overview

Noncovalent interactions such as biotin-avidin have been used for decades in protein research, but only recently have synthetic supramolecular strategies received attention in chemical biology. Supramolecular interactions display the bio-orthogonality that is necessary for protein modification in complex mixtures, and can undergo complexation at the low concentrations, physiological pH, and temperature range stipulations necessary for biological applications. Complexes with mid-range binding affinities can be tailored for protein immobilization, purification, and recycling, while tight binding pairs are used for protein modification. Additionally, researchers have shown that supramolecular interactions can be used to direct covalent modification strategies, affording site-specific modification of natural amino acids, or promoting reactions with non-native residues.

This work adapts supramolecular strategies for applications in protein modification and nanomedicine. By utilizing cucurbituril host-guest chemistry, a supramolecular-promoted azide-alkyne click reaction was developed for protein modification. This reaction proceeds in mild conditions with few equivalents of substrate, and allows for the conjugation of both synthetic substrates and complex biomolecules onto protein surfaces. Cucurbituril host-guest chemistry was additionally explored as a selective protecting group for either the N terminus in the case of cucurbit[7]uril (CB7), or as a lysine protecting group in the case of cucurbit[6]uril (CB6). Supramolecular interactions were further utilized in the design of new turn-on diagnostic probes for ^{129}Xe hyperCEST NMR/MRI. We reported a method for blocking interactions between ^{129}Xe and cucurbit[6]uril hosts until activation by a specific chemical event. CB6 rotaxane probes prevented ^{129}Xe interaction with the CB6 macrocycle until a cleavage event released the CB6, which then produced a ^{129}Xe@CB6 NMR signal. Lastly, protein-based nanomaterials were explored for applications in chemotherapeutic drug delivery to glioblastomas, with a special focus on how the morphology of protein-based nanomaterials can affect their biodistributions and drug delivery efficacies. Throughout this work, supramolecular interactions are interwoven with applications and techniques in chemical biology to achieve new advances that would not be possible otherwise.

1.9 References

(1) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chemie Int. Ed.* **2009**, *48* (38), 6974–6998.

(2) Witus, L. S.; Francis, M. B. Using Synthetically Modified Proteins to Make New Materials.

Acc. Chem. Res. **2011**, *44* (9), 774–783.
(3) Roberts, A. M.; Ward, C. C.; Nomura, D. K. Activity-Based Protein Profiling for Mapping and Pharmacologically Interrogating Proteome-Wide Ligandable Hotspots. *Curr. Opin. Biotechnol.* **2017**, *43*, 25–33.
(4) Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. Site-Selective Protein-Modification Chemistry for Basic Biology and Drug Development. *Nat. Chem.* **2015**, *8* (2), 103–113.
(5) Ma, X.; Zhao, Y. Biomedical Applications of Supramolecular Systems Based on Host–Guest Interactions. *Chem. Rev.* **2015**, *115* (15), 7794–7839.
(6) Hu, J.; Liu, S. Engineering Responsive Polymer Building Blocks with Host–Guest Molecular Recognition for Functional Applications. *Acc. Chem. Res.* **2014**, *47* (7), 2084–2095.
(7) Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115* (22), 12320–12406.
(8) Webber, M. J.; Langer, R. Drug Delivery by Supramolecular Design. *Chem. Soc. Rev.* **2017**, *46*, 6600–6620.
(9) Laitinen, O. H.; Nordlund, H. R.; Hytönen, V. P.; Kulomaa, M. S. Brave New (Strept)avidins in Biotechnology. *Trends Biotechnol.* **2007**, *25* (6), 269–277.
(10) Shetty, D.; Khedkar, J. K.; Park, K. M.; Kim, K. Can We Beat the Biotin–avidin Pair?: cucurbit[7]uril-Based Ultrahigh Affinity Host–guest Complexes and Their Applications. *Chem. Soc. Rev.* **2015**, *44* (23), 8747–8761.
(11) Nguyen, T.; Joshi, N. S.; Francis, M. B. An Affinity-Based Method for the Purification of Fluorescently-Labeled Biomolecules. *Bioconjug. Chem.* **2006**, *17* (4), 869–872.
(12) Chung, J. A.; Wollack, J. W.; Hovlid, M. L.; Okesli, A.; Chen, Y.; Mueller, J. D.; Distefano, M. D.; Taton, T. A. Purification of Prenylated Proteins by Affinity Chromatography on Cyclodextrin-Modified Agarose. *Anal. Biochem.* **2009**, *386* (1), 1–8.
(13) Kwant, R. L.; Jaffe, J.; Palmere, P. J.; Francis, M. B. Controlled Levels of Protein Modification through a Chromatography-Mediated Bioconjugation. *Chem. Sci.* **2015**, *6* (4), 2596–2601.
(14) Chinai, J. M.; Taylor, A. B.; Ryno, L. M.; Hargreaves, N. D.; Morris, C. A.; Hart, P. J.; Urbach, A. R. Molecular Recognition of Insulin by a Synthetic Receptor. *J. Am. Chem. Soc.* **2011**, *133* (23), 8810–8813.
(15) Lee, J. W.; Shin, M. H.; Mobley, W.; Urbach, A. R.; Kim, H. I. Supramolecular Enhancement of Protein Analysis via the Recognition of Phenylalanine with Cucurbit[7]uril. *J. Am. Chem. Soc.* **2015**, *137* (48), 15322–15329.
(16) Shaurya, A.; Dubicki, K. I.; Hof, F. Chemical Agents for Binding Post-Translationally Methylated Lysines and Arginines. *Supramol. Chem.* **2014**, *26* (7–8), 583–590.
(17) Hof, F. Host-Guest Chemistry That Directly Targets Lysine Methylation: Synthetic Host Molecules as Alternatives to Bio-Reagents. *Chem. Commun.* **2016**, *52*, 10093–10108.
(18) Lee, D.-W.; Park, K. M.; Banerjee, M.; Ha, S. H.; Lee, T.; Suh, K.; Paul, S.; Jung, H.; Kim, J.; Selvapalam, N.; et al. Supramolecular Fishing for Plasma Membrane Proteins Using an Ultrastable Synthetic Host–guest Binding Pair. *Nat. Chem.* **2011**, *3* (2), 154–159.
(19) Murray, J.; Sim, J.; Oh, K.; Sung, G.; Lee, A.; Shrinidhi, A.; Thirunarayanan, A.; Shetty, D.; Kim, K. Enrichment of Specifically Labeled Proteins by an Immobilized Host Molecule. *Angew. Chemie Int. Ed.* **2017**, *56* (9), 2395–2398.
(20) Ilha Hwang; Kangkyun Baek; Minseon Jung; Youngkook Kim; Kyeng Min Park; Don-

Wook Lee; Narayanan Selvapalam, A.; Kim, K. Noncovalent Immobilization of Proteins on a Solid Surface by Cucurbit[7]uril-Ferrocenemethylammonium Pair, a Potential Replacement of Biotin−Avidin Pair. *J. Am. Chem. Soc.* **2007**, *129* (14), 4170–4171.

(21) Uhlenheuer, D. A.; Wasserberg, D.; Haase, C.; Nguyen, H. D.; Schenkel, J. H.; Huskens, J.; Ravoo, B. J.; Jonkheijm, P.; Brunsveld, L. Directed Supramolecular Surface Assembly of SNAP-Tag Fusion Proteins. *Chem. - A Eur. J.* **2012**, *18* (22), 6788–6794.

(22) Young, J. F.; Nguyen, H. D.; Yang, L.; Huskens, J.; Jonkheijm, P.; Brunsveld, L. Strong and Reversible Monovalent Supramolecular Protein Immobilization. *ChemBioChem* **2010**, *11* (2), 180–183.

(23) Fragoso, A.; Caballero, J.; Almirall, E.; Villalonga, R.; Cao, R. Immobilization of Adamantane-Modified Cytochrome *c* at Electrode Surfaces through Supramolecular Interactions. *Langmuir* **2002**, *18* (13), 5051–5054.

(24) Rosen, C. B.; Kwant, R. L.; MacDonald, J. I.; Rao, M.; Francis, M. B. Capture and Recycling of Sortase A through Site-Specific Labeling with Lithocholic Acid. *Angew. Chemie Int. Ed.* **2016**, *55* (30), 8585–8589.

(25) Schmitt, J.; Hess, H.; Stunnenberg, H. G. Affinity Purification of Histidine-Tagged Proteins. *Mol. Biol. Rep.* **1993**, *18* (3), 223–230.

(26) Mero, A.; Ishino, T.; Chaiken, I.; Veronese, F. M.; Pasut, G. Multivalent and Flexible PEG-Nitrilotriacetic Acid Derivatives for Non-Covalent Protein Pegylation. *Pharm. Res.* **2011**, *28* (10), 2412–2421.

(27) Kapanidis, A. N.; Ebright, Y. W.; Ebright, R. H. Site-Specific Incorporation of Fluorescent Probes into Protein: Hexahistidine-Tag-Mediated Fluorescent Labeling with (Ni2+:Nitrilotriacetic Acid)n−Fluorochrome Conjugates. *J. Am. Chem. Soc.* **2001**, *123* (48), 12123–12125.

(28) Kim, T. H.; Swierczewska, M.; Oh, Y.; Kim, A.; Jo, D. G.; Park, J. H.; Byun, Y.; Sadegh-Nasseri, S.; Pomper, M. G.; Lee, K. C.; et al. Mix to Validate: A Facile, Reversible PEGylation for Fast Screening of Potential Therapeutic Proteins In Vivo. *Angew. Chemie Int. Ed.* **2013**, *52* (27), 6880–6884.

(29) Rosen, C. B.; Kodal, A. L. B.; Nielsen, J. S.; Schaffert, D. H.; Scavenius, C.; Okholm, A. H.; Voigt, N. V.; Enghild, J. J.; Kjems, J.; Tørring, T.; et al. Template-Directed Covalent Conjugation of DNA to Native Antibodies, Transferrin and Other Metal-Binding Proteins. *Nat. Chem.* **2014**, *6* (9), 804–809.

(30) Webber, M. J.; Appel, E. A.; Vinciguerra, B.; Cortinas, A. B.; Thapa, L. S.; Jhunjhunwala, S.; Isaacs, L.; Langer, R.; Anderson, D. G. Supramolecular PEGylation of Biopharmaceuticals. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (50), 14189–14194.

(31) Antonik, P. M.; Eissa, A. M.; Round, A. R.; Cameron, N. R.; Crowley, P. B. Noncovalent PEGylation via Lectin–Glycopolymer Interactions. *Biomacromolecules* **2016**, *17* (8), 2719–2725.

(32) Salmaso, S.; Bersani, S.; Scomparin, A.; Mastrotto, F.; Caliceti, P. Supramolecular Bioconjugates for Protein and Small Drug Delivery. *Isr. J. Chem.* **2010**, *50* (2), 160–174.

(33) Cao, L.; Shi, X.; Cui, Y.; Yang, W.; Chen, G.; Yuan, L.; Chen, H. Protein–polymer Conjugates Prepared via Host–guest Interactions: Effects of the Conjugation Site, Polymer Type and Molecular Weight on Protein Activity. *Polym. Chem.* **2016**, *7* (32), 5139–5146.

(34) Hirotsu, T.; Higashi, T.; Abu Hashim, I. I.; Misumi, S.; Wada, K.; Motoyama, K.; Arima, H. Self-Assembly PEGylation Retaining Activity (SPRA) Technology via a Host–Guest Interaction Surpassing Conventional PEGylation Methods of Proteins. *Mol. Pharm.* **2017**,

14 (2), 368–376.
(35) Gubeli, R. J.; Sonzini, S.; Podmore, A.; Ravn, P.; Scherman, O. A.; van der Walle, C. F. Selective, Non-Covalent Conjugation of Synthetic Peptides with Recombinant Proteins Mediated by Host–guest Chemistry. *Chem. Commun.* **2016**, *52* (22), 4235–4238.
(36) Biedermann, F.; Rauwald, U.; Zayed, J. M.; Scherman, O. A. A Supramolecular Route for Reversible Protein-Polymer Conjugation. *Chem. Sci.* **2011**, *2* (2), 279–286.
(37) Sonzini, S.; Marcozzi, A.; Gubeli, R. J.; van der Walle, C. F.; Ravn, P.; Herrmann, A.; Scherman, O. A. High Affinity Recognition of a Selected Amino Acid Epitope within a Protein by Cucurbit[8]uril Complexation. *Angew. Chemie Int. Ed.* **2016**, *55* (45), 14000–14004.
(38) de Vink, P. J.; Briels, J. M.; Schrader, T.; Milroy, L. G.; Brunsveld, L.; Ottmann, C. A Binary Bivalent Supramolecular Assembly Platform Based on Cucurbit[8]uril and Dimeric Adapter Protein 14-3-3. *Angew. Chemie Int. Ed.* **2017**, *56* (31), 8998–9002.
(39) Uhlenheuer, D. A.; Young, J. F.; Nguyen, H. D.; Scheepstra, M.; Brunsveld, L. Cucurbit[8]uril Induced Heterodimerization of Methylviologen and Naphthalene Functionalized Proteins. *Chem. Commun.* **2011**, *47* (24), 6798.
(40) Li, X.; Bai, Y.; Huang, Z.; Si, C.; Dong, Z.; Luo, Q.; Liu, J. A Highly Controllable Protein Self-Assembly System with Morphological Versatility Induced by Reengineered Host–guest Interactions. *Nanoscale* **2017**, *9*, 7991–7997.
(41) Hou, C.; Li, J.; Zhao, L.; Zhang, W.; Luo, Q.; Dong, Z.; Xu, J.; Liu, J. Construction of Protein Nanowires through Cucurbit[8]uril-Based Highly Specific Host-Guest Interactions: An Approach to the Assembly of Functional Proteins. *Angew. Chemie Int. Ed.* **2013**, *52* (21), 5590–5593.
(42) van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. Supramolecular Chemistry Targeting Proteins. *J. Am. Chem. Soc.* **2017**, *139* (40), 13960–13968.
(43) Hou, C.; Huang, Z.; Fang, Y.; Liu, J. Construction of Protein Assemblies by Host–guest Interactions with Cucurbiturils. *Org. Biomol. Chem.* **2017**, *15* (20), 4272–4281.
(44) Sakamoto, S.; Kudo, K. Supramolecular Control of Split-GFP Reassembly by Conjugation of β-Cyclodextrin and Coumarin Units. *J. Am. Chem. Soc.* **2008**, *130* (29), 9574–9582.
(45) Ghosh, S.; Isaacs, L. Biological Catalysis Regulated by Cucurbit[7]uril Molecular Containers. *J. Am. Chem. Soc.* **2010**, *132* (12), 4445–4454.
(46) Bruns, C. J.; Liu, H.; Francis, M. B. Near-Quantitative Aqueous Synthesis of Rotaxanes via Bioconjugation to Oligopeptides and Proteins. *J. Am. Chem. Soc.* **2016**, *138* (47), 15307–15310.
(47) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. π-Clamp-Mediated Cysteine Conjugation. *Nat. Chem.* **2015**, *8* (2), 120–128.
(48) Dai, P.; Zhang, C.; Welborn, M.; Shepherd, J. J.; Zhu, T.; Van Voorhis, T.; Pentelute, B. L. Salt Effect Accelerates Site-Selective Cysteine Bioconjugation. *ACS Cent. Sci.* **2016**, *2* (9), 637–646.
(49) Brewster, M. E.; Loftsson, T. Cyclodextrins as Pharmaceutical Solubilizers. *Adv. Drug Deliv. Rev.* **2007**, *59* (7), 645–666.
(50) Walker, S.; Oun, R.; McInnes, F. J.; Wheate, N. J. The Potential of Cucurbit[n]urils in Drug Delivery. *Isr. J. Chem.* **2011**, *51* (5–6), 616–624.
(51) Saleh, N.; Koner, A. L.; Nau, W. M. Activation and Stabilization of Drugs by Supramolecular P K_a Shifts: Drug-Delivery Applications Tailored for Cucurbiturils. *Angew. Chemie Int. Ed.* **2008**, *47* (29), 5398–5401.

(52) Appel, E. A.; Forster, R. A.; Rowland, M. J.; Scherman, O. A. The Control of Cargo Release from Physically Crosslinked Hydrogels by Crosslink Dynamics. *Biomaterials* **2014**, *35* (37), 9897–9903.
(53) Mann, J. L.; Yu, A. C.; Agmon, G.; Appel, E. A. Supramolecular Polymeric Biomaterials. *Biomater. Sci.* **2018**, *6* (1), 10–37.
(54) Ambrogio, M. W.; Thomas, C. R.; Zhao, Y.-L.; Zink, J. I.; Stoddart, J. F. Mechanized Silica Nanoparticles: A New Frontier in Theranostic Nanomedicine. *Acc. Chem. Res.* **2011**, *44* (10), 903–913.
(55) Kim, C.; Agasti, S. S.; Zhu, Z.; Isaacs, L.; Rotello, V. M. Recognition-Mediated Activation of Therapeutic Gold Nanoparticles inside Living Cells. *Nat. Chem.* **2010**, *2* (11), 962–966.
(56) Czapar, A. E.; Steinmetz, N. F. Plant Viruses and Bacteriophages for Drug Delivery in Medicine and Biotechnology. *Curr. Opin. Chem. Biol.* **2017**, *38*, 108–116.
(57) Dedeo, M. T.; Finley, D. T.; Francis, M. B. Viral Capsids as Self-Assembling Templates for New Materials. *Prog. Mol. Biol. Transl. Sci.* **2011**, *103*, 353–392.
(58) Wen, A. M.; Steinmetz, N. F. Design of Virus-Based Nanomaterials for Medicine, Biotechnology, and Energy. *Chem. Soc. Rev.* **2016**, *45* (15), 4074–4126.
(59) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537* (7620), 320–327.
(60) Butterfield, G. L.; Lajoie, M. J.; Gustafson, H. H.; Sellers, D. L.; Nattermann, U.; Ellis, D.; Bale, J. B.; Ke, S.; Lenz, G. H.; Yehdego, A.; et al. Evolution of a Designed Protein Assembly Encapsulating Its Own RNA Genome. *Nature* **2017**, *552* (7685), 415–420.
(61) Lin, Y.-R.; Koga, N.; Tatsumi-Koga, R.; Liu, G.; Clouser, A. F.; Montelione, G. T.; Baker, D. Control over Overall Shape and Size in de Novo Designed Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (40), E5478-85.
(62) Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C.; Huang, P.-S.; Ravichandran, R.; et al. Design of a Hyperstable 60-Subunit Protein Icosahedron. *Nature* **2016**, *535* (7610), 136–139.
(63) Bruns, C. J.; Stoddart, J. F. Rotaxane-Based Molecular Muscles. *Acc. Chem. Res.* **2014**, *47* (7), 2186–2199.
(64) Wikoff, W. R.; Liljas, L.; Duda, R. L.; Tsuruta, H.; Hendrix, R. W.; Johnson, J. E. Topologically Linked Protein Rings in the Bacteriophage HK97 Capsid. *Science* **2000**, *289* (5487), 2129–2133.

Chapter 2

Cucurbit[6]uril-promoted click chemistry for protein modification

ABSTRACT: Azide-alkyne cycloaddition is a powerful reaction for the formation of bioconjugates. When catalyzed by Cu(I) or strain promotion, this cycloaddition is considered to be a "click" reaction with many applications in chemical biology and materials science. We report a new type of azide-alkyne click chemistry for the synthesis of protein conjugates using cucurbit[6]uril (CB6) supramolecular chemistry. CB6-promoted azide-alkyne cycloaddition has been previously used for the synthesis of rotaxanes, but has not been applied to the development of complex bioconjugates. By developing new substrates for CB6 click that do not contain any cross-reactive functional groups and by optimizing reaction conditions, we converted CB6 click chemistry from a rotaxane synthesis tool into a useful bioconjugation technique. Using these new parameters, we synthesized a series of protein conjugates including protein-peptide, protein-DNA, protein-polymer, and protein-drug conjugates. We further demonstrated that CB6 click can be used in conjunction with strain-promoted azide-alkyne cycloaddition to generate distinct bioconjugates in protein mixtures. Development of new hydrogen-bonding substrates for the CB6 click reaction led to enhanced reaction rates and modification yields. CB6 click is a promising new reaction for the development of protein conjugates, and can be applied toward the synthesis of complex biomaterials for a wide range of applications.

Portions of the work described in this chapter have been reported in the following publication:
JA Finbloom, K Han, CC Slack, AL Furst, MB Francis. Cucurbit[6]uril-promoted click chemistry for protein modification. *J. Am. Chem. Soc.* 2017, *139;* 9691-9697.

2.1 Site-specific protein modification and cucurbit[6]uril-promoted click chemistry

Protein modification with synthetic molecules allows for the development of useful analytical platforms, therapeutic agents, and imaging tools with functions that cannot be achieved by the individual components alone.[1–3] The highly valuable set of chemical reactions used to functionalize biomolecules must take place in aqueous systems at near-neutral pH and with low concentrations of substrates.[1–3] Of these strategies, the highly exothermic Huisgen 1,3 dipolar azide-alkyne cycloaddition (AAC) has emerged as a particularly powerful reaction to couple two components together under the most demanding of circumstances.[4] When catalyzed by Cu(I), Cu-AAC is considered to be a "click" reaction,[5] owing to its speed of reactivity, high conversion rates, and virtually complete functional group tolerance.[5,6] Application of the copper catalyst also establishes the 1,4-substitution pattern of the triazole product.

Since its first report in 2002, Cu-AAC has been widely used in the fields of materials chemistry and chemical biology.[7] However, copper can perturb protein function, cause DNA damage, and is toxic to cells, rendering it less desirable for many biological applications.[7,8] As an alternative, Bertozzi and coworkers have developed a strain-promoted azide-alkyne cycloaddition (SP-AAC) that has suitable kinetics and reactivity for biological applications without the need for copper as a catalyst.[9] This reaction is typically performed between an azide installed on the biomolecule and a cyclooctyne on the synthetic substrate. This method overcomes many of the biological limitations of Cu-AAC, albeit through the use of substrates that can be limited by instability, reduced water solubility, and synthetic accessibility.[7,8]

An unexplored bioconjugation alternative for SP-AAC is the cucurbit[6]uril-promoted AAC (CB6-AAC), Figure 2.1. In this reaction, CB6 forms a heteroternary complex with propargylamine and azidoethylamine derivatives, whereby the protonated amine moieties form a hydrogen bonding network with the carbonyl portal of the CB6. This aligns the azide and alkyne components to facilitate triazole product formation. The concept of CB6-AAC was first reported by Mock and coworkers in 1983, and affords the 1,4-regioisomer of the triazole product selectively.[10–12]

Since Mock's discovery and subsequent kinetic studies, CB6-AAC has seen use for the synthesis of pseudorotaxanes and rotaxanes.[13,14] In particular, Stoddart and coworkers recently reported a cooperative capture method for the synthesis of CB6 rotaxanes, in which a cyclodextrin or pillarene host forms a hydrogen bonding network with the CB6.[14] This templating of substrates increases CB6 solubility in pure water and allows for quantitative rotaxane synthesis on the minutes to hours timescale.[14,15]

Many features of the CB6-AAC reaction suggest it will have the high functional group tolerance and water compatibility that bioconjugation strategies require. However, to date it has not been explored in the context of biomolecule modification. Herein we report the use of CB6 click chemistry as an efficient strategy for the synthesis of highly functionalized protein conjugates. We first adapted the reactive substrates to allow facile introduction on biomolecules, and optimized the conditions to produce protein conjugates in hours using relatively low equivalents of the synthetic coupling partners. When this reaction is performed with bulky substrates, protein-rotaxane conjugates are synthesized due to the formation of a mechanical bond[14] between the CB6

host and the triazole axle of the conjugate (Figure 2.1). Protein-rotaxane conjugates are a relatively unexplored class of protein-based materials, with our lab recently reporting on the first synthesis of such conjugates using thiol-maleimide chemistry.[16] Using this optimized CB6 click reaction, we synthesized a series of protein conjugates, ranging in substrate scope from small molecule and synthetic polymer coupling partners to biomolecules, such as DNA and peptides.

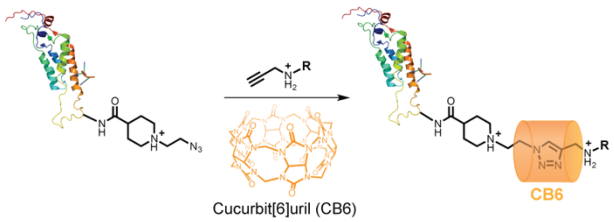


Figure 2.1. A protein bioconjugation strategy using CB6 click chemistry. An azidoethylpiperidine can be installed via NHS ester chemistry, providing a functional handle for subsequent modification via the CB6 click reaction.

2.2 Development of a CB6 click bioconjugation strategy.

In order to apply CB6 click for bioconjugation, new substrates had to be developed that did not contain cross-reactive amines. While azidoethylpyridinium substrates were previously used for CB6 click in the synthesis of rotaxanes,[14] they do not offer the synthetic modularity desired for bioconjugation. We therefore used previously unreported piperidine derivatives as the substrates. The synthesis of azidoethylpiperidine NHS ester **1** provided an azide CB6 click handle that can be installed onto proteins easily using NHS ester chemistry (Figure 2.2a). This would not be possible with the reactive secondary amines of azidoethylamine derivatives previously used in the synthesis of CB6 rotaxanes.

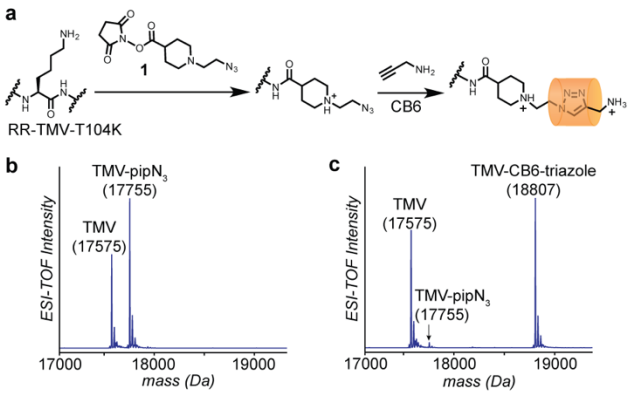


Figure 2.2. Development of CB6 click as a bioconjugation strategy. (a) NHS ester **1** can be reacted with free amines on proteins to install click handles. Following incubation with a propargylamine derivative and CB6, the protein-CB6-triazole product is formed. (b) RR-TMV-T104K double disks were used as a model protein for reaction development studies. TMV assemblies were modified to ~60% (20 additions per double disk) using 10 equiv of NHS ester **1**, as observed by LC/MS. (c) Virtually complete conversion to the triazole-CB6 product was observed after 24 h of incubation with 10 equiv of CB6 and propargylamine at 37 °C.

Optimization studies were performed on a recombinantly expressed, double arginine mutant of the tobacco mosaic virus capsid protein (RR-TMV) that assembles from 34 monomers into a stable disk-shaped nanomaterial. The specific protein used for these studies also had a T104K mutation that, after double disk assembly, provided modification sites on the protein surface. NHS ester **1** was conjugated to the RR-TMV-T104K coat proteins with ~60% conversion using 10 equiv of NHS ester (Figure 2.2b). We then monitored via LC/MS the disappearance of the TMV-pipN_3 peak and the emergence of a new TMV-CB6-triazole peak for subsequent CB6 click optimization studies. Incubation of TMV-pipN_3 with 10 equiv of CB6 and propargylamine at 37 °C in 50 mM BisTRIS, pH 6, led to approximately 95% conversion after 24 h, as measured by comparing the TMV-pipN_3 peak to the newly formed TMV-CB6-triazole peak (Figure 2.2c).

Interestingly, the observed mass included a CB6 molecule that was still associated with the protein, suggesting that the product existed as a stable pseudorotaxane. In line with this interpretation, subsequent incubation with a known CB6 guest (spermine) caused a mixture of TMV-CB6-triazole and TMV-triazole peaks to be observed by LC/MS (Figure S2.1). Although CB6 is retained on the protein conjugate product, CB6 is not anticipated to interfere with any downstream applications. Cucurbiturils have been used in a wide range of drug delivery and other biological applications, and no toxicity or unwanted downstream effects have been observed.[12] Importantly, no cyclodextrin or secondary supramolecular host other than CB6 was required in this reaction, in contrast to the cooperative capture method.[14] This may be due to the substantially increased CB6 solubility in buffer, relative to the pure water solvent that was used previously for rotaxane synthesis.

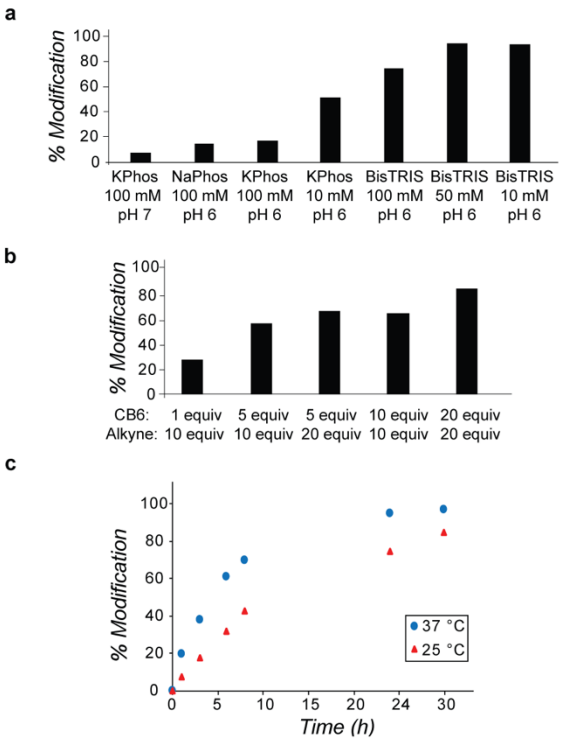


Figure 2.3. Optimization studies of CB6 click chemistry on TMV coat protein disks modified with an azidoethylpiperidine handle. (a) Buffer and pH conditions were screened for CB6 click reactions on TMV. All reactions were performed with 5 equiv of CB6 and 10 equiv of propargylamine for 24 h at 37 °C. (b) The reagent equivalents were screened for the CB6 click reaction on TMV. All reactions were carried out for 6 h at 37 °C in 50 mM BisTRIS, pH 6. (c) A time course study was performed for the CB6 click reaction on TMV. Both reactions were conducted with 5 equiv of CB6 and 10 equiv of alkyne in 50 mM BisTRIS, pH 6.

Studies of buffer conditions with CB6 click revealed BisTRIS to be an ideally suited buffer for this reaction, leading to near-quantitative conversion within 24 h when the components were incubated at 37 °C (Figure 2.3a). While the reaction does proceed to a degree in both sodium phosphate (NaPhos) and potassium phosphate (KPhos) buffers, the presence of large concentrations of small ions impedes the speed of the CB6 click reaction. Both Na^+ and K^+ are known to interact with the carbonyl portal of cucurbiturils.[17,18] Additionally, higher overall buffer strengths may lead to an inhibitory effect on the speed of the reaction, again due to unwanted interactions between the buffer ions and the CB6 ring. The reaction proceeds better at pH 6 vs 7, as lowering the pH increases the ratio of protonated to non-protonated amines on the substrates. These interactions are presumably essential for CB6 binding and subsequent triazole formation. A study of substrate equivalents revealed that between 5 and 20 equiv (50 – 200 μM) of propargylamine and CB6 were suitable to achieve good levels of conversion after 6 h (Figure 2.3b). A time course study was run at both 37 °C and ambient temperature. Near quantitative conversion was achieved at 37 °C, with high degrees of conversion (~85%) still being observed at ambient temperatures (Figure 2.3c). In all conditions tested, the CB6 click substrates displayed excellent solubility. This enhanced water solubility over the cyclooctyne SP-AAC substrates could lend to increased ease of protein modification.

2.3 PEGylation of proteins via CB6 click chemistry.

Polymer-protein conjugates are a class of materials that have seen increasing use in pharmaceuticals and drug delivery applications.[19,20] Polyethylene glycol (PEG) is the most commonly used polymer in the synthesis of these materials, and was therefore chosen as a model polymer substrate for protein conjugation via CB6 click chemistry. To facilitate substrate design for this reaction, propargylpiperidine NHS ester **2** was synthesized (Figure 2.4a). This substrate allows for the conversion of any amine-bearing small molecule or polymer into a reactive alkyne handle for CB6 click chemistry. Additionally, this substrate can be attached to proteins to install a propargyl CB6 click handle instead of the azide handle **1**. Lysozyme was reacted with NHS ester **1**, leading to 2-5 modifications of lysozyme with the azidoethylpiperidine handle (Figure 2.4c). Lysozyme-pipN_3 was then reacted with 10 equiv of CB6 and 10 equiv of PEG_{5k}-propargylpiperidine **3a** for 24 h at 37 °C. Approximately 34% PEGylation was observed via gel electrophoresis (Figure 2.4d). PEGylation was tested with other proteins and similarly modest levels of conversion were observed.

The limited modification may be due to the presence of two piperidine substrates, rather than the azidoethylpiperidine and the propargylamine substrates used in optimization experiments. To test this hypothesis, another PEG substrate **3b** was synthesized using commercially available propargylglycine (Figure 2.4a). Significantly enhanced modification of lysozyme was observed after incubation with CB6 and PEG **3b**, with near quantitative conversion to multiple PEGylation products closely resembling the mass spectrum ladder of the azidoethylpiperidine modification of lysozyme. These yield differences reflect a known trend in binding efficiencies of propargylamine derivatives for CB6.[14,21] When a single alkyl substitution occurs on the propargylamine derivative leading to a secondary amine substrate, only minimal perturbation in CB6 binding affinities is observed. However, when two alkyl substitutions occur, the subsequent tertiary amine displays over 1000-fold weaker binding affinities for CB6.[21] Therefore, CB6 click substrates containing the propargylglycine derivative, rather than the propargylpiperidine derivative, lead to excellent conversions to the protein-CB6-triazole products.

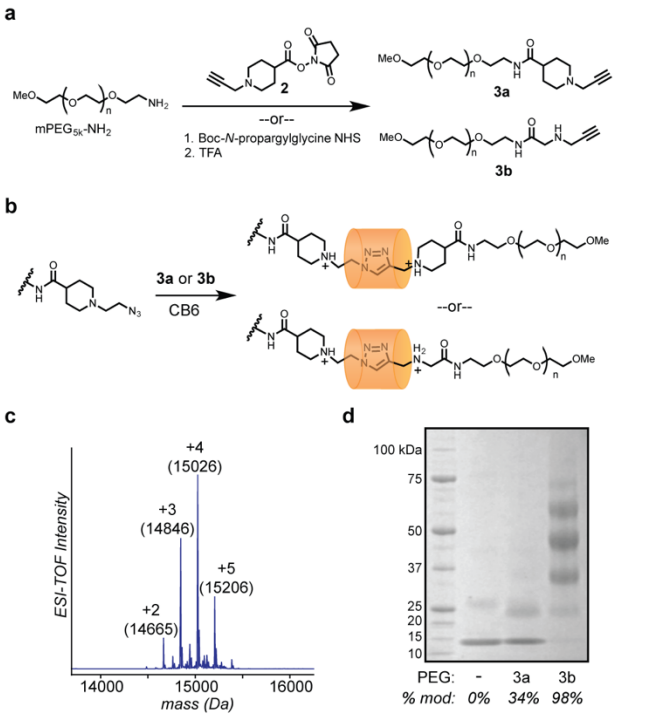


Figure 2.4. PEGylation of lysozyme via CB6 click. (a) Methoxy-terminated PEG_{5k}-amine was reacted with NHS ester **2** to form PEG_{5k}-propargylpiperidine **3a**, while modification with propargylglycine led to PEG_{5k}-propargylglycine **3b**. (b) Incubation of lysozyme-pipN_3 with **3a** or **3b** in the presence of CB6 leads to modification of the protein. (c) Lysozyme was treated with 10 equiv of NHS ester **1** to install an azide handle for subsequent CB6 click chemistry. (d) Incubation of lysozyme-pipN_3 with 10 equiv of CB6 and 10 equiv of **3a** or **3b** in 50 mM BisTRIS, pH 6 for 24 h at 37 °C led to ~34% PEGylation with **3a** and ~98% PEGylation with **3b**, as observed by gel electrophoresis.

2.4 Protein-peptide and protein-DNA conjugation via CB6 click chemistry

Bioconjugates that combine different classes of biomolecules are valuable in the fields of biomaterials and chemical biology. As such, it was important to determine that the CB6 click tolerates peptide and DNA substrates for protein conjugation. To do this, a heptameric peptide **4** containing an RGDS sequence was synthesized via solid phase peptide synthesis (Figure 2.5a). RGD peptides bind to the integrin receptors of cells, and are widely used for targeted delivery of chemotherapeutics.[22] *N*-propargylglycine was incorporated at the N terminus of peptide **4** as a reactive handle for CB6 click chemistry. Treatment of TMV-pipN$_3$ with 10 equiv of peptide **4** and 10 equiv of CB6 led to 70% conversion to the protein-peptide conjugate after 24 h at 37 ºC.

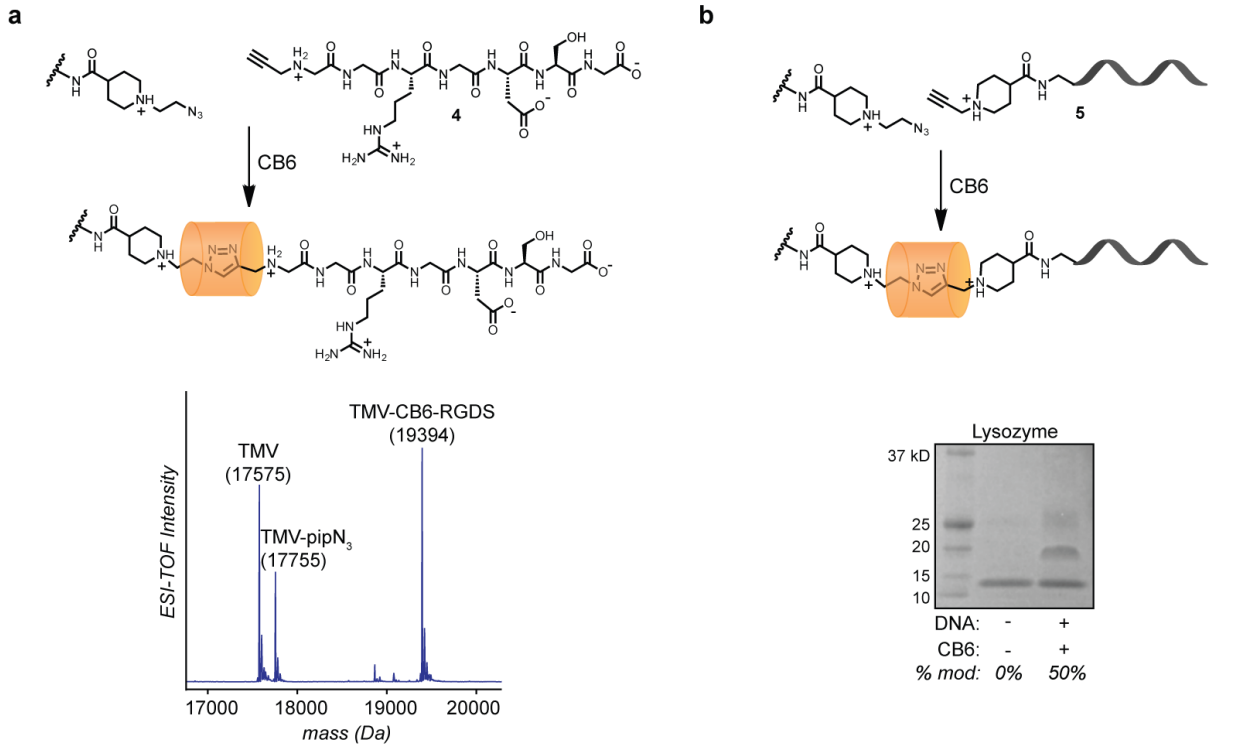


Figure 2.5. Using CB6-click to prepare protein-biomolecule conjugates. (a) An RGDS peptide **4** was synthesized, bearing a propargylglycine amino acid for CB6 click chemistry onto proteins. TMV-pipN$_3$ was incubated with 10 equiv of **4** and CB6 for 24 h at 37 °C. Approximately 70% modification was observed by LC/MS. (b) A 22 nt 5'-aminoDNA strand was reacted with NHS ester **2** to form DNA **5** for subsequent CB6 click chemistry onto proteins. Lysozyme-pipN$_3$ was incubated with 1 equiv of **5** and CB6 for 24 h at 37 °C. Approximately 50% modification was observed by gel electrophoresis.

Protein-DNA conjugates are used for many applications, including therapeutic targeted delivery with aptamers[23] and the immobilization of proteins through DNA complementation.[24] We therefore sought to determine if CB6 click would be a suitable reaction for protein-DNA conjugation. A 22-base oligomer ssDNA-amine bearing a binding motif for the TATA binding protein (TBP) transcription factor was reacted with propargylpiperidine-NHS **2** to yield DNA **5** with a reactive alkyne for CB6 click (Figure 2.5b). Lysozyme was modified with 1 equiv of NHS ester **1** to yield lysozyme-pipN$_3$ with predominantly a single azidopiperidine modification (Figure S2.2). Lysoyzme-pipN$_3$ was incubated with 1 equiv of **5** and 1 equiv of CB6. After 24 h at 37 ºC,

approximately 50% conversion to singly and doubly modified lysozyme-CB6-DNA was observed by gel electrophoresis (Figure 2.5b). When TMV was modified with DNA **5**, only modest modification was observed (Figure S3), likely due to the charge repulsion present between the DNA strands on the tightly packed TMV double disks. These results demonstrate that CB6 click chemistry can be used with a variety of substrates, ranging from small molecules and polymers, to complex biomolecules, such as peptides and DNA.

2.5 CB6 click for the development of protein-drug conjugates

Protein-drug conjugates are becoming increasingly important in the areas of medicine and chemical biology. For example, antibody-drug conjugates are used to target chemotherapeutic drugs to sites of interest.[25,26] Additionally, protein-based nanomaterials can be used as nanocarriers with potential applications in drug delivery.[27] Using CB6 click chemistry, we developed protein-drug conjugates using the RR-TMV double disk and doxorubicin (DOX) as a model chemotherapeutic. DOX-SS-propargylamine **7** was synthesized from DOX-SPDP **6** and *N*-propargylcysteamine to install a CB6 click handle onto DOX via a redox-cleavable disulfide linker (Figure 2.6a). Doxorubicin disulfide derivatives have been previously synthesized through reaction at the glycosamine of the drug, and the cleaved DOX-SH have been shown to retain activity against cancer cells.[28,29]

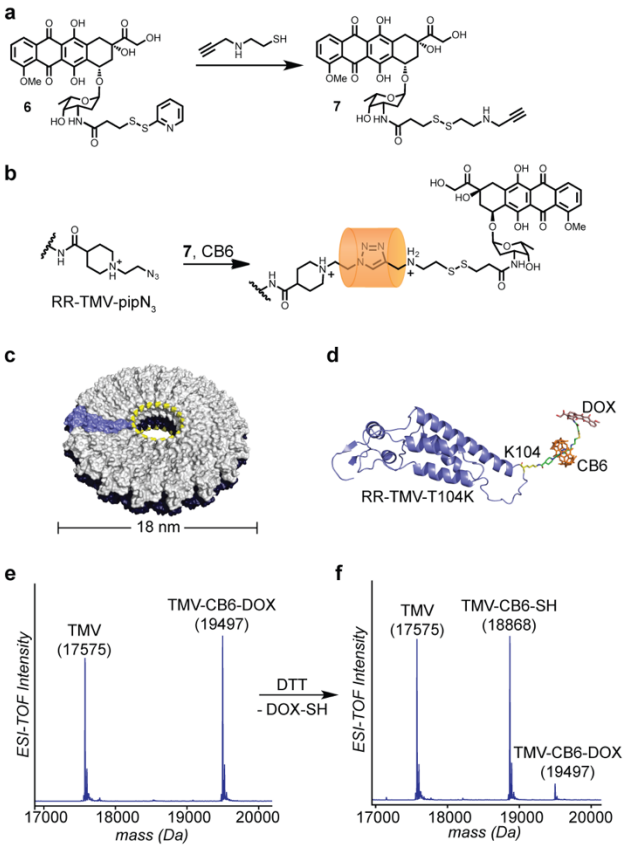


Figure 2.6. Synthesis of doxorubicin-protein conjugates via CB6 click chemistry. (a) DOX-SS-propargylamine **7** was synthesized to contain a CB6 click handle and a cleavable disulfide linker for drug release. (b) TMV double disks were modified with azidoethylpiperidine at the T104K sites, followed by subsequent modification with CB6 and **7**. (c) RR-TMV assembles into double disks composed of 34 monomers. (d) Model of a TMV-CB6-doxorubicin (DOX) monomer (based on PDB ID 1EI7). (e) Incubation of TMV-pipN$_3$ with 2 equiv of CB6 and **7** for 1 h at 37 °C led to 99% conversion to the TMV-CB6-DOX product. (f) Cleavage and release of ~90% DOX-SH was achieved after 2 h of incubation with 5 mM DTT.

DOX-SS-propargylamine **7** was reacted with TMV-pipN$_3$ and CB6 to synthesize TMV-CB6-DOX conjugates. Interestingly, the reaction proceeded to 99% conversion after 1 h incubation at 37 °C while only using 2 equiv of CB6 and **7** (Figure 2.6). Additionally, this reaction proceeded quantitatively in both BisTRIS and phosphate buffers. To test the modularity of this reaction, human serum albumin (HSA) was also conjugated to doxorubicin via CB6 click chemistry. HSA-doxorubicin conjugates are used clinically, as are other HSA-drug conjugates,[30,31] and therefore HSA was an important protein system to test with this chemistry. HSA-DOX conjugation via CB6

click also demonstrated the enhanced reactivity observed with the TMV-DOX conjugation (Figure S2.4). This enhancement in reactivity may be due to a hydrogen bond network formed between the DOX and the CB6, similar to the cooperative capture rotaxane synthesis using cyclodextrins or pillarenes and CB6.[14]

To determine if hydrogen bonding from the anthracene core of the DOX led to the rate enhancement, 3,4-dihydroxybenzyl propargylamine was synthesized and reacted with TMV-pipN_3. Interestingly, a 2-fold enhancement in modification was observed for this substrate when compared to propargylamine (Figure S2.5). It is therefore likely that the phenol groups on the DOX align with the CB6 carbonyl portal to form a hydrogen-bond network, leading to the observed rate enhancement.

The TMV-CB6-DOX conjugates were treated with DTT to cleave the disulfide linkage and release the DOX-SH payload. The release of ~90% DOX-SH was achieved after 2 h at ambient temperature, as monitored by LC/MS (Figure 2.6f). The CB6 click chemistry is therefore a particularly promising reaction for the synthesis of protein-conjugates, as the presence of the CB6 does not interfere with release of the chemotherapeutic cargo.

2.6 The orthogonality of CB6 click for protein modification

It is often important in bioconjugate synthesis to have multiple functional handles for two orthogonal reactions to occur. One of the most common bioconjugation techniques is thiol-maleimide coupling, and we therefore sought to confirm that the CB6 click reaction is orthogonal to reactive thiols. To test this, an RR-TMV mutant bearing a reactive cysteine at position 123 (123C) and the reactive lysine at position 104 (104K) was reacted with NHS ester **1** to install a pipN_3 handle for CB6 click (Figure S2.6). We have previously reported on the dual-modification of TMV disks with FRET dyes for light harvesting applications.[32,33] To make a FRET pair using the CB6 click and maleimide-thiol reactions, a propargylamine derivative of Alexa Fluor 488 was synthesized. *N*-propargylcysteamine was reacted with AF488-maleimide to form the AF488-propargylamine product **8** (Figure 2.7a). The propargylcysteamine derivative allows for facile installation of CB6 click handles onto any maleimide molecule of interest. The TMV disks were reacted with AF488-propargylamine **8** and CB6, followed by reaction with AF594-maleimide (Figure 2.7b). This modification strategy yielded TMV disks that were orthogonally labeled with AF488 donors at the 104K position and AF594 acceptors at the 123C position

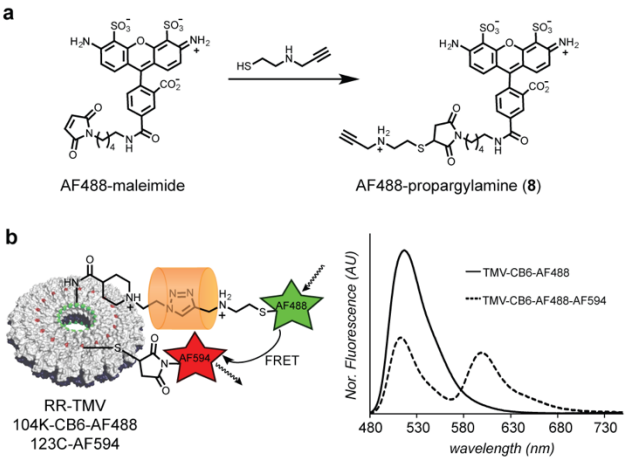


Figure 2.7. Development of FRET system using CB6 click chemistry. (a) Alexa Fluor 488 C5 maleimide was reacted with propargylcysteamine to form AF488-propargylamine **8**. (b) RR-TMV disks were modified with Alexa Fluor 488 (AF488) donors via CB6 click chemistry, followed by modification with Alexa Fluor 594 maleimide acceptors. Excitation at 450 nm led to a significant FRET response (54% efficiency) of the TMV-CB6-AF488-AF594 conjugates, as determined by comparing the emission spectra.

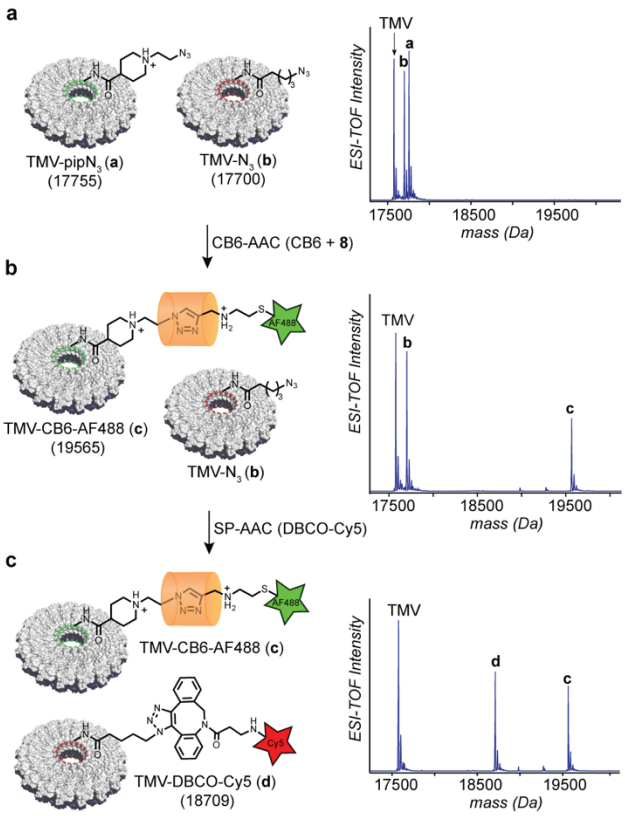


Figure 2.8. SP-AAC and CB6-AAC complementarity for the modification of protein mixtures. (a) Two TMV populations bearing pipN$_3$ or alkyl-N$_3$ handles were mixed together for modification via CB6-AAC followed by SP-AAC. (b) LC/MS spectra revealed specific modification of the azidopiperidine handles with CB6 and propargylamine-AF488 **8**, while no TMV-N$_3$ modification was observed. (c) Subsequent reaction of the TMV mixture with DBCO-Cy5 led to specific TMV-N$_3$ labeling via SP-AAC. Enlarged versions of the LC/MS traces appear in Figure S2.8.

(Figure S2.6). Fluorescence spectroscopy revealed a strong FRET response (54% FRET efficiency as measured by quenching of AF488 fluorescence) for the TMV disks upon excitation at 450 nm (Figure 2.7b).

We next sought to take advantage of the inherent substrate specificity of the CB6 click reaction to demonstrate its complementarity to the SP-AAC reaction. Since CB6 only binds to propargylamine and azidoethylamine derivatives, it should therefore be possible to conduct the CB6 click reaction in the presence of an alkyl azide on the protein surface. To test this, azidopentanoic acid NHS ester was synthesized[34] and reacted with TMV to install an alkyl azide handle (TMV-N$_3$, Figure S2.7). TMV-N$_3$ was then mixed 1:1 with TMV-pipN$_3$ and reacted with CB6 and **8** (Figure 2.8a). Complete conversion to the TMV-CB6-AF488 product was observed, while no side reactions between **8** and TMV-N$_3$ were observed (Figure 2.8b). Subsequent incubation with DBCO-Cy5 modified the TMV-N$_3$, forming two separate protein conjugates in one pot (Figure 2.8c). Incubation of the starting TMV mixture with DBCO-Cy5 alone led to modification of both TMV-pipN$_3$ and TMV-N$_3$ (Figure S2.9), as both of these substrates are reactive toward strained alkynes. Taken together, these results demonstrate that the CB6 click is a bioorthogonal reaction that offers complementarity to the SP-AAC for the separate modification of protein mixtures in solution.

2.7 Development of hydrogen-bonding substrates for CB6 click rate enhancement

While the CB6 click reaction allows for modification of proteins with relatively few equivalents of reagent, the rate of the reaction is relatively slow when compared to other bioconjugation reactions such as maleimide-thiol modification. We had observed that doxorubicin, as well as 3,4 dihydroxybenzyl derivatives displayed improved reaction rates, and we therefore designed new substrates for the reaction based on those findings. It had previously been observed that hydrogen bonding between CB6 and propargylamine derivatives can improve the reaction rates, and so a

series of phenol-propargylamine derivatives was synthesized to determine if the reaction rate was dependent on the position of the hydrogen-bonding hydroxyl group (Figure 2.9). The reactions were quenched via competitive CB6 binding with spermine at 3 h and the reaction conversions were analyzed by LC/MS. All the substrates tested displayed higher reaction rates than the propargylamine-benzene control, with the *para* derivative displaying the highest modification (95% conversion by 3 h).

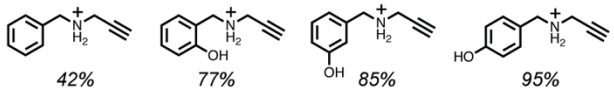


Figure 2.9. New hydrogen-bonding substrates for CB6 click chemistry. Three different hydroxybenzyl-propargylamine substrates were synthesized with the hydroxyl in the 2, 3, or 4 position. LC/MS analysis at 3 h incubation revealed that all hydroxyl-containing substrates performed better than the benzene-propargylamine control, with the para derivative performing the best.

To assess the effects of hydrogen bonding on the CB6 click reaction, computational searches for hydrogen-bonding interactions between a model triazole product onto a pipN_3 peptide and CB6 were performed. Hydrogen-bonding in the lowest energy conformation of the CB6 click product was only observed for the *ortho* derivative (Figure 5.10a). To ascertain why the *para* derivative led to such marked rate enhancement, a molecular dynamics simulation was carried out in the presence of two *para* derivatives and two CB6 hosts. In this simulation, the *para* substrates underwent pi-stacking and formed a 2:2 complex with the CB6 hosts. In this conformation, strong H-bonding was observed between the derivatives and the CB6 hosts. This finding was further observed in lowest energy conformational searches with a 2:2 ratio of substrate to CB6 (Figure 2.10b). While this phenomenon was also observed in the case of the *ortho* and *meta* derivative, it indicates a possible pathway for the rate enhancement observed experimentally.

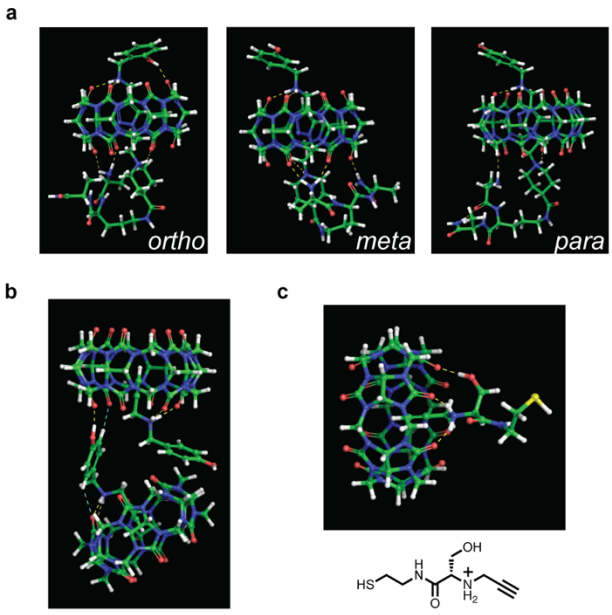


Figure 2.10. (a) Macromodel conformational search revealed strong hydrogen bonding for *ortho*-substituted propargylamine-phenol triazole products. No such hydrogen bonding was observed in the cases of *meta* and *para* derivatives. (b) A 2:2 complex was observed between the *para* derivative and CB6 using Macromodel computational searching. (c) Propargylamine-serine derivatives exhibited strong hydrogen-bonding to CB6 hosts as observed in Macromodel.

A series of hydroxyl-bearing derivatives that had not been tested experimentally was modeled to determine if any would exhibit strong hydrogen bonding with CB6. A serine derivative that also contained a thiol for maleimide functionalization displayed strong hydrogen bonding with CB6 in molecular dynamics simulations and in conformational searches (Figure 2.10c). In future work, this promising derivative will be synthesized and tested experimentally to confirm the computational findings.

2.8 Conclusions

The development of new reactions for the synthesis of protein conjugates is of paramount importance to the fields of chemical biology and biomedicine. We report here the use of CB6 to promote azide-alkyne cycloaddition for the modification of proteins. Using piperidine derivatives of propargylamine and azidoethylamine, CB6 click handles can be easily installed onto virtually any substrate of interest. The reaction itself is operationally simple and occurs under sufficiently mild conditions to maintain biomolecular structures. We tested this system with a variety of proteins and coupling partners, and demonstrated its utility for the synthesis of complex protein bioconjugates. The reaction is particularly effective for the installation of small molecules and peptides, and still provides good yields when large macromolecules are used. Importantly, this strategy can achieve modification of specific subsets of azides in the presence of others.

As with other click reactions, it is anticipated that further improvements in reactivity can be achieved through substrate modifications. The highly efficient coupling of doxorubicin to proteins suggests that this will be the case. Preliminary findings have indicated that phenol derivatives and serine-based derivatives might provide the rate enhancement needed for more advanced applications. Future work will also focus on the design and installation of unnatural amino acids bearing CB6 click handles. This development will facilitate site-selective protein modification, and expand CB6 click chemistry for *in vivo* applications in drug delivery and disease imaging.

2.9 Materials and methods

General Methods

All solvents and reagents, including cucurbit[6]uril hydrate ($CB6 \cdot XH_2O$), boc-*N*-(propargyl)glycine, methoxy-PEG_{5k}-NH_2, DNA-NH_2, DBCO-Cy5, serum albumin, and lysozyme were purchased from commercial suppliers and used without further purification. Ethyl 1-(2-chloroethyl)piperidine-4-carboxylate,[35] ethyl-1-(propargyl)piperidine-4-carboxylate,[36] trityl-cysteamine,[37] and azidopentanoic acid NHS ester,[S34] were prepared according to literature procedures. Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} (E. Merck) and visualized under a UV lamp at 254 nm. Column chromatography was carried out on silica gel 60 (E. Merck, 230–400 mesh).

Instrumentation and sample analysis

Nuclear magnetic resonance (NMR). All NMR spectra were recorded on Bruker Avance 600 spectrometers with working frequencies of 600 MHz for ^{1}H NMR, and 150 MHz for ^{13}C NMR. Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity, coupling constant (Hz), and integration. Data for ^{13}C NMR are reported in terms of chemical shift. Chemical shifts are referenced to the residual non-deuterated solvents for ^{1}H ($CDCl_3$: δ = 7.27 ppm, MeOD: δ = 3.31 ppm) and ^{13}C ($CDCl_3$: δ = 77.0 ppm, MeOD: δ = 49.0 ppm) nuclei.

Mass spectrometry. Peptides and protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an Agilent

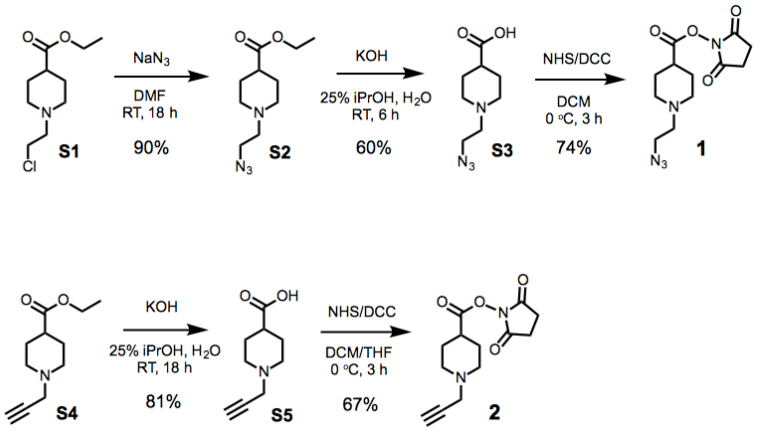
6224 Time-of Flight (TOF) LC/MS system equipped with a Turbospray ion source. Percent modification was determined through integration of MS peaks and comparing the drop in protein-azide peak integration to the increase in the click product peak. The integration of the completely unmodified protein peak served as an internal standard in determining the percent modification.

Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a Novex apparatus (Bio-Rad, Hercules, CA), using a 4-12% precast linear gradient polyacrylamide gel (Bio-Rad). All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 45-50 min at 200 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). ImageJ was used to determine the level of modification by optical densitometry.

Fluorescence Spectroscopy. Protein-dye conjugates were analyzed for FRET via fluorescence spectroscopy using a Fluoromax-2 spectrofluorometer (SA Instruments, Edison, NJ). Samples were normalized by AF488 absorbance at 488 nm prior to fluorescent analysis. Excitation at 450 nm provided efficient excitation of the AF488 donor while avoiding the direct excitation of the AF594 acceptor.

Computational Modeling. Macromodel was used to perform conformational searches on minimized structures of all the derivatives. Minimized structures were subjected to conformational search calculations using the Macromodel conformational search tool with the OPLS-2005 force field and mixed Low Mode (LMOD) and Monte Carlo Multiple Continuum (MCMM) method. Molecular dynamics simulations of each system were performed in orthorhombic 0.15 M NaCl water boxes with Cl^- to neutralize any charge. Systems were simulated for 10 ns using an NPT ensemble at 300 K and 1.013 bar in Desmond.

Synthetic Procedures



Scheme 2.1. Synthesis of piperidine substrates used for CB6 click chemistry.

Ethyl 1-(2-azidoethyl)piperidine-4-carboxylate (S2). A portion of sodium azide (195 mg, 3.0 mmol) was added to a solution of ethyl 1-(2-chloroethyl)piperidine-4-carboxylate[S1] **S1** (220 mg, 1.0 mmol) in DMF (5 mL). After stirring at ambient temperature for 18 h, the solvent was diluted with ddH_2O (15 mL) and extracted 3x with EtOAc. The organic layers were collected and washed with ddH_2O (2x) and brine (1x). The organic layers were then dried with Na_2SO_4 and the solvent was removed *in vacuo* to afford the product as a yellow oil (204 mg, 90%). ^{1}H NMR (600 MHz, $CDCl_3$) δ = 4.11 (q, *J* = 7.1 Hz, 2H), 3.31 (t, *J* = 6.1 Hz, 2H), 2.86 (d, *J* = 11.5 Hz, 2H), 2.55 (t, *J* = 6.1 Hz, 2H), 2.26 (t, *J* = 11.1 Hz, 1H), 2.09 (t, *J* = 12.2 Hz, 2H), 1.88 (d, *J* = 13.9 Hz, 2H), 1.82 – 1.70 (m, 2H), 1.22 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ = 174.86, 60.24, 57.28, 52.97, 48.35, 40.93, 29.63, 28.11, 14.15. HRMS (ESI-TOF-MS): *m/z* calc'd for $C_{10}H_{19}O_2N_4$ [*M* + H]$^+$ 227.1503, observed 227.1501.

1-(2-Azidoethyl)piperidine-4-carboxylic acid (S3). A portion of KOH (455 µL of 2 M solution, 0.91 mmol) was added to a solution of **S2** (113 mg, 0.50 mmol) in 25% iPrOH/H_2O (5 mL). After stirring at ambient temperature for 6 h, the solution was treated with 1 M HCl until pH 5.5 was reached. The iPrOH was removed *in vacuo* and the product was purified via a C_{18} Sep-Pak according to manufacturer's instructions. The solvent was removed *in vacuo* to afford the product as a tan solid (54 mg, 60%). ^{1}H NMR (600 MHz, MeOD) δ = 2.11 (t, *J* = 6.0 Hz, 2H), 1.76 (m, 2H), 1.46 (t, *J* = 5.9 Hz, 2H), 1.17 (t, *J* = 12 Hz, 2H), 0.83 (t, *J* = 10.6 Hz, 1H), 0.48 (d, *J* = 17.8 Hz, 2H), 0.33 (t, *J* = 12.6 Hz, 2H). ^{13}C NMR (150 MHz, MeOD) δ = 178.07, 55.64, 52.29, 46.25, 40.28, 26.67. *m/z* calc'd for $C_8H_{15}O_2N_4$ [*M* + H]$^+$ 199.1190, observed 199.1187.

1-(2-Azidoethyl)piperidine-4-N-hydroxysuccinimidyl ester (1). A solution of **S3** (40 mg, 0.20 mmol) in DCM (1.5 mL) was stirred at 0 °C for 5 min. A portion of *N*-hydroxysuccinimide (26 mg, 0.23 mmol) was added and the solution was stirred for an addition 5 min. Dicyclohexylcarbodiimide (46 mg, 0.23 mmol) was then added and the solution was stirred on ice for an additional 3 h. After 3 h, the solution was cooled at -20 °C and subjected to multiple rounds of purification via 0.2 µm syringe filtration to remove the urea byproduct. The solvent was removed *in vacuo* to afford the product as a tan solid (44 mg, 74%). ^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.38 (s, 2H), 2.92 (d, *J* = 12 Hz, 2H), 2.83 (s, 4H), 2.75-2.68 (m, 1H), 2.63 (s, 2H), 2.35-2.22 (m, 2H), 2.08 (s, 2H), 2.03 – 1.89 (m, 2H). ^{13}C NMR (150 MHz, $CDCl_3$) δ = 169.82, 168.99, 57.09, 52.18, 48.13, 33.86, 27.68, 25.56. *m/z* calc'd for $C_{12}H_{18}O_4N_5$ [*M* + H]$^+$ 296.1353, observed 296.1352.

1-(Propargyl)piperidine-4-carboxylic acid (S5). A portion of KOH (300 µL of a 2 M solution, 0.60 mmol) was added to a solution of ethyl 1-(propargyl)piperidine-4-carboxylate[S2] **S4** (65 mg, 0.33 mmol) in 25% iPrOH/H_2O (5 mL). After stirring at ambient temperature for 18 h, the solution was treated with 1 M HCl until pH 5.5 was reached. The iPrOH was removed *in vacuo* and the product was purified via a C_{18} Sep-Pak according to manufacturer's instructions. The solvent was removed *in vacuo* to afford the product as a colorless solid (45 mg, 81%). ^{1}H NMR (600 MHz, MeOD) δ = 3.51 (d, J= 2.4 Hz, 2H), 3.10 (d, *J* = 11.8 Hz, 2H), 2.87 (s, 1H), 2.56 (t, *J* = 11.0 Hz, 2H), 2.36 – 2.28 (m, 1H), 2.03-1.97 (m, 2H), 1.81 (q, *J* = 13.7 Hz, 2H). ^{13}C NMR (150 MHz, MeOD) δ = 177.74, 76.01, 75.25, 51.18, 45.83, 40.19, 27.22. *m/z* calc'd for $C_9H_{14}O_2N_1$ [*M* + H]$^+$ 168.1019, observed 168.1017.

1-(Propargyl)piperidine-4-N-hydroxysuccinimidyl ester (2). A solution of **S5** (30 mg, 0.18 mmol) in 25% DCM/THF (4 mL) was stirred at 0 °C for 5 min. A portion of *N*-hydroxysuccinimide (23 mg, 0.20 mmol) was added and the solution was stirred for an addition 5 min. Dicyclohexylcarbodiimide (41.2 mg, 0.20 mmol) was then added and the solution was stirred on ice for an additional 3 h. After 3 h, the solvent was removed *in vacuo* and the resulting residue was resuspended in DCM (2 mL) and cooled at -20 °C, followed by multiple rounds of purification using a 0.2 μm syringe filter to remove excess urea. The solvent was removed *in vacuo* to afford the product as a tan solid (32 mg, 67%). ^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.32 (d, *J* = 2.4 Hz, 2H), 2.88 (d, *J* = 11.7 Hz, 2H), 2.83 (s, 4H), 2.70 – 2.63 (m, 1H), 2.35 (t, *J* = 10.6 Hz, 2H), 2.25 (d, *J* = 2.4 Hz, 1H), 2.08 (d, *J* = 16.8 Hz, 2H), 2.02 – 1.87 (m, 2H). ^{13}C NMR (150 MHz, $CDCl_3$) δ = 169.99, 168.99, 73.22, 51.00, 47.04, 38.08, 33.92, 27.94, 25.56. *m/z* calc'd for $C_{13}H_{17}O_4N_2$ $[M + H]^+$ 265.1183, observed 265.1183.

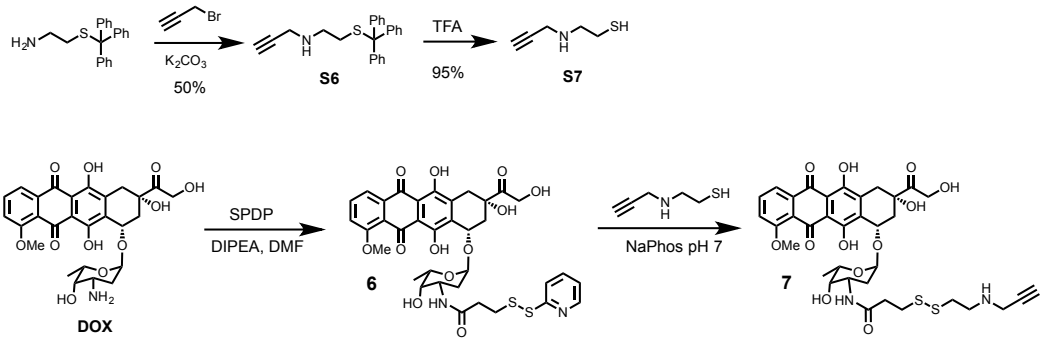
Propargylpiperidine-PEG$_{5k}$ (3a). To a solution of methoxy-terminated PEG_{5k}-NH_2 (10 mg, 0.002 mmol) in DCM (0.5 mL), was added **2** (2.85 mg, 0.01 mmol) and triethylamine (2.02 mg, 0.02 mmol). The reaction was stirred at ambient temperature for 6 h. After 6 h, a portion of the solution tested negative on a TLC plate via ninhydrin stain, indicating that the reaction was complete. The DCM was then removed under a stream of N_2, and the residue was resuspended in ddH_2O (1 mL) and subjected to multiple rounds of spin concentration using a 3000 MWCO spin concentrator. The resulting solution was lyophilized to afford **3a** as a colorless powder (5.4 mg, 53%).

N-(propargyl)glycine-PEG$_{5k}$ (3b). To a solution of boc-*N-(*propargyl)glycine (106.5 mg, 0.5 mmol) in DCM (2 mL), was added *N*-hydroxysuccinimide (64.4 mg, 0.56 mmol). The solution was stirred at 0 °C for 5 min. Dicyclohexylcarbodiimide (115.4 mg, 0.56 mmol) was then added and the solution was stirred on ice for an additional 4 h. After 3 h, the solvent was removed *in vacuo* and the resulting residue was resuspended in DCM (2 mL) and cooled at -20 °C, followed by purification using a 0.2 μm syringe filter to remove excess urea. The solvent was removed *in vacuo* and the NHS ester was used without further purification. To a solution of methoxy-terminated PEG_{5k}-NH_2 (10 mg, 0.002 mmol) in DCM (0.5 mL), was added the NHS ester of boc-*N*-(propargyl)glycine (3.1 mg, 0.01 mmol) and triethylamine (2.02 mg, 0.02 mmol). The reaction was stirred at ambient temperature for 6 h. After 6 h, a portion of the solution tested negative on a TLC plate via ninhydrin stain, indicating that the reaction was complete. The DCM was then removed under a stream of N_2, and the residue was resuspended in ddH_2O (1 mL), filtered using a 0.2 μm spin filter, and subjected to multiple rounds of spin concentration using a 3000 MWCO spin concentrator. The resulting solution was lyophilized to afford **3b** as a colorless powder (5.9 mg, 59%).

N-propargylGly-GRGDSG peptide (4). Peptide **4** was synthesized using standard Fmoc solid phase peptide synthesis (SPPS) on Gly-Wang resin (Anaspec, Fremont, CA). Deprotection of the Fmoc groups was performed with two 10 min incubations in a 30% v/v piperidine in dimethylformamide (DMF) solution. Couplings were carried out using 10 equiv of amino acid with 10 equiv of 2-(6-chloro-1-*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), and 20 equiv of *N*,*N*-diisopropylethylamine (DIPEA) in DMF for 20 min. The N terminus of the peptide was modified on resin with Boc-protected *N*-propargylglycine (ChemImpex, Wood Dale, IL) using the same coupling conditions. The peptide was cleaved from the resin after incubation with a solution of 95:2.5:2.5 TFA:TIPS:H_2O for 4 h. Excess TFA was removed under a stream of

N_2 and the peptide was precipitated in cold diethyl ether. The precipitate was redissolved in H_2O and purified by RP-HPLC, eluting in an aqueous gradient 5% to 95% MeCN/0.1% TFA in H_2O/0.1% TFA over 45 min at a flow rate of 3.0 mL/min. After lyophilization, peptide **4** was obtained as colorless powder (7.5 mg, 58%). HRMS (ESI-TOF-MS): *m/z* calcd for **4** $C_{24}H_{39}N_{10}O_{11}$ $[M+H]^+$ 643.27, observed 643.29.

Propargylpiperidine-DNA (5). 100 mM KPhos pH 8 (60 µL) was mixed with 150 nmol of DNA-amine (5'-[AmC3]-GGCGTCTATAAAGCGATCGCGA) dissolved in 60 µL of MilliQ water. DMF (75 µL) was then added, along with propargylpiperidine-NHS **2** (20 equiv) dissolved in 50 µL DMSO. The mixture was shaken overnight and purified by NAP-5 column (70% yield). MALDI-TOF MS spectra were acquired with 3-hydroxypicolinic acid matrix. HRMS (MALDI-TOF MS): calcd: 7071.0, observed: 7071.6.



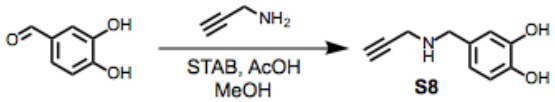
Scheme 2.2. Synthesis of DOX-SS-propargylamine **7**.

N-(2-(tritylthio)ethyl)propargylamine (S6). To a solution of trityl-protected cysteamine[S3] (480 mg, 1.5 mmol) in DMF (3 mL) was added K_2CO_3 (207 mg, 1.5 mmol) and a solution of propargyl bromide (58.5 mg, 0.5 mmol) in DMF (2 mL). The solution was stirred at ambient temperature for 2 h. The solvent was then removed *in vacuo* and the reaction mixture was reconstituted into ddH_2O and extracted with EtOAc (3x). The organic fractions were collected, dried with Na_2SO_4 and the product was then purified on silica gel with a gradient of EtOAc:hexanes (10-80% EtOAc) to afford the product as a colorless powder (90 mg, 50%). ^{1}H NMR (600 MHz, $CDCl_3$) δ = 7.42 (d, *J* = 8.6 Hz, 6H), 7.29 (d, *J* = 7.3 Hz, 6H), 7.21 (t, *J* = 7.3 Hz, 3H), 3.29 (s, 2H), 2.64 (t, *J* = 6.6 Hz, 2H), 2.36 (t, *J* = 6.6 Hz, 2H), 2.16 (s, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ = 144.79, 129.55, 127.83, 126.60, 81.83, 71.37, 66.53, 47.01, 37.73, 32.04. *m/z* calc'd for $C_{24}H_{24}NS$ $[M + H]^+$ 368.1626, observed 368.1624.

N-propargyl cysteamine (S7). Compound **S6** (30 mg, 0.08 mmol) was dissolved in a 95:2.5:2.5 (v/v) mixture of TFA/TIPS/H_2O (0.5 mL). The solution was stirred at ambient temperature for 30 min, and the solvent was removed under a stream of N_2. The resulting residue was resuspended in ddH_2O and extracted with Et_2O. The aqueous layer was collected and lyophilized to afford the product as a yellow oil (9 mg, 95%). ^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.86 (d, *J* = 2.5 Hz, 2H), 3.26 (t, *J* = 6.6 Hz, 2H), 2.89 (t, *J* = 2.5 Hz, 1H), 2.76 (t, *J* = 6.6 Hz, 2H). ^{13}C NMR (150 MHz, $CDCl_3$) δ = 78.22, 72.86, 48.89, 36.31, 19.76. *m/z* calc'd for C_5H_9NS $[M]^+$ 115.0456, observed 115.0458.

Doxorubicin-SPDP (6). To a solution of doxorubicin-HCl (2.9 mg, 0.005 mmol) in DMF (100 µL) was added DIPEA (1.29 mg, 0.01 mmol) and a solution of SPDP (2.3 mg, 0.0075 mmol) in DMF (100 µL). The reaction was vortexed at ambient temperature for 6 h, followed by purification via RP-HPLC, eluting in an aqueous gradient 10% to 90% MeCN/0.1% TFA in H_2O/0.1% TFA over 20 min at a flow rate of 3.0 mL/min. After lyophilization, **6** was obtained as a red powder (2.2 mg, 59%). HRMS (MALDI-TOF MS): *m/z* calc'd for $C_{35}H_{37}O_{12}N_2S_2$ $[M + H]^+$ 741.17, observed 740.23.

Doxorubicin-SS-propargylamine (7). Compound **6** (38 µL of a 50 mM stock in DMSO) was dissolved in a 50/50 mixture of MeCN and 100 mM NaPhos pH 7 (125 µL). Compound **S7** (22.7 µL of a 100 mM stock in NaPhos pH 7) was then added to the solution. The reaction mixture was gently vortexed at ambient temperature for 3 h, followed by purification via RP-HPLC eluting in an aqueous gradient 10% to 90% MeCN/0.1% TFA in H_2O/0.1% TFA over 20 min at a flow rate of 3.0 mL/min. After lyophelization, DOX-SS-propargylamine (**7**) was obtained as a red powder (0.7 mg, 50%). HRMS (MALDI-TOF MS): *m/z* calc'd for $C_{35}H_{41}O_{12}N_2S_2$ $[M + H]^+$ 745.20, observed 745.37.



Scheme S2.3. Synthesis of 3,4-Dihydroxybenzyl-propargylamine

3,4-Dihydroxybenzyl-propargylamine (S8). To a solution of 3,4-dihydroxybenzaldehyde (139 mg, 1.0 mmol) in dichloroethane (8 mL), was added a solution of propargylamine (82.5 mg, 1.5 mmol) in dichloroethane (2 mL). To the reaction mixture was added AcOH (91 mg, 1.5 mmol) and sodium triacetoxyborohydride (318 mg, 1.5 mmol). The reaction mixture was stirred at ambient temperature overnight, and the reaction was monitored by TLC. The reaction was quenched with saturated $NaHCO_3$ and extracted into EtOAc (3x). The organic fractions were collected, dried with Na_2SO_4 and the solvent was removed *in vacuo* to afford the product as a red powder (56 mg, 31%). ^{1}H NMR (600 MHz, MeOD) δ = 6.77 (s, 1H), 6.72 (d, *J* = 8.0 Hz, 1H), 6.65 (d, *J* = 8.0 Hz, 1H), 3.69 (s, 2H), 3.34 (d, *J* = 2.2 Hz, 2H), 2.66 (s, 1H). ^{13}C NMR (150 MHz, MeOD) δ = 144.89, 144.37, 129.54, 119.93, 115.61, 114.81, 80.24, 72.13, 50.94, 35.71. HRMS (ESI-TOF MS): *m/z* calc'd for $C_{10}H_{12}O_2N$ $[M + H]^+$ 178.0863, observed 178.0864.

Alexa Fluor 488 propargylamine (8). To a solution of AF488-maleimide (200 µL, 2.5 mM) in 10 mM NaPhos pH 7, was added a solution of *N*-(propargyl)-cysteamine **S7** as a 100 mM stock solution in 10 mM NaPhos pH 7, for a final concentration of 12.5 mM. The reaction mixture was vortexed at room temperature for 2 h and the reaction was monitored via MALDI-TOF MS. The product was then purified using NAP-5 columns (GE Healthcare) according to the manufacturer's instructions. The product was frozen and stored as a 310 µM stock solution in ddH_2O. HRMS (MALDI-TOF MS): *m/z* calc'd for $C_{35}H_{36}O_{12}N_5S_3$. $[M + H]^+$ 814.87, observed 814.54.

Protein modification

*General procedure for modification of proteins with azidoethylpiperidine-NHS **1***. To a solution of protein (150 μM) in 50 mM NaPhos buffer pH 8, 5-10 equiv of NHS ester **1** (100 mM stock solution in DMSO) was added. The solution was incubated for 2-4 h at ambient temperature. The solution was then spin concentrated 3-5 times into the appropriate buffer for subsequent modification.

General procedure for modification of proteins via CB6-promoted click chemistry. To a solution of azidoethylpiperidine modified protein (10-50 μM) in the reaction buffer was added 1-20 equiv of CB6 (500 μM stock solution in reaction buffer), followed by addition of 1-20 equiv of propargylamine or propargylamine derivative (1 mM stock in ddH_2O). The reaction was incubated at 37 °C and the conversion was monitored by LC/MS or gel electrophoresis. In some cases, the protein conjugate was purified by spin concentration 3-5 times into 10 mM NaPhos buffer, pH 7.

Doxorubicin modification of RR-TMV via CB6-promoted click chemistry. To 20 μL of a 10 μM protein solution (based on monomer) of azidoethylpiperidine-modified RR-TMV (SYS N-terminus, T104K mutation) in 50 mM BisTRIS pH 6 was added 2 equiv of CB6 (0.8 μL of 500 μM stock solution), followed by 2 equiv of **7** (0.8 μL of 500 μM stock solution). The reaction was incubated for 1 h at 37 °C, and the conversion was monitored by LC/MS.

Doxorubicin cleavage from RR-TMV. To a solution of TMV-CB6-DOX (10 μM based on protein monomer) in 10 mM NaPhos pH 7 was added DTT for a final DTT concentration of 5 mM. The solution was incubated at ambient temperature and the cleavage of DOX-SH was monitored by LC/MS.

Modification of RR-TMV with AF488-AF594 FRET pair. To a solution of azidoethylpiperidine-modified RR-TMV (SYS N-terminus, T104K S123C mutation, 17 μL of a 145 μM protein solution) in 50 mM BisTRIS pH 6 was added an additional 20 μL of reaction buffer. To the protein solution was added 5 equiv of CB6 (25 μL of 500 μM stock solution), followed by 5 equiv of **8** (40 μL of 310 μM stock solution). The reaction was incubated for 1.5 h at 37 °C. For the AF488 control, the TMV-CB6-AF488 product was isolated via multiple rounds of spin concentration using a 30,000 MWCO spin concentrator. For the FRET pair reaction, 1 equiv of AF594-maleimide was added and the reaction mixture was incubated at ambient temperature for 1 h, followed by purification via multiple rounds of spin concentration using a 30,000 MWCO spin concentrator.

CB6-AAC and SP-AAC Modification of TMV Mixture. A population of RR-TMV modified with NHS ester **1** (TMV-pipN_3), and a population of RR-TMV modified with azidopentanoic acid NHS ester (TMV-N_3) were mixed together to achieve equal concentrations of 10 μM of each protein in 50 mM BisTRIS pH 6 (50 μL total volume). To this mixture was added 5 equiv of CB6 (7 μL of 500 μM stock solution), followed by 5 equiv of **8** (11.3 μL of 310 μM stock solution). The reaction was incubated overnight at 37 °C. After overnight incubation, 5 equiv of DBCO-Cy5 was added and the reaction mixture was incubated at 37 °C for an additional 3 h. At each stage, the reaction was monitored via LC/MS.

2.10 References

(1) Witus, L. S.; Francis, M. B. Using Synthetically Modified Proteins to Make New Materials. *Acc. Chem. Res.* **2011**, *44* (9), 774–783.

(2) Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. Site-Selective Protein-Modification Chemistry for Basic Biology and Drug Development. *Nat. Chem.* **2015**, *8* (2), 103–113.

(3) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chemie Int. Ed.* **2009**, *48* (38), 6974–6998.

(4) Huisgen, R.; Szeimies, G.; Möbius, L. 1.3-Dipolare Cycloadditionen, XXXII. Kinetik Der Additionen Organischer Azide an CC-Mehrfachbindungen. *Chem. Ber.* **1967**, *100* (8), 2494–2507.

(5) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chemie Int. Ed.* **2002**, *41* (14), 2596–2599.

(6) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie Int. Ed.* **2001**, *40* (11), 2004–2021.

(7) McKay, C. S.; Finn, M. G. Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chem. Biol.* **2014**, *21* (9), 1075–1101.

(8) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. A Comparative Study of Bioorthogonal Reactions with Azides. *ACS Chem. Biol.* **2006**, *1* (10), 644–648.

(9) Codelli, J. A.; Baskin, J. M.; Agard, N. J.; Bertozzi, C. R. Second-Generation Difluorinated Cyclooctynes for Copper-Free Click Chemistry. *J. Am. Chem. Soc.* **2008**, *130* (34), 11486–11493.

(10) Mock, W. L.; Irra, T. A.; Wepsiec, J. P.; Manimaran, T. L. Cycloaddition Induced by Cucurbituril. A Case of Pauling Principle Catalysis. *J. Org. Chem.* **1983**, *48* (20), 3619–3620.

(11) Mock, W. L.; Irra, T. A.; Wepsiec, J. P.; Adhya, M. Catalysis by Cucurbituril. The Significance of Bound-Substrate Destabilization for Induced Triazole Formation. *J. Org. Chem.* **1989**, *54* (22), 5302–5308.

(12) Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115* (22), 12320–12406.

(13) Angelos, S.; Yang, Y.-W.; Patel, K.; Stoddart, J. F.; Zink, J. I. pH-Responsive Supramolecular Nanovalves Based on Cucurbit[6]uril Pseudorotaxanes. *Angew. Chem. Int. Ed.* **2008**, *47*, 2222–2226.

(14) Hou, X.; Ke, C.; Fraser Stoddart, J. Cooperative Capture Synthesis: Yet Another Playground for Copper-Free Click Chemistry. *Chem. Soc. Rev.* **2016**, *45* (14), 3766–3780.

(15) Ke, C.; Smaldone, R. A.; Kikuchi, T.; Li, H.; Davis, A. P.; Stoddart, J. F. Quantitative Emergence of Hetero[4]rotaxanes by Template-Directed Click Chemistry. *Angew. Chemie Int. Ed.* **2013**, *52* (1), 381–387.

(16) Bruns, C. J.; Liu, H.; Francis, M. B. Near-Quantitative Aqueous Synthesis of Rotaxanes via Bioconjugation to Oligopeptides and Proteins. *J. Am. Chem. Soc.* **2016**, *138* (47), 15307–15310.

(17) Buschmann, H.-J.; Cleve, E.; Schollmeyer, E. Cucurbituril as a Ligand for the Complexation of Cations in Aqueous Solutions. *Inorganica Chim. Acta* **1992**, *193* (1), 93–97.

(18) Masson, E.; Ling, X.; Joseph, R.; Kyeremeh-Mensah, L.; Lu, X. Cucurbituril Chemistry: A Tale of Supramolecular Success. *RSC Adv.* **2012**, *2*, 1213–1247.

(19) Webber, M. J.; Appel, E. A.; Vinciguerra, B.; Cortinas, A. B.; Thapa, L. S.; Jhunjhunwala, S.; Isaacs, L.; Langer, R.; Anderson, D. G. Supramolecular PEGylation of Biopharmaceuticals. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (50), 14189–14194.

(20) Peer, D.; Karp, J. M.; Hong, S.; Farokhzad, O. C.; Margalit, R.; Langer, R. Nanocarriers as an Emerging Platform for Cancer Therapy. *Nat. Nanotechnol.* **2007**, *2* (12), 751–760.

(21) Mock, W. L. Cucurbituril. In *Topics in Current Chemistry*; Springer Berlin Heidelberg, 1995; pp 1–24.

(22) Danhier, F.; Breton, A. Le; Préat, V. RGD-Based Strategies To Target Alpha(v) Beta(3) Integrin in Cancer Therapy and Diagnosis. *Mol. Pharm.* **2012**, *9* (11), 2961–2973.

(23) Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. Dual-Surface Modified Virus Capsids for Targeted Delivery of Photodynamic Agents to Cancer Cells. *ACS Nano* **2010**, *4* (10), 6014–6020.
(24) Palla, K. S.; Hurlburt, T. J.; Buyanin, A. M.; Somorjai, G. A.; Francis, M. B. Site-Selective Oxidative Coupling Reactions for the Attachment of Enzymes to Glass Surfaces through DNA-Directed Immobilization. *J. Am. Chem. Soc.* **2017**, *139* (5), 1967–1974.
(25) Schrama, D.; Reisfeld, R. A.; Becker, J. C. Antibody Targeted Drugs as Cancer Therapeutics. *Nat. Rev. Drug Discov.* **2006**, *5* (2), 147–159.
(26) Chari, R. V. J.; Miller, M. L.; Widdison, W. C. Antibody-Drug Conjugates: An Emerging Concept in Cancer Therapy. *Angew. Chemie Int. Ed.* **2014**, *53* (15), 3796–3827.
(27) Wen, A. M.; Steinmetz, N. F. Design of Virus-Based Nanomaterials for Medicine, Biotechnology, and Energy. *Chem. Soc. Rev.* **2016**, *45*, 4074–4126.
(28) Chen, H.; Wang, Z.; Zong, S.; Wu, L.; Chen, P.; Zhu, D.; Wang, C.; Xu, S.; Cui, Y. SERS-Fluorescence Monitored Drug Release of a Redox-Responsive Nanocarrier Based on Graphene Oxide in Tumor Cells. *ACS Appl. Mater. Interfaces* **2014**, *6* (20), 17526–17533.
(29) Xu, R.; Fisher, M.; Juliano, R. L. Targeted Albumin-Based Nanoparticles for Delivery of Amphipathic Drugs. *Bioconjug. Chem.* **2011**, *22* (5), 870–878.
(30) Elsadek, B.; Kratz, F. Impact of Albumin on Drug Delivery — New Applications on the Horizon. *J. Control. Release* **2012**, *157* (1), 4–28.
(31) Kratz, F. Albumin as a Drug Carrier: Design of Prodrugs, Drug Conjugates and Nanoparticles. *J. Control. Release* **2008**, *132* (3), 171–183.
(32) Dedeo, M. T.; Duderstadt, K. E.; Berger, J. M.; Francis, M. B. Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. *Nano Lett.* **2010**, *10* (1), 181–186.
(33) Miller, R. A.; Presley, A. D.; Francis, M. B. Self-Assembling Light-Harvesting Systems from Synthetically Modified Tobacco Mosaic Virus Coat Proteins. *J. Am. Chem. Soc.* **2007**, *129* (11), 3104–3109.
(34) Tae Seok Seo; Zengmin Li; Ruparel, H.; Ju, J. Click Chemistry to Construct Fluorescent Oligonucleotides for DNA Sequencing. *J. Org. Chem.* **2002**, *68* (2), 609–612.
(35) Lainé, D. I.; McCleland, B.; Thomas, S.; Neipp, C.; Underwood, B.; Dufour, J.; Widdowson, K. L.; Palovich, M. R.; Blaney, F. E.; Foley, J. J.; et al. Discovery of Novel 1-Azoniabicyclo[2.2.2]octane Muscarinic Acetylcholine Receptor Antagonists. *J. Med. Chem.* **2009**, *52* (8), 2493–2505.
(36) Cailly, T.; Dumas, N.; Millet, P.; Lemaître, S.; Fabis, F.; Charnay, Y.; Rault, S. Synthesis and Characterization of a Iodine-125-Labeled pyrrolo[1,2-a]thieno[3,2-E]pyrazine and Evaluation as a Potential 5-HT4R SPECT Tracer. *Eur. J. Med. Chem.* **2010**, *45* (11), 5465–5467.
(37) Maltese, M. Reductive Demercuration in Deprotection of Trityl Thioethers, Trityl Amines, and Trityl Ethers. *J. Org. Chem.* **2001**, *66*, 7615–7625.

2.11 Additional Figures

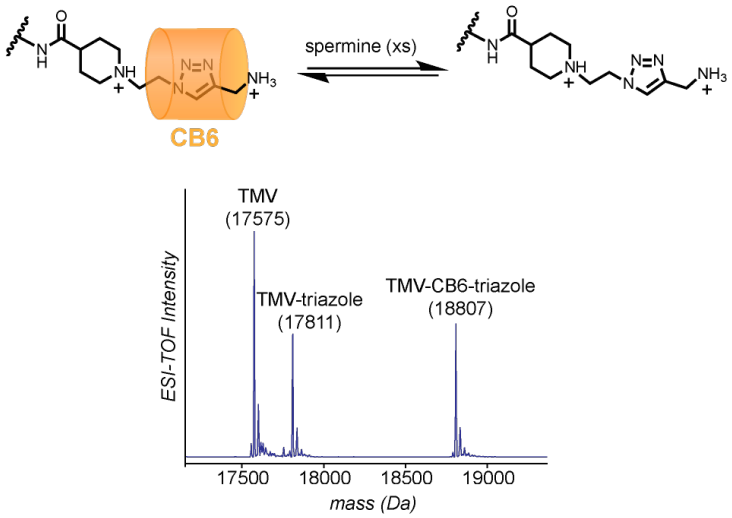


Figure S2.1. TMV-triazole exhibits reversible complexation by CB6. Incubation of 10 μM TMV-CB6-triazole with 1 mM spermine led to a mixture of TMV-triazole and TMV-CB6-triazole, as observed by LC/MS.

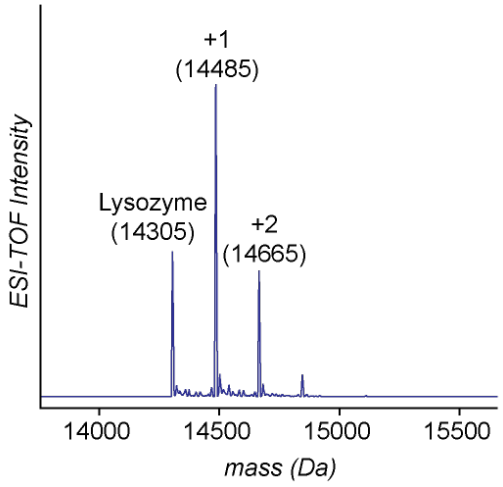


Figure S2.2. Lysozyme modification with NHS ester **1** for subsequent DNA modification. Lysozyme was modified with 1.05 equiv of NHS ester **1** for 2 h to afford predominantly 1 modification.

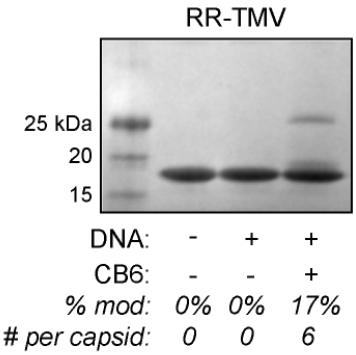


Figure S2.3. Modification of RR-TMV with DNA via CB6 click chemistry. TMV-pipN_3 was incubated with 10 equiv of CB6 and DNA **5** to afford 17% protein modification, corresponding to an average of 6 DNA strands per capsid. The modest modification percent was likely due to charge repulsions between the DNA strands on the capsids. Incubation of TMV-pipN_3 and DNA **5** without CB6 afforded no detectible modification.

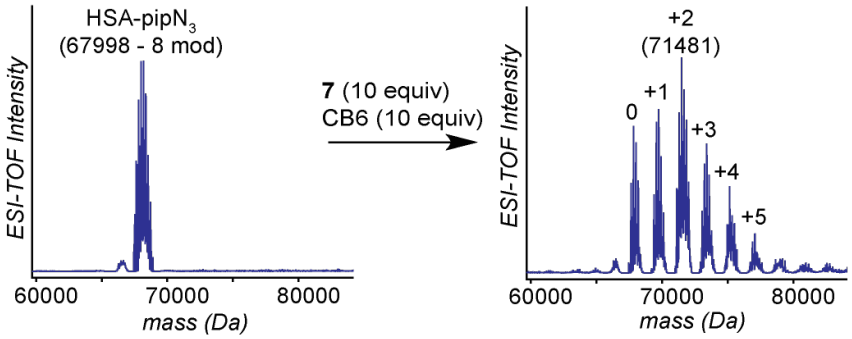


Figure S2.4. Synthesis of HSA-CB6-DOX conjugates. Human serum albumin (HSA) was incubated with 10 equiv of NHS ester **1** to achieve an average of 8 modifications (67998 Da). As there were an average of 8 modifications per HSA protein, HSA-pipN_3 was subsequently treated with 10 equiv of DOX-SS-propargylamine **7** and 10 equiv of CB6 for 1 h at 37 °C. Multiple modifications to HSA-CB6-DOX were observed, centered around 2 additions of DOX per protein.

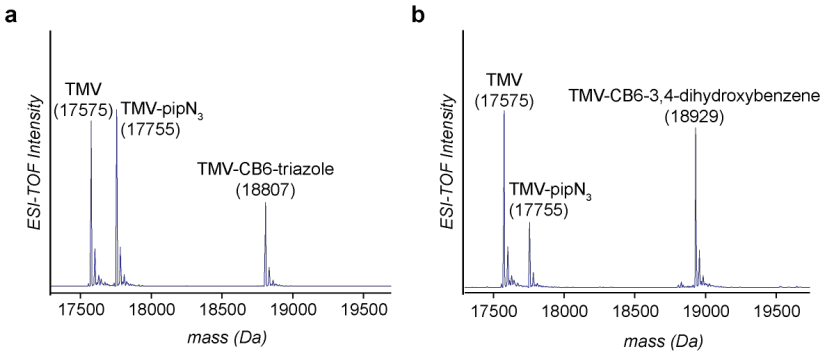


Figure S2.5. CB6 click rate enhancement by 3,4-dihydroxybenzyl-propargylamine. (a) Incubation of TMV-pipN$_3$ with 10 equiv of CB6 and propargylamine for 3 h at 37 °C. Approximately 35% modification was observed by LC/MS. (b) Incubation of TMV-pipN$_3$ with 10 equiv of CB6 and 3,4-dihydroxybenzyl-propargylamine **S8** for 3 h at 37 °C. Approximately 70% modification was observed by LC/MS.

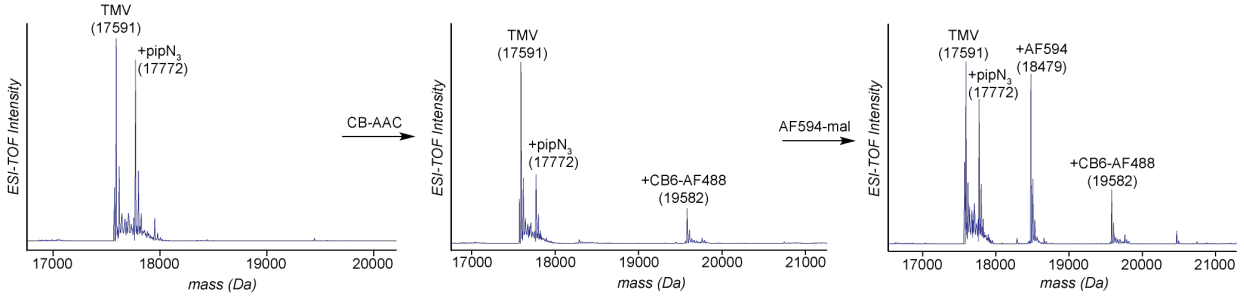


Figure S2.6. Modification of RR-TMV bearing reactive lysine 104K and reactive cysteine 123C to install a FRET pair. TMV-pipN$_3$ was incubated with 5 equiv of AF488-propargylamine **8** and CB6 for 1.5 h at 37 °C to afford 30% modification. Subsequent modification with 1 equiv of AF594-maleimide for 1 h afforded approximately 45% modification of previously unmodified TMV.

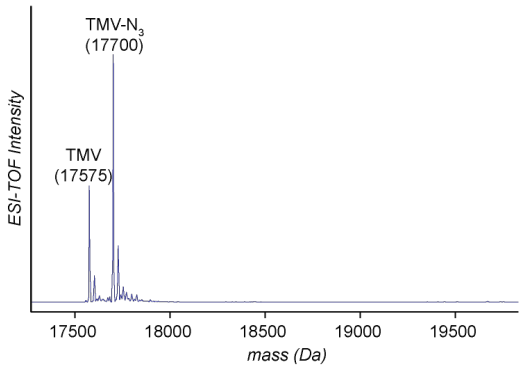


Figure S2.7. Installation of alkyl-azide onto TMV for SP-AAC. Incubation of TMV with 10 equiv of azidopentanoic acid NHS ester for 3 h led to 70% modification, as observed by LC/MS.

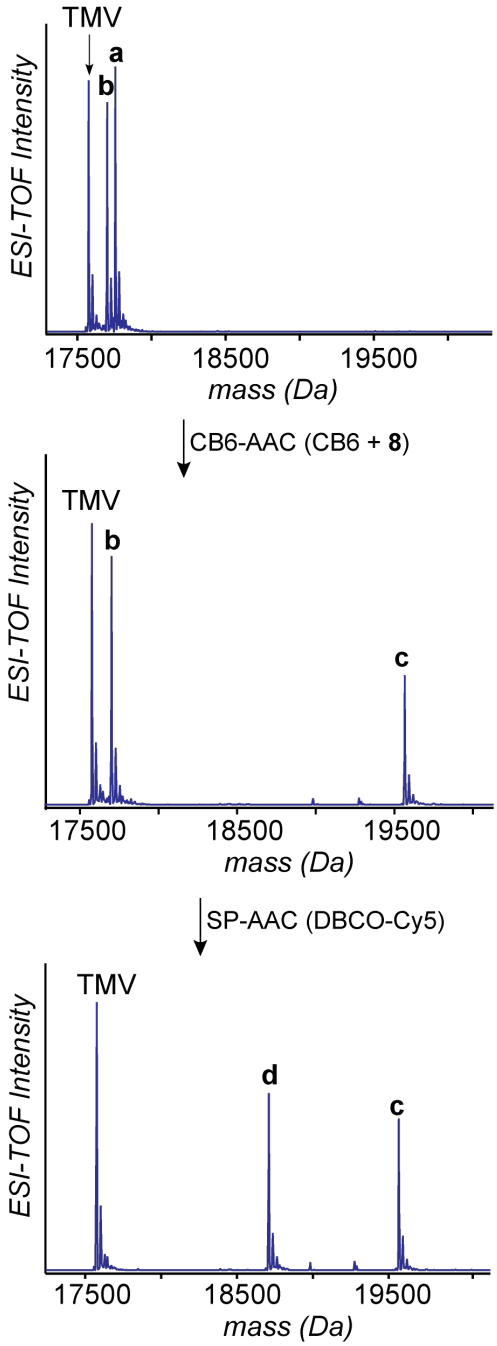


Figure S2.8. LC/MS traces for mixed click protein modification. Incubation of a TMV-N_3 (b) and TMV-pipN_3 (a) mixture with 5 equiv of AF488-propargylamine **8** for 12 h at 37 °C afforded complete modification of the TMV-pipN_3 to the TMV-CB6-AF488 product (c). No modification of TMV-N_3 was observed by LC/MS. Subsequent reaction with DBCO-Cy5 for 3 h led to conversion of TMV-N_3 to TMV-DBCO-Cy5 (d). These LC/MS traces also appears in Figure 2.8.

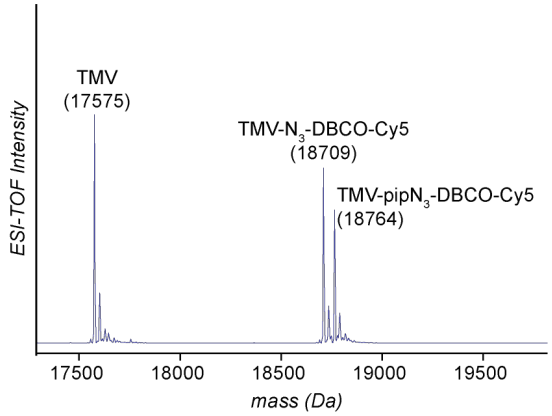


Figure S2.9. Incubation of a TMV-N_3 and TMV-pipN_3 mixture with 5 equiv of DBCO-Cy5 led to complete conversion of both TMV-N_3 and TMV-pipN_3 to the corresponding triazole-DBCO-Cy5 products.

Chapter 3

Cucurbiturils as transient supramolecular lysine and N-terminal protecting groups for peptide and protein modification

ABSTRACT: Protecting groups for native protein residues could offer an efficient and practical strategy to achieve greater homogeneity in protein conjugation, but as of yet have not been realized for wide-spread use. To that end, we report the use of cucurbituril macrocylces as supramolecular protecting groups for use in protein modification. Cucurbit[7]uril (CB7) was tested for N-terminal Phe protection, while cucurbit[6]uril (CB6) was tested for general lysine protection. Both CB7 and CB6 demonstrated efficient protection as observed by a decrease in overall acylation levels, and in a redirection to other nucleophilic sites in the case of CB7 protection. Dual CB6 and CB7 protection showed promise for insulin conjugation, as insulin contains an N-terminal Phe and a reactive lysine on its B chain, whereas the A chain of insulin only contains an N-terminal Gly. When incubated with both CB6 and CB7, redirection of acylation from the B chain to the A chain was observed for the PEGylation of insulin, switching from 1.5:1 B:A ratio without CB protection, to a 2.8:1 A:B with CB6 and CB7 protection. Cucurbiturils therefore show promise as transient protecting groups for use in the acylation of proteins with activated esters.

3.1 Protecting groups for protein and peptide modification

Protein modification relies on the conjugation of a molecule of interest to a specific amino acid residue on the biomolecule. One of the most common conjugation reactions involves the modification of lysine residues with activated esters.[1–3] While this strategy is convenient and reliable in modifying proteins, it is typically difficult to achieve lysine modification without N-terminal amine acylation, or to target a specific lysine over others.[1–3] Protecting groups for amino acids containing nucleophilic amines that are tolerated in biologically-relevant conditions are therefore desirable for protein modification applications. By specifically protecting lysines over protein N termini, activated esters could be used to modify the N terminus of a protein site-specifically. This would function as an alternative to N-terminal specific reactions that can require technical syntheses to attain the small molecule coupling partners necessary for the reactions. Alternatively, site-specific protection of N termini would afford specific modification of lysine residues without unwanted N-terminal acylation. To our knowledge, no such N-terminal protecting group has been reported, and while there are several reports of lysine protecting groups, they require the installation of noncanonical protected amino acids into the protein backbone.[4–6] It would therefore be desirable to have a modification strategy that involves the protection and subsequent deprotection of native amino acids for protein modification. Both the installation and deprotection of the protecting group would need to be carried out in biologically-relevant conditions that maintain the protein and peptide structure.

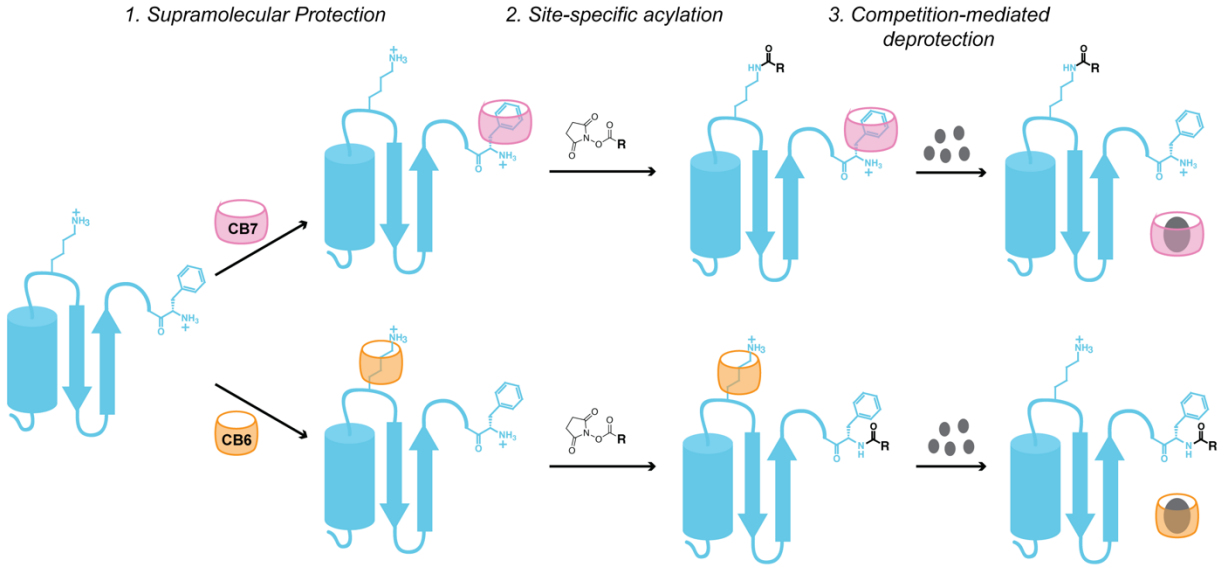


Figure 3.1. Cucurbiturils as supramolecular protecting groups for lysines and N termini. When bound to cucurbituril macrocyles, protein and peptide residues bearing nucleophilic amines are protected from modification from activated esters. Competition-mediated displacement of the CBs offers a deprotection mechanism for the bound residues.

While covalent protecting group strategies are traditionally used for amino acid protection, supramolecular interactions could offer a versatile platform for functional group protection. Host-guest chemistry targeting specific amino acids would allow for noncovalent protection and subsequent competition-mediated deprotection of nucleophilic amino acids. Cucurbiturils (CBs) are a particularly promising class of macrocycles for amino acid protection, as CBs are known to

interact with certain peptide sequences preferentially.[7–10] For CB binding to occur, a protonated amine group binds to the CB carbonyl portal, while a hydrophobic region binds to the inner cavity of the macrocycle (Figure 3.1). Of the CBs, cucurbit[7]uril (CB7) has been the most extensively studied for its protein interactions.[7,9–12] Specifically, CB7 is capable of binding to N-terminal phenylalanine residues with high affinity ($K_a = 1x10^6\ M^{-1}$).[13,14] CB6 is also known to interact with proteins, albeit to a lesser extent. CB6 interactions with lysine side chains have been observed, with binding affinities reported on the order of $10^2\ M^{-1}$.[8] We sought to use CB-protein interactions as supramolecular protecting groups for nucleophilic amines. We hypothesized that CBs could function as protecting groups for lysines and N terminal residues by increasing steric hindrance to activated esters upon CB binding to amino acids. Additionally, for amines to bind to CB portals, the amine must be protonated and therefore would have decreased nucleophilicity toward activated esters. Cucurbit[7]uril was tested as an N-terminal protecting group that could facilitate selective lysine modification, while cucurbit[6]uril was tested for lysine protection to facilitate N-terminal modification with activated esters (Figure 3.1).

3.2 Cucurbiturils as lysine and N-terminal protecting groups on a model peptide

To evaluate whether cucurbiturils could be used as protecting groups, an FGGSWKG peptide was synthesized via solid phase peptide synthesis. FGGSWKG functions as a model peptide for selective CB protection, as it contains both an N-terminal phenylalanine for CB7 protection, and

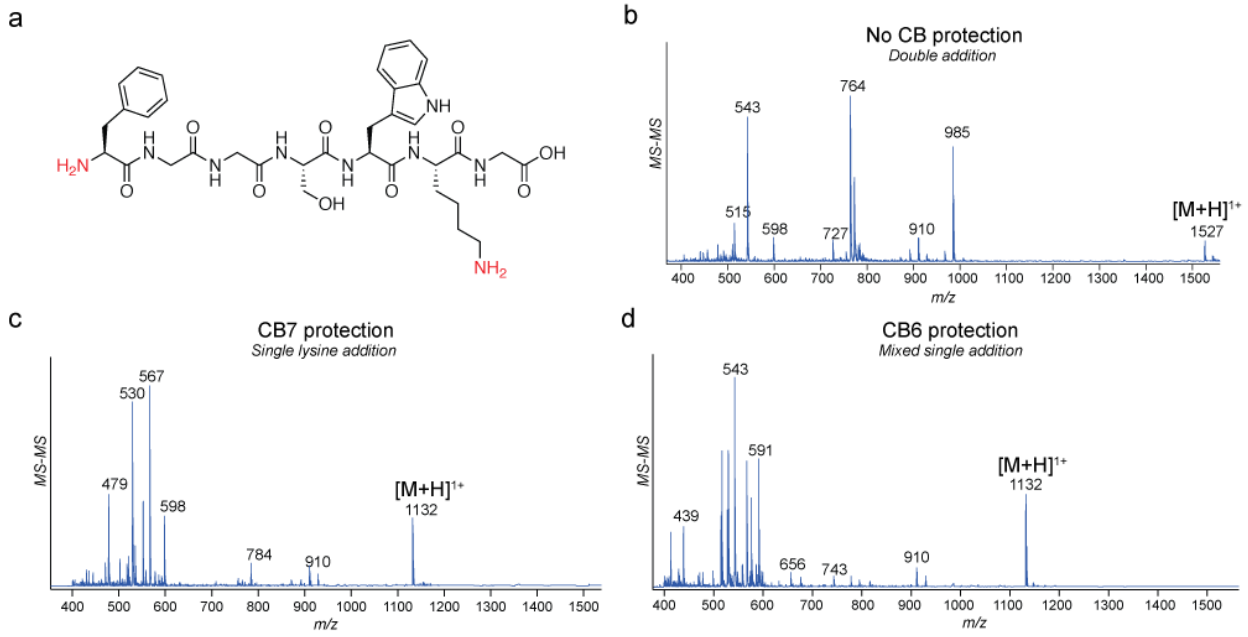


Figure 3.2. Evaluation of CB protecting groups on a model peptide. (a) An FGGSWKG peptide was synthesized by solid phase peptide synthesis and tested for CB protection against activated esters. (b) LC/MS-MS spectra of FGGSWKG modified with 4 equiv methoxy-PEG_8-NHS without CB7 and (c) with 10 equiv of CB7 protection. With CB7 protection only single modification was observed, predominantly at the lysine residue. FGGSWK(PEG)G fragment y_6^{2+} was observed at 530 m/z, and fragment x_3^{2+} was observed at 479 m/z. Fragment y_5 and y_6 were observed at 784 and 598 m/z, respectively. (d) Upon incubation with CB6, modest protection was achieved. Second addition was observed in the LC trace but to a lesser extent than without CB protection. NHS ester modification of both the lysine and the N-terminal Phe were observed in the single addition product. The peak at 591 m/z corresponds to y_2 fragment of F(PEG)GGSWKG, and 543 m/z corresponds to the b_1 fragment. See Supporting Figure S3.2 for LC traces of all three reactions.

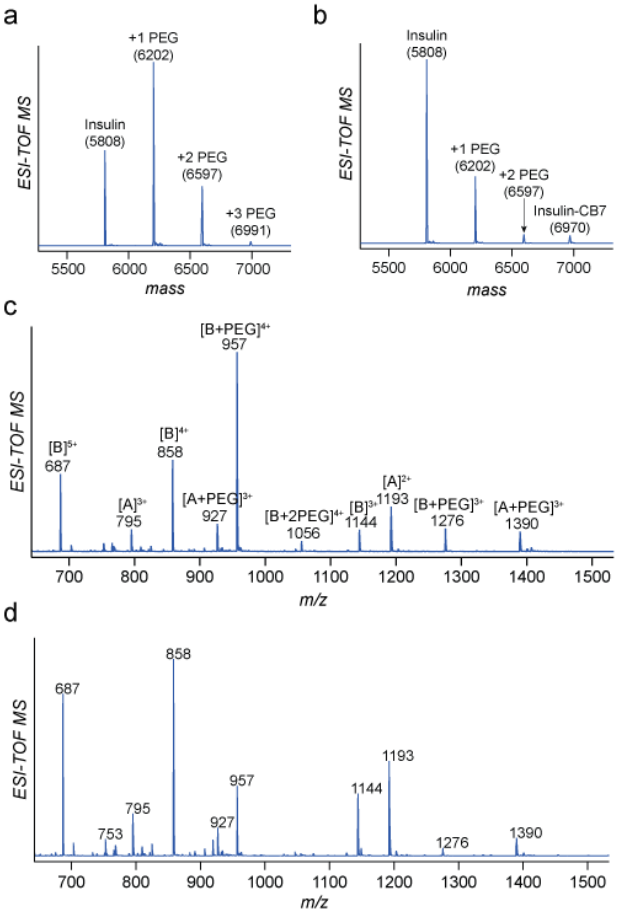


Figure 3.3. N-terminal protection of insulin with CB7. (a) Insulin was modified with 2.5 equiv of PEG_8-NHS ester leading to 2-4 modifications. (b) After protection with 3 equiv of CB7, PEG_8 modification was lowered to predominantly single addition. (c) DTT cleavage of insulin that had not received CB7 protection revealed single modification on the A chain, as well as single and double modification on the B chain. (d) After DTT cleavage of CB7-protected insulin, only single modification of A and B chains were observed. Approximately 35% reduction in B chain modification was observed, whereas approximately 10% reduction in A chain modification was observed.

a lysine for CB6 protection. To evaluate the use of CBs as protecting groups, FGGSWKG was incubated with cucurbituril to facilitate CB protection, followed by incubation with a PEG_8-NHS ester as a model activated ester. After incubation, the crude reaction was monitored by LC/MS-MS to determine protection efficiency (Figure 3.2). Modification with PEG_8-NHS and no CB incubation led to double modification of the peptide as determined by LC/MS-MS (Figure 3.2b). When CB7 was incubated with the peptide before subsequent NHS ester modification, no double modification was observed, with predominant single addition at the lysine residue (Figure 3.2c). Thus, CB7 afforded complete N-terminal protection of the FGGSWKG peptide, leading to near-selective lysine acylation with PEG_8-NHS.

When CB6 was used for lysine protection, only modest protection was observed by MS-MS. While some double modification was observed, it was significantly less than that observed with no CB incubation (Figure S3.2). Additionally, LC/MS-MS spectra corresponding to singly modified peptide indicated modification at both lysine and N-terminal residues (Figure 3.3d). This modest protection is likely due to the weaker binding affinity of CB6 for lysine residues when compared to CB7 affinity for N-terminal phenylalanine residues.[8] It is also known that CB6 is more sensitive to buffer conditions than CB7, and additional optimization of reaction conditions could afford enhanced lysine protection levels with CB6.

3.3 Cucurbit[7]uril as an N-terminal protecting group for insulin bioconjugation

Insulin is one of the most used biopharmaceuticals on the market today. Improving insulin delivery and efficacy if therefore of critical importance in the field of biomedicine. Previous attempts at PEGylating insulin have yielded conjugates with increased half-lives and biodistribution, thereby increasing their *in vivo* activity.[15,16] Importantly, high degrees of PEGylation of insulin at multiple sites can lead to a decrease in bioactivity, and conjugation of certain synthetic small molecules onto the more reactive B chain N terminus have likewise led to

a decrease in insulin activity.[15,17,18] Thus, more selective conjugation of insulin could lead to enhanced uptake, bioavailability, and efficacy of insulin conjugates for diabetes treatment. It would therefore be advantageous if NHS esters, which are widely used small molecules for bioconjugation, could be used to generate insulin bioconjugates with cargo attachments at specific sites on the protein surface.

Insulin contains an N-terminal phenylalanine on the insulin B chain, which is known to associate strongly with cucurbit[7]uril ($K_a = 10^6\ M^{-1}$).[13] Therefore, insulin is an ideal candidate for CB7 protection, as protection with CB7 could afford insulin conjugates with modifications primarily located at the Lys29 of the B chain and at the A chain N terminus. Incubation of insulin with PEG_8-NHS ester led to complete acylation of insulin, with a protein modification ladder centered around double modification (Figure 3.3a). When insulin was incubated with excess CB7 prior to treatment with PEG_8-NHS ester, single modification was dominant with a minor amount of double acylation. Additionally, a peak at 6970 Da was observed in the mass spectrum, corresponding to complexation between the native protein and CB7 (Figure 3.3b). This observed decrease in reactivity with CB7 protection is possibly due to the protection of the B chain N terminus, which is known to be the most reactive nucleophilic amine of the protein.[15,18] To determine the sites of acylation, insulin was reduced with DTT to separate the A and B chains, followed by analysis via LC/MS. Without CB7 protection, modification with PEG_8-NHS led to similar levels of modification at both the A and B chains, with double acylation observed on the B chain of insulin, which corresponds to modifications at both the N terminus and Lys29 (Figure 3.3c). When protected via CB7-Phe binding, higher degrees of modification at the A chain were observed, as were MS peaks corresponding to a single modification at the B chain, likely at the Lys29 position (Figure 3.3d). Future work will focus on performing an enzymatic digest of the modified insulin to confirm that B chain modification was indeed at the Lys29 position.

To evaluate further the potential of CB7 to function as an N-terminal protecting group for insulin, a reaction conditions screen was performed to determine the best CB and NHS ester equivalents for complete N-terminal protection while also allowing the most reactivity at the other nucleophilic amines present on insulin (Figure 3.4). Insulin was incubated with varying equivalents of CB7 prior to addition of either 2.5 or 10 equiv of PEG_8-NHS. Without any CB7 protection, near quantitative modification of insulin with a majority double addition was observed with 10 equiv of NHS ester. However, upon addition of CB7, protection was observed, affording majority single addition to the A chain of insulin, with some B chain addition. While CB7 protection of the most-reactive residue may account for some of the decrease in modification, the strong redirection of modification to the A chain N terminus, rather than to the B chain lysine, indicates that at high concentrations, CB7 may be loosely bound to Lys29 of insulin in addition to N-terminal Phe, causing protection at both nucleophilic B chain

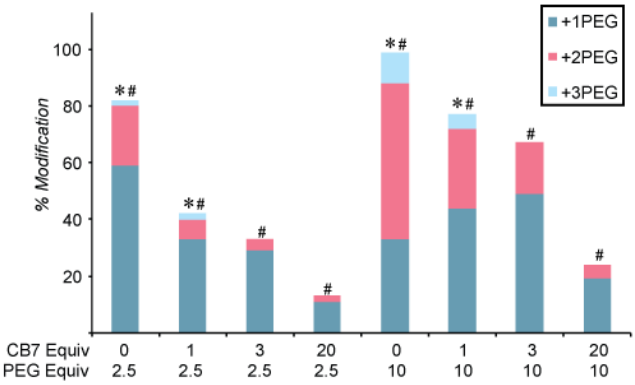


Figure 3.4. Equivalents screen for CB protection of insulin N terminus. Insulin was incubated with varying equivalents of CB7 followed by incubation with 2.5 or 10 equiv of PEG_8-NHS. Percent modification was determined by LC/MS analysis of intact and DTT-reduced insulin. #A chain modification is dominant. *B chain modification is dominant. *#Both A and B chain modifications were observed to similar degrees.

amines. In comparison, CB7-Gly interactions, such as those that may occur at the A chain N-terminal Gly, are much weaker than either CB7-Lys or CB7-Phe.[8] We would therefore hypothesize that negligible protection of the A chain N terminus would occur, which corresponds to the results from the equivalents screen. CB7 protection of insulin therefore causes redirection of NHS ester modification from the B chain to the A chain when higher equivalents of CB7 are used. This phenomenon was observed with both high and low equivalents of NHS ester. Thus, CB7 protection of insulin offers the potential to redirect NHS modification from the B chain of insulin to the A chain via strong N-terminal Phe protection and moderate protection of Lys29.

3.4 Cucurbit[6]uril as a lysine protecting group for protein modification

While examples of lysine protecting groups for protein modification have been reported, these strategies require the instillation of noncanonical amino acids onto the protein surface.[4–6] Therefore, it would be desirable to have a convenient method for native lysine protection. Cucurbit[6]uril is known to interact with protein surface lysines with reported binding affinities on the order of 10^2 M^{-1}.[8] While this is a much weaker interaction than that of CB7 for N-terminal Phe, it may be enough to provide some lysine protection if CB6 is present in sufficient excess in solution. While our initial test of the FGGSWKG peptide revealed very modest lysine protection via CB6, we hypothesized that buffer conditions could heavily influence CB6 binding to lysine, as it is known that small ions such as potassium and sodium can bind to the portal of CB6 and

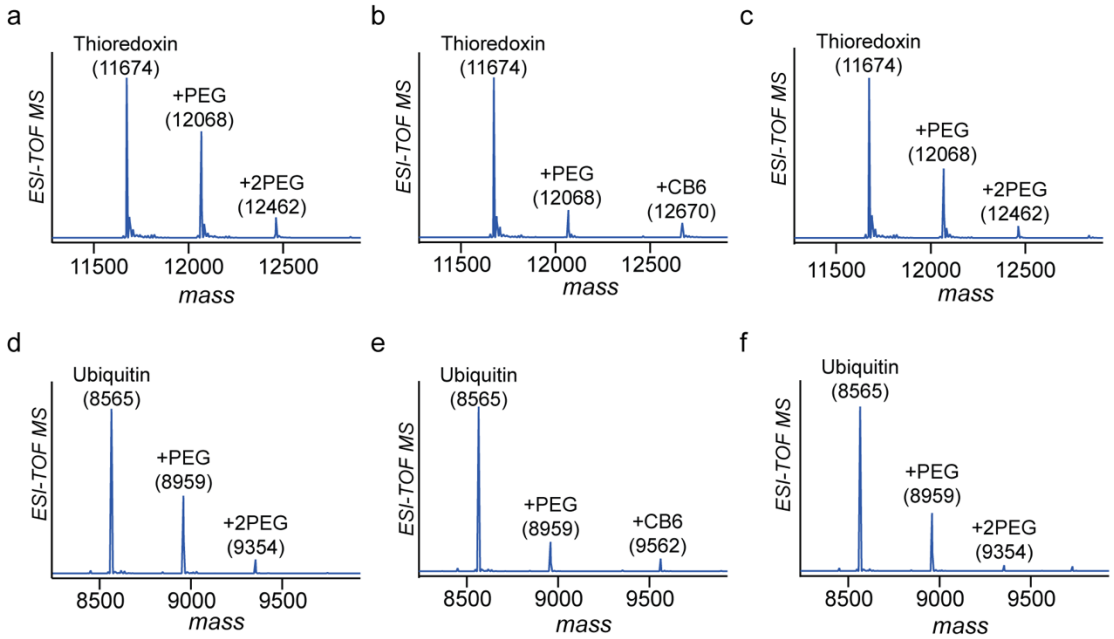


Figure 3.5. Cucurbituril-mediated protection of lysines on protein surfaces. PEG_8-NHS (10 equiv) was used to modify thioredoxin (Ser N terminus) or ubiquitin (Met N terminus) with or without CB protection. CBs should not interact strongly with either protein N terminus and should rather bind exclusively to accessible lysine residues on the protein surface. (a,d) Incubation of thioredoxin and insulin with NHS ester led to 1-2 PEG modifications without CB protection. (b,e) CB6 incubation (25 equiv) led to marked protection of the reactive lysines, and a protein-CB6 peak was observed for both proteins. (c,f) Incubation with CB7 (50 equiv) afforded modest protection, as double modification was still observed, and only trace protein-CB7 interactions were observed by ESI-TOF MS.

interfere with CB6 host-guest chemistry.[7,19] We therefore used Bis-TRIS buffer going forward, as we had previously shown that CB6 click chemistry is tolerated well with that buffer system (Figure 2.3).

Using Bis-TRIS pH 7 for all subsequent lysine protection reactions, we tested CB6 as a supramolecular protecting group against both thioredoxin and ubiquitin proteins. Thioredoxin contains 10 lysines, as well as a solvent accessible Ser N terminus. Ubiquitin has 5 reactive lysines that are used in the ubiquitination process, but has a buried Met N terminus that is nonreactive. Interestingly, CB6 has been shown to bind to ubiquitin lysines and can induce reversible conformational changes of the protein upon binding, depending on the lysine binding site.[20] For both proteins tested, significant protection of reactive lysines were observed, as the overall modification percentages were lowered and no second PEG addition was observed. Additionally, a protein-CB6 peak was observed in the mass spectra for both thioredoxin and ubiquitin (Figure 3.5b,d).

When the proteins were incubated with CB7, only modest protection was afforded, and only trace protein-CB7 peaks were observed in the mass spectra (Figure 3.5c,f). Thus, CB6 and CB7 can function complementarily to one another. Whereas CB6 offers lysine protection, CB7 can function as a more selective protecting group for aromatic N termini. Future work will focus on proteomic analyses of both ubiquitin and thioredoxin conjugates to determine the exact modification site in the presence and absence of CB6 protection.

3.5 Redirection of insulin modification to the A chain via dual CB6-CB7 protection

We next set out to determine if a combination of CB6-Lys and CB7-Phe N-terminal protection of insulin could lead to redirection of insulin acylation to the A chain N terminus. Insulin was incubated with either no protection, CB6, CB7, or a combination of the two, prior to modification with 5 equiv PEG_8-NHS in Bis-TRIS buffer (Figure 3.6). When reacted without any CB protection, B chain modification was favored by a ratio of 1.5:1 as observed by LC/MS of DTT-reduced conjugates. This is in agreement with literature findings which report greater B chain reactivity than A chain.[15,18] When incubated with CB6 alone, B chain favorability was increased to 1.9:1 modification over A chain, while overall modification dropped from 88% in the unprotected case, to 57% when protected by CB6. While CB6 is predicted to bind to Lys29 on the B chain of insulin, it is possible that CB6 binding of insulin at Lys29 leads to mild protection of the A chain N terminus through steric interference, as the two residues are located within 1 nm of one another. This

Protecting Group	Insulin modification	Insulin chain selectivity
None	*52%* (single) *33 %* (double) *3%* (triple)	*0.67:1* **(A:B)**
CB6 (5 equiv)	*40%* (single) *17%* (double)	*0.52:1* **(A:B)**
CB7 (5 equiv)	*16%* (single) *3%* (double)	*1.3* **:** *1* **(A:B)**
CB6 + CB7 (2.5 equiv each)	*17%* (single) *3%* (double)	*2.4* **:** *1* **(A:B)**
CB6 + CB7 (5 equiv each)	*10%* (single) *1%* (double)	*2.8* **:** *1* **(A:B)**

Figure 3.6. Redirection of insulin chain modificaiton via CB6 and CB7 protection. Incubation of insulin with CB6 alone prior to PEGylation led to modest overall protection with moderate enhancement in B chain modification preference. Incubation with CB7 alone led to significant protection with a redirection to A chain modification preference. Incubation with CB6 and CB7 led to 4-fold enhancement in A chain acylation over B chain.

interference would therefore afford an increase in acylation favorability toward the B chain N terminus. Incubation with CB7 alone led to a shift in favorability from the B chain to the A chain by a ratio of 1.3:1 A:B modification. When coupled with CB6 protection, that ratio was increased to 2.4:1 in the case of 2.5 equiv each of CB6 and CB7, and 2.8:1 when 5 equiv of both CB6 and CB7 were used. Interestingly, the overall modification percent remained similar between CB7 alone at 5 equiv and 2.5 equiv each of CB6+CB7 at ~20% overall insulin modification. Thus, incubation with CB6 and CB7 in unison shows promise as a means of blocking both the B chain N-terminal Phe and B chain Lys29 sites, thereby redirecting NHS ester acylation to the N terminus of the A chain.

3.6 Conclusions

Protecting groups for protein modification could afford more selective and homogenous protein conjugates without the use of complex reagents, or the installation of noncanonical amino acids into the protein backbone. By using supramolecular host-guest interactions, we were able to demonstrate protection of nucleophilic protein residues with cucurbituril macrocycles. CB6 was used for lysine protection, whereas CB7 was used for the protection of aromatic N-terminal residues. When incubated with CBs, reduced modification levels with activated esters were observed, indicating protection of nucleophilic amine residues. In the case of CB7 protection of insulin N-terminal Phe, a redirection to the A chain N terminus was observed. This redirection was enhanced when both CB7 and CB6 were incubated with insulin to afford protection at both the B chain N terminus and at the B chain Lys29 position. Future work will focus on optimizing reaction conditions to achieve more complete protection of lysines with CB6 and aromatic N termini with CB7. Additionally, proteomic analyses of CB-protected conjugates will be performed to confirm the sites of protection. Lastly, using these supramolecular protecting groups, homogeneous protein conjugates will be developed and tested for improved activity against their heterogeneous counterparts.

3.7 Materials and methods

All solvents and reagents, including cucurbit[6]uril, cucurbit[7]uril, insulin, thioredoxin, and ubiquitin, were purchased from commercial suppliers and used without further purification. A C18 column was used for analytical and semi-preparative reverse phase high performance liquid chromatography (RP-HPLC) on an Agilent 1100 Series Capillary LC. Runs were eluted with H_2O/MeCN (0.1 % v/v TFA) and monitored using a UV-Vis detector. UV-Vis spectra were obtained using a NanoDrop (Thermo Scientific).

Mass spectrometry. Protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an Agilent 6224 Time-of Flight (TOF) LC/MS system equipped with a Turbospray ion source. LC/MS-MS of peptides was performed using an Agilent 1100 series liquid chromatograph (Agilent Technologies, USA) that was connected in line with a 3200 Q Trap LC/MS-MS system.

General procedure for CB protection of peptide and protein samples. To a solution of protein (20 μM) or peptide (100 μM) in the reaction buffer was added 1-20 equiv of CB6 or CB7 (1-5 mM stock solution in reaction buffer). The protein-CB sample was allowed to incubate for 1 h at room temperature, followed by addition of 1-20 equiv of NHS ester (1-5 mM stock in DMSO). The reaction was incubated at room temperature for 3 h, at which point 50 equiv of hydroxylamine was added to quench the reaction and disrupt any weaker protein NHS ester modifications. In some cases, DTT (50 mM stock in ddH_2O) was added to a final concentration of 5 mM. After quenching and/or DTT treatment, the conversion was monitored by LC/MS. In some cases, the protein conjugate was spin concentrated 5 times into reaction buffer to remove excess NHS ester, followed by CB7 deprotection with 50 equiv of ferrocenemethanol (50 mM stock in DMSO), and purified by spin concentration 5 times into reaction buffer.

FGGSWKG peptide synthesis. FGGSKWG was synthesized using standard Fmoc solid phase peptide synthesis (SPPS) on Gly-Wang resin (Anaspec, Fremont, CA). Deprotection of the Fmoc groups was performed with two 10 min incubations in a 30% v/v piperidine in dimethylformamide (DMF) solution. Couplings were carried out using 10 equiv of amino acid with 10 equiv of 2-(6-chloro-1-*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), and 20 equiv of *N*,*N*-diisopropylethylamine (DIPEA) in DMF for 20 min. The N terminus of the peptide was modified on resin with *N*-Boc-protected Phe using the same coupling conditions. The peptide was cleaved from the resin after incubation with a solution of 95:2.5:2.5 TFA:TIPS:H_2O for 2 h. Excess TFA was removed under a stream of N_2 and the peptide was precipitated in cold diethyl ether. The precipitate was redissolved in H_2O and purified by RP-HPLC, eluting in an aqueous gradient 5% to 95% MeCN/0.1% TFA in H_2O/0.1% TFA over 45 min at a flow rate of 3.0 mL/min. After lyophilization, FGGSWKG was obtained as a colorless powder. HRMS (ESI-QTRAP-MS): *m/z* calcd for FGGSWKG $C_{35}H_{48}N_9O_9$ $[M+H]^+$ 737.82, observed 738.30.

3.8 References

(1) Rosen, C. B.; Francis, M. B. Targeting the N Terminus for Site-Selective Protein Modification. *Nat. Chem. Biol.* **2017**, *13* (7), 697–705.

(2) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chemie Int. Ed.* **2009**, *48* (38), 6974–6998.

(3) Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. Site-Selective Protein-Modification Chemistry for Basic Biology and Drug Development. *Nat. Chem.* **2015**, *8* (2), 103–113.

(4) Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; et al. Palladium-Triggered Deprotection Chemistry for Protein Activation in Living Cells. *Nat. Chem.* **2014**, *6* (4), 352–361.

(5) Davis, L.; Chin, J. W. Designer Proteins: Applications of Genetic Code Expansion in Cell Biology. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (3), 168–182.

(6) Gautier, A.; Deiters, A.; Chin, J. W. Light-Activated Kinases Enable Temporal Dissection of Signaling Networks in Living Cells. *J. Am. Chem. Soc* **2011**, *133*, 2124–2127.

(7) Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115* (22), 12320–12406.

(8) Urbach, A. R.; Ramalingam, V. Molecular Recognition of Amino Acids, Peptides, and

Proteins by Cucurbit[n]uril Receptors. *Isr. J. Chem.* **2011**, *51* (5–6), 664–678.
(9) Sonzini, S.; Marcozzi, A.; Gubeli, R. J.; van der Walle, C. F.; Ravn, P.; Herrmann, A.; Scherman, O. A. High Affinity Recognition of a Selected Amino Acid Epitope within a Protein by Cucurbit[8]uril Complexation. *Angew. Chemie Int. Ed.* **2016**, *55* (45), 14000–14004.
(10) Gubeli, R. J.; Sonzini, S.; Podmore, A.; Ravn, P.; Scherman, O. A.; van der Walle, C. F. Selective, Non-Covalent Conjugation of Synthetic Peptides with Recombinant Proteins Mediated by Host–guest Chemistry. *Chem. Commun.* **2016**, *52* (22), 4235–4238.
(11) Walker, S.; Oun, R.; McInnes, F. J.; Wheate, N. J. The Potential of Cucurbit[n]urils in Drug Delivery. *Isr. J. Chem.* **2011**, *51* (5–6), 616–624.
(12) Hou, C.; Huang, Z.; Fang, Y.; Liu, J. Construction of Protein Assemblies by Host–guest Interactions with Cucurbiturils. *Org. Biomol. Chem.* **2017**, *15* (20), 4272–4281.
(13) Chinai, J. M.; Taylor, A. B.; Ryno, L. M.; Hargreaves, N. D.; Morris, C. A.; Hart, P. J.; Urbach, A. R. Molecular Recognition of Insulin by a Synthetic Receptor. *J. Am. Chem. Soc.* **2011**, *133* (23), 8810–8813.
(14) Lee, J. W.; Shin, M. H.; Mobley, W.; Urbach, A. R.; Kim, H. I. Supramolecular Enhancement of Protein Analysis via the Recognition of Phenylalanine with Cucurbit[7]uril. *J. Am. Chem. Soc.* **2015**, *137* (48), 15322–15329.
(15) Hinds, K. D.; Kim, S. W. Effects of PEG Conjugation on Insulin Properties. *Adv. Drug Deliv. Rev.* **2002**, *54* (4), 505–530.
(16) Webber, M. J.; Appel, E. A.; Vinciguerra, B.; Cortinas, A. B.; Thapa, L. S.; Jhunjhunwala, S.; Isaacs, L.; Langer, R.; Anderson, D. G. Supramolecular PEGylation of Biopharmaceuticals. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (50), 14189–14194.
(17) Hentz, N. G. Synthesis and Characterization of Insulin-Fluorescein Derivatives for Bioanalytical Applications. *Anal. Chem.* **1997**, *69* (24), 4994–5000.
(18) Chan, A. O.-Y.; Ho, C.-M.; Chong, H.-C.; Leung, Y.-C.; Huang, J.-S.; Wong, M.-K.; Che, C.-M. Modification of N-Terminal α-Amino Groups of Peptides and Proteins Using Ketenes. *J. Am. Chem. Soc.* **2012**, *134* (5), 2589–2598.
(19) Buschmann, H.-J.; Cleve, E.; Schollmeyer, E. Cucurbituril as a Ligand for the Complexation of Cations in Aqueous Solutions. *Inorganica Chim. Acta* **1992**, *193* (1), 93–97.
(20) Lee, J. W.; Heo, S. W.; Lee, S. J. C.; Ko, J. Y.; Kim, H.; Kim, H. I. Probing Conformational Changes of Ubiquitin by Host-Guest Chemistry Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (1), 21–29.

3.9 Additional Figures

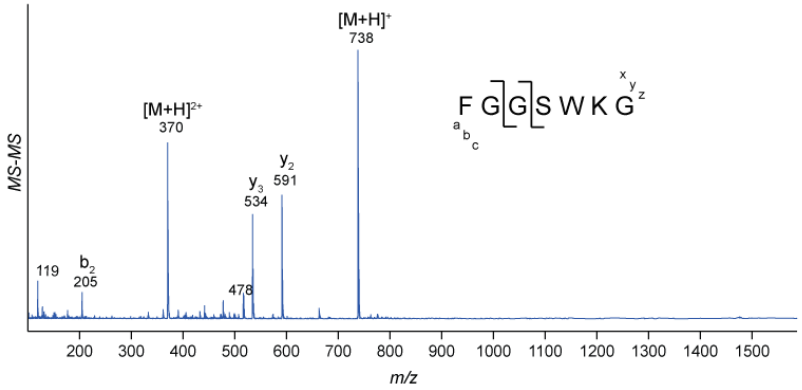


Figure S3.1. LC/MS-MS analysis of FGGSWKG peptide. FGGSWKG $(M+H)^+$ was observed at 738 m/z, as were the y_3 fragment at 534 m/z and the y_2 fragment at 591 m/z.

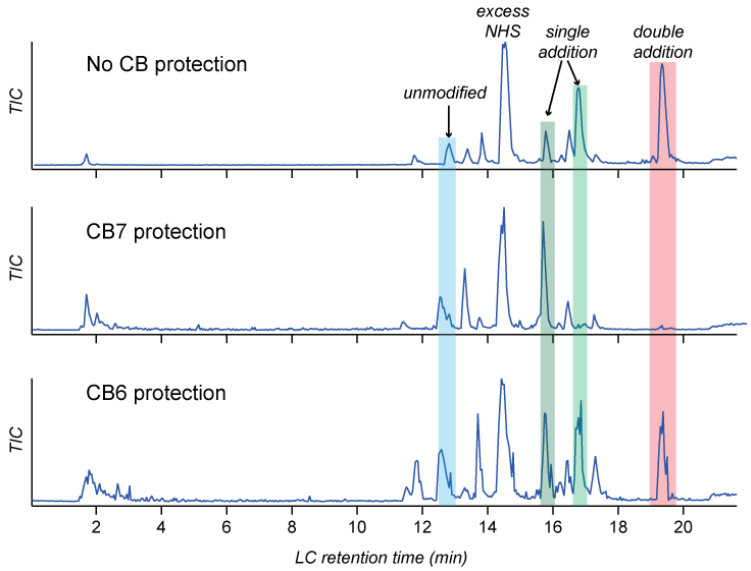


Figure S3.2. Liquid chromatograms of FGGSWKG modification by PEG_8-NHS, (a) without CB protection, (b) with CB7 protection, and (c) with CB6 protection.

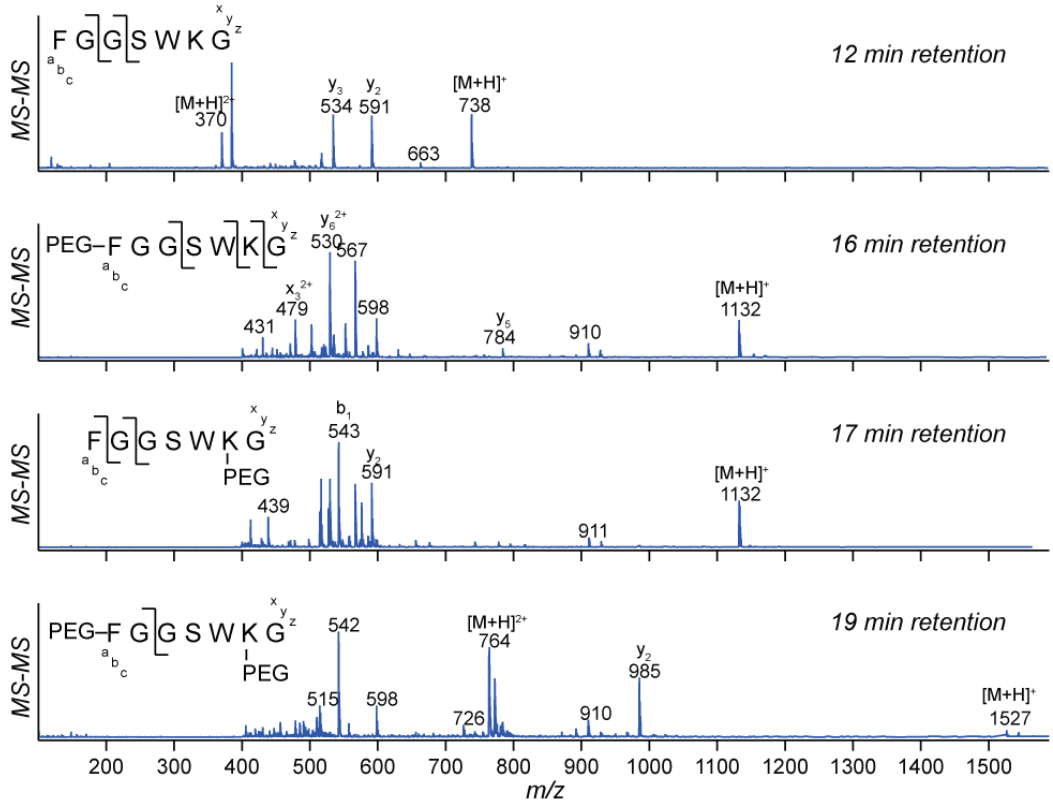


Figure S3.3. Representative MS-MS spectra of FGGSWKG modification peaks at 12, 16, 17, and 19 min.

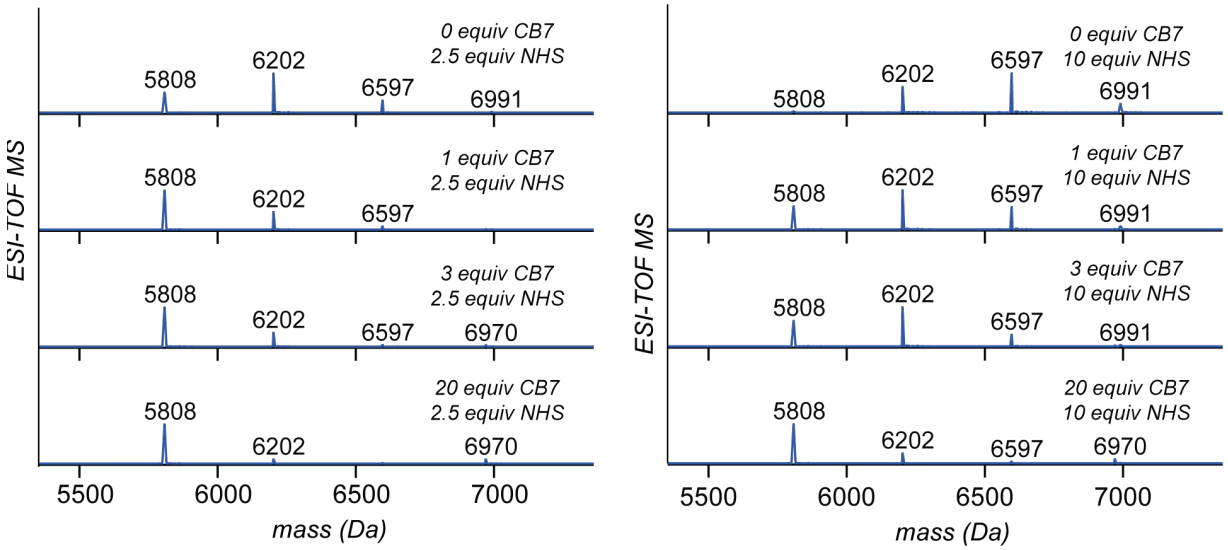


Figure S3.4. LC/MS spectra of insulin modified with varying equivalents of CB7 and PEG_8-NHS. Insulin (5808 Da) was modified with PEG_8-NHS, leading to either single (6202 Da), double (6597 Da), or triple (6991) PEGylation. Insulin-CB7 peaks were observed at 6970 Da.

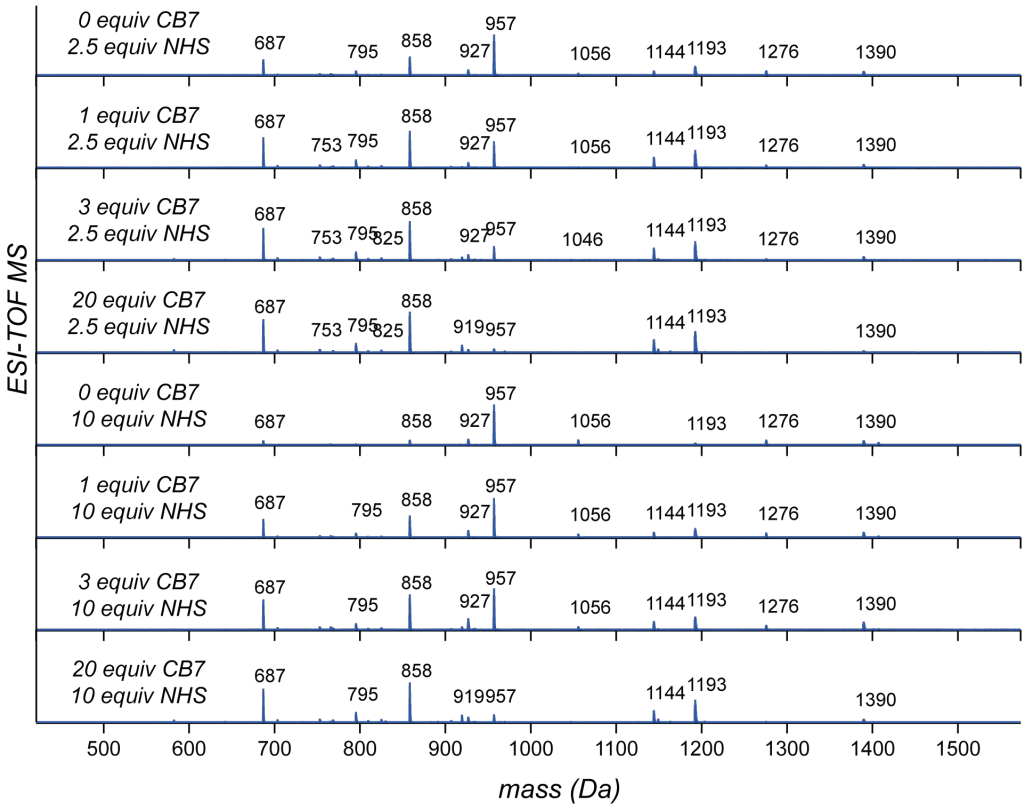


Figure S3.5. LC/MS spectra of insulin modified with varying equivalents of CB7 and PEG_8-NHS after DTT reduction.

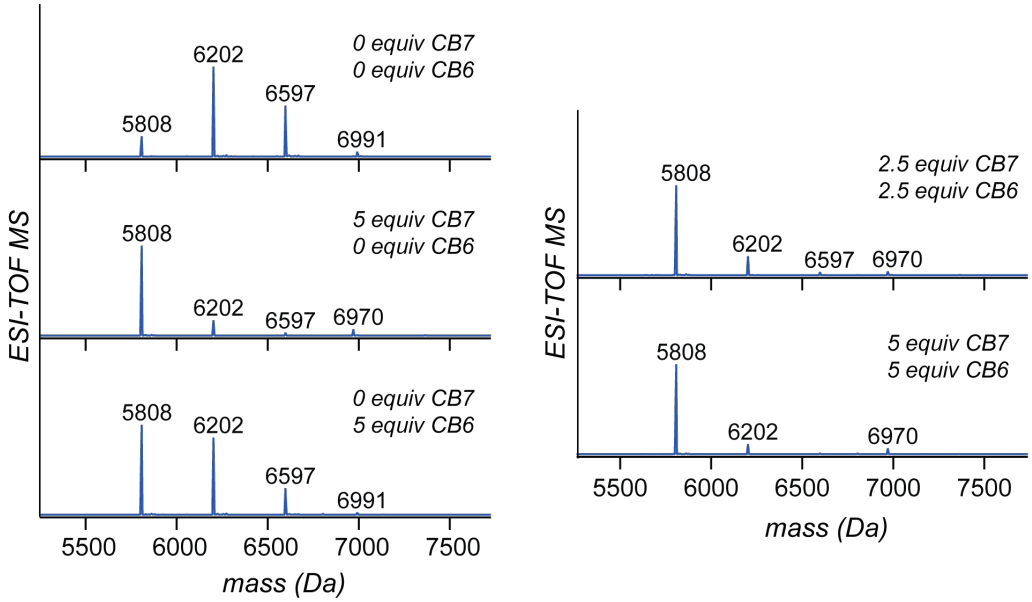


Figure S3.6. Modification of insulin with CB6 and CB7 protection. Insulin was incubated with varying equivalents of CB6, CB7, or both, prior to modification with 5 equiv PEG_8-NHS.

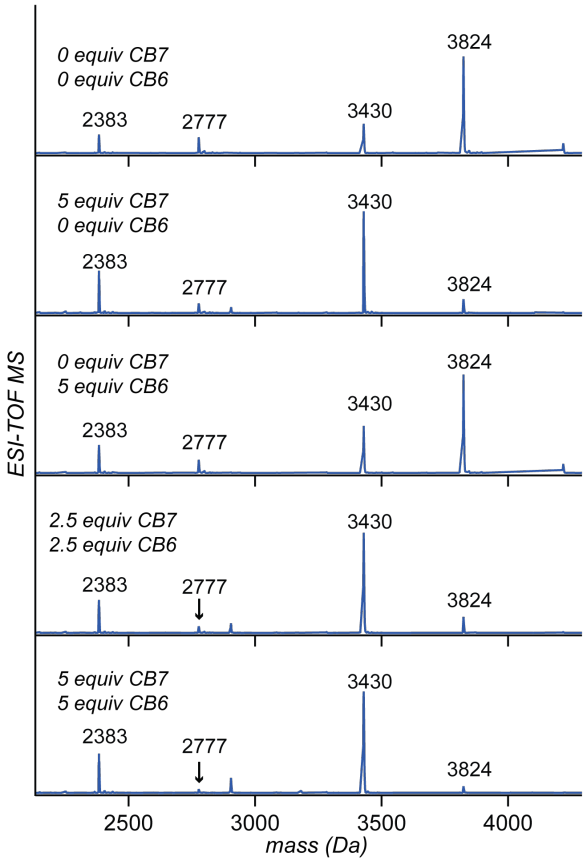


Figure S3.7. Distribution of A and B chain modification of insulin with CB6 and CB7 protection. Insulin was incubated with varying equivalents of CB6, CB7, or both, prior to modification with 5 equiv PEG_8-NHS. After 3 h incubation, DTT was added to reduce disulfides connecting the A and B chains. Deconvolution of charged states led to resolved spectra shown. The insulin A chain has lower TIC when compared to B chain. Modification percentages were calculated by via peak integration of modified A vs. unmodified A, and modified B vs. unmodified B.

Chapter 4

Rotaxane probes for disease detection by ^{129}Xe hyperCEST NMR

ABSTRACT: Sensitive and precise methods for disease detection are necessary for effective disease diagnosis and treatment. ^{129}Xe hyperCEST NMR is a promising analytical tool due to its high sensitivity and molecular detection capabilities, but the reports of turn-on probes for hyperCEST have been limited. We report a new class of activated hyperCEST agents using cucurbit[6]uril (CB6) rotaxanes, whereby a hyperCEST response is produced upon the cleavage of a CB6 rotaxane by a chemical event. The subsequent increase in ^{129}Xe@CB6 interactions produces a hyperCEST response far above background levels. We first report on a model system whereby turn-on is initiated through ester hydrolysis. We then took advantage of this platform for the design and implementation of CB6 rotaxanes as probes for protease detection, whereby a protease of interest cleaves a peptide recognition sequence on the rotaxane, increasing the accessibility of CB6 for xenon interactions. We further report a CB6 rotaxane probe for the detection of matrix metalloproteinase 2 (MMP-2), which is overexpressed in cancer tissue. The turn-on rotaxanes successfully produced hyperCEST responses when incubated with its target enzyme, demonstrating that CB6 rotaxane probes can be used for protease detection. We next designed a CB6 rotaxane that combines a hyperCEST probe with drug delivery applications, by synthesizing a protein-drug conjugate using CB6 promoted click chemistry. This protein-drug conjugate contained a cleavable linker that facilitated drug release and subsequent release of the CB6 macrocycle for Xe binding and hyperCEST detection. Thus, CB6 rotaxanes are promising new diagnostic tools for sensitive and modular disease detection.

Portions of the work described in this chapter have been reported in the following publications:

CC Slack, JA Finbloom, K Jeong, CJ Bruns, DE Wemmer, A Pines, MB Francis. Rotaxane probes for protease detection by ^{129}Xe hyperCEST NMR. *Chem. Commun.* 2017, 53; 1076-1079.

JA Finbloom, CC Slack, CJ Bruns, K Jeong, DE Wemmer, A Pines, MB Francis. Rotaxane-mediated suppression and activation of cucurbit[6]uril for molecular detection by ^{129}Xe hyperCEST NMR. *Chem. Commun.* 2016. 52; 3119-22

4.1 ^{129}Xe hyperCEST NMR for disease diagnostics

Localized molecular detection with high sensitivity and selectivity is of paramount importance in early disease detection. Current techniques such as magnetic resonance imaging (MRI), the clinical version of nuclear magnetic resonance (NMR), and x-ray computed tomography benefit from being noninvasive imaging techniques with high resolution and excellent tissue penetration. However, these techniques lack molecular information that could provide valuable insight for disease diagnosis. While advances in responsive ^{1}H NMR contrast agents that turn on only in select environments have led to the detection of enzymes,[1] signalling molecules,[2] and redox conditions,[3] none of these advances offer direct molecular detection with high signal contrast. Hyperpolarized xenon MRI provides a sensitive complement to ^{1}H MRI. The use of hyperpolarization increases signal strengths by several orders of magnitude, and there is no natural background ^{129}Xe signal in biological environments.[4,5] Xenon participates in a range of supramolecular interactions, which produce distinct molecular signals for different xenon environments inside various hosts. This allows xenon NMR to be used for molecular imaging applications.[6,7] Coupling hyperpolarization with chemical exchange saturation transfer (CEST) between xenon in free and host-encapsulated forms, termed ^{129}Xe hyperCEST NMR, allows highly sensitive molecular detection at concentration levels that are required for the early detection of disease biomarkers.[8]

Taking advantage of this sensitivity, hyperCEST has been used to detect cancer markers,[9–11] small molecule analytes,[12] and cell surface glycans.[13] These methods rely on either the targeted delivery of xenon hosts to a region of interest, or small chemical shift differences between bound and unbound xenon sensors. A desirable alternative approach would be to suppress the ^{129}Xe@host signal completely until the sensor reaches a region of interest, where it is selectively activated. Previous examples of ^{129}Xe hyperCEST have relied on cryptophane-A (CryA) xenon hosts that are hydrophobic, costly, and difficult to functionalize.

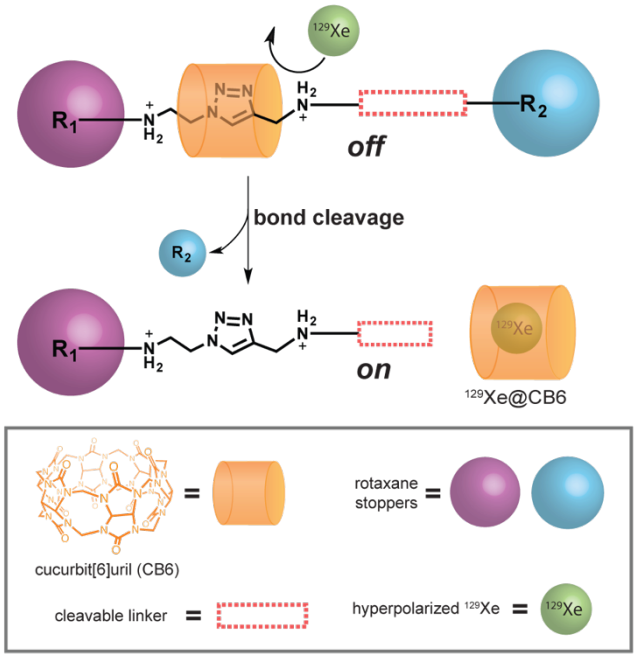


Figure 4.1. General strategy for activating CB6-rotaxanes for ^{129}Xe NMR. A rotaxane with a triazole-diammonium-CB6 center and a cleavable linker blocks ^{129}Xe-CB6 interactions and thus suppresses ^{129}Xe@CB6 NMR signals. Once cleaved, the CB6 ring can dissociate and serve as a host for hyperpolarized ^{129}Xe, producing a ^{129}Xe@CB6 NMR signal.

Cucurbit[6]uril (CB6) is a promising xenon host for ^{129}Xe NMR detection, as it produces a distinctive ^{129}Xe@CB6 signal, has better exchange parameters for hyperCEST when compared to CryA, is soluble in most buffers and biological environments, and is commercially available.[14] While CB6 has only recently been reported for use with ^{129}Xe hyperCEST NMR,[15,16] it is attracting significant attention over CryA-based sensors due to these advantages. However, a major limitation of CB6 sensors is the difficult chemical functionalization to generalize them for diverse spectroscopic applications.[14,17]

While the covalent functionalization of CB6 remains a challenge, its ability to participate in a wide range of supramolecular host-guest interactions has been well studied.[14] Competing CB6 guests have been shown to suppress ^{129}Xe@CB6 signals to varying degrees, although in each case an excess of competing molecules was needed to suppress the ^{129}Xe@CB6 signal.[15,18] Recently, Schröder and coworkers took advantage of CB6 host-guest interactions to design an enzyme-sensing platform based on a competition between ^{129}Xe and an enzyme product for binding of the CB6 cavity.[18] However, it would be beneficial to have a more modular ^{129}Xe turn-on sensor platform that does not rely on a specific molecule's affinity toward cucurbituril. It would likewise be valuable to create a system that will block ^{129}Xe@CB6 interactions with greater control, as such an approach could eliminate background signals until the CB6 reaches a region of interest, where it is then released to produce a ^{129}Xe@CB6 signal.

Herein, we report the design, synthesis, and implementation of an activatable CB6 rotaxane platform for ^{129}Xe hyperCEST NMR (Figure 4.1). In this system, the ^{129}Xe@CB6 interactions are blocked by a mechanical bond, where the "dumbbell" stopper component of a rotaxane prevents ^{129}Xe from accessing the CB6 cavity. Upon cleavage of the rotaxane stopper, dissociation of the CB6 ring allows it to host ^{129}Xe, producing a ^{129}Xe@CB6 signal. Thus, CB6 rotaxanes afford a new turn-on probe approach to ^{129}Xe hyperCEST disease detection and may be used in a wide range of detection applications.

4.2 Design and synthesis of CB6 rotaxane probes

The rotaxanes used in this work were synthesized by CB6-promoted azide-alkyne 1,3-dipolar cycloaddition, which is routinely used for CB6-rotaxane synthesis, allows for increased solubility of CB6 through complexation, and is a facile method for the preparation of rotaxanes with diverse functional groups for various applications.[14,19,20] Dmochowski and coworkers previously reported that CB6 incubated with excess putrescine guests led to a suppression of ^{129}Xe@CB6 NMR signals.[15] However, in equal concentrations of competitors relative to CB6, little change in the ^{129}Xe@CB6 hyperCEST NMR signal was observed. Schröder and coworkers similarly observed that excess concentrations of cadaverine were able to suppress ^{129}Xe@CB6 signals.[18] Since both of these competing guests were alkane-diammonium ions, the triazole-diammonium recognition unit of our rotaxane platform might therefore pose a challenge to CB6 detection if it were able to overwhelm ^{129}Xe@CB6 interactions through complex competition. To test if this competition could suppress ^{129}Xe@CB6 signals, we synthesized CB6 complex **1** via a CB6-promoted azide-alkyne 1,3 dipolar cycloaddition (Figure 4.2). Complex **1** lacks the stoppers of a full rotaxane, allowing the

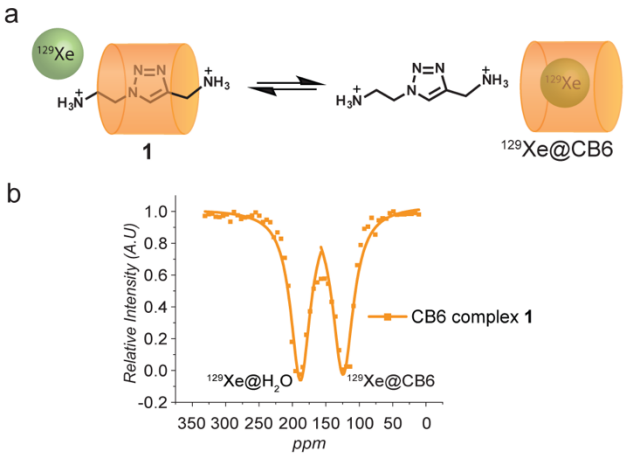


Figure 4.2. Complex **1** demonstrates competitive host-guest interactions between **1** and ^{129}Xe@CB6. (a) Scheme outlining the dynamic host-guest interactions of **1** and ^{129}Xe. (b) ^{129}Xe HyperCEST NMR spectrum of 100 μM CB6 complex **1**. A strong ^{129}Xe@CB6 peak was observed.

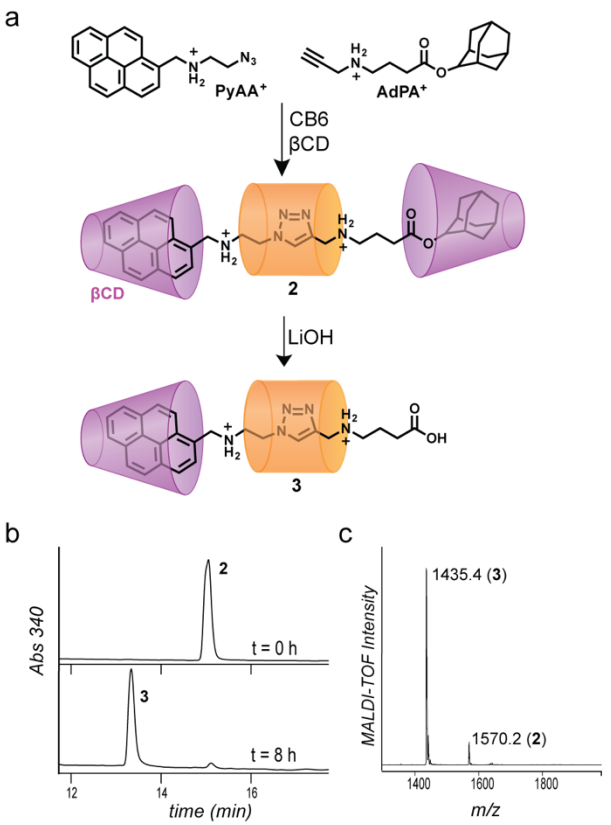


Figure 4.3. Rotaxane **2** exhibits responsive cleavage. (a) Rotaxane **2** containing a labile ester bond was synthesized with β-cyclodextrin (βCD) caps to improve water solubility and to accelerate the rotaxane capture reaction. (b) HPLC traces monitoring the cleavage rate of **2** (15 min) when exposed to 10 equiv of LiOH at 40 °C. Near complete conversion to **3** (13 min) was observed by 8 h. (c) MALDI-TOF MS confirmed the presence of ester-hydrolysis product **3** at 1435 m/z.

CB6 ring and triazole-diammonium axle components to exchange easily in solution. This construct therefore represents a model for the post-cleavage product of an activated CB6-rotaxane, where the CB6 host and the triazole-diammonium guest are in equal concentrations. A strong ^{129}Xe@CB6 peak was observed for 100 μM CB6 complex **1** by ^{129}Xe hyperCEST NMR, approaching complete saturation of the dissolved xenon signal, similar to 10 μM free CB6 in solution (Figure S4.2). While there was some signal suppression due to competition between complexes **1** and ^{129}Xe@CB6, xenon was still able to exchange rapidly between the bulk water and the CB6 molecules, giving a strong hyperCEST response. These observations are in agreement with previous competition studies.[15,18]

Once it was confirmed that a competitive triazole-diammonium guest would not significantly block ^{129}Xe-CB6 interactions, we set out to synthesize a chemically activated CB6-rotaxane that could completely suppress the ^{129}Xe@CB6 NMR signal until undergoing a controlled cleavage and subsequent release of CB6 for xenon binding. Rotaxane **2** was synthesized with pyrene-functionalized 2-azidoethylamine (**PyAA**$^+$) and an adamantyl-ester functionalized propargylamine (**AdPA**$^+$) as rotaxane stoppers (Figure 4.3). β-cyclodextrin (βCD) caps were added to improve the solubility of the end groups and to accelerate the rotaxane capture reaction.[19] Compared to previous xenon sensors, which have relied on direct CryA conjugates, this synthetic strategy provides a facile and modular approach. To confirm that the βCD caps would not interact with xenon and affect the ^{129}Xe hyperCEST response of **2**, βCD alone in solution was tested and revealed no measurable background ^{129}Xe hyperCEST signals.

The ^{129}Xe hyperCEST response of **2** revealed no detectable ^{129}Xe@CB6 signal even at 100 μM (Figure 4.4). Free CB6 in solution is easily detected at low nanomolar concentrations;[15,16] thus it is revealing that even at relatively high concentrations of **2**, no CB6 was detected. These results demonstrate that ^{129}Xe is completely prevented from exchanging with the CB6 cavity and producing a hyperCEST response when CB6 is locked to the rotaxane complex. In contrast to previous work that explored supramolecular CB6 competition between xenon and competing guests,[15,18,21] rotaxane **2** demonstrated complete suppression of ^{129}Xe@CB6 signals without using excess concentrations of the triazole-diammonium guest relative to CB6.

To create an activated CB6-rotaxane for ^{129}Xe NMR, **2** was synthesized with a labile ester group that can be hydrolysed to **3** and lead to the release of CB6. Treatment with 10 equiv of LiOH led to the complete ester hydrolysis of **2** by 8 h, as confirmed by high performance liquid chromatography (HPLC) and mass spectrometry (MS) (Figure 4.3b,c). After treatment with LiOH, rotaxane **2** was activated, and a significant ^{129}Xe@CB6 signal was observed at levels nearing complete saturation, similar to CB6 complex **1** (Figure 4.4). Thus, even in complexes containing a single bulky rotaxane stopper and a terminal carboxylate on the axle portion of the rotaxane, it is possible to maintain the association and exchange kinetics between xenon and CB6 that are necessary to produce a hyperCEST response.

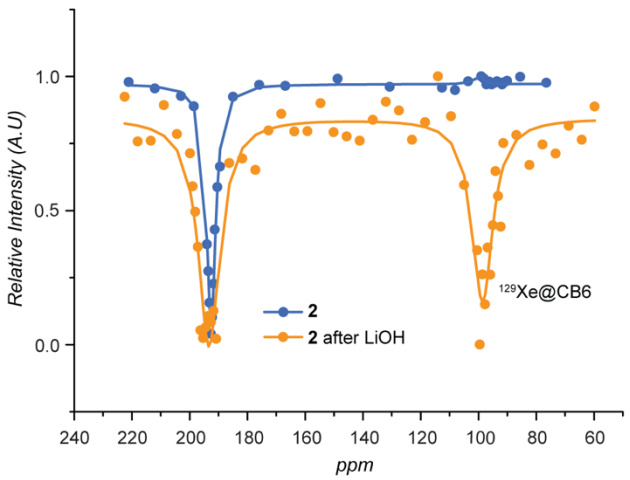


Figure 4.4. Rotaxane **2** can be activated for ^{129}Xe hyperCEST NMR. ^{129}Xe-hyperCEST NMR spectra of 100 μM **2** displayed no significant ^{129}Xe@CB6 signal. After exposure to 10 equiv of LiOH at 40 °C a substantial increase in the ^{129}Xe@CB6 peak was observed.

4.3 Rotaxane probes for detection of matrix metalloproteinases

Proteases are widely used markers for disease detection, as certain proteases are overexpressed in diseased tissues and cells.[1,22] Proteases have the added benefits of being catalytically active, such that only a small amount is required to activate the probe. They can also be highly specific in their substrate recognition motifs. While proteases have been used in several detection platforms, the reports of such probes for ^{129}Xe NMR have been limited. To create CB6 rotaxane probes for protease detection, our design incorporates peptide sequences into the axle component of the rotaxane (Figure 4.5). These peptides can be synthesized to contain any sequence, including protease recognition sequences of interest. After proteolysis events, a portion of the peptide remains on the axle component of the rotaxane, leaving a post-cleavage product mechanically unlocking CB6, which can be released from the rotaxane becoming CEST active. In theory, this design can be used to detect any of a wide variety of proteases using ^{129}Xe hyperCEST NMR if the CB6 becomes more accessible for ^{129}Xe interactions following proteolysis.

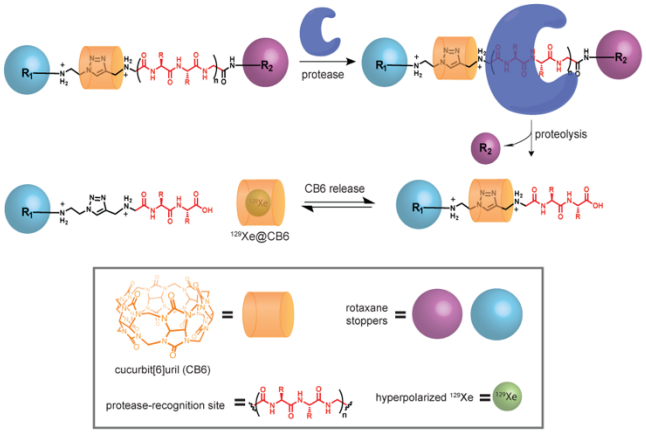


Figure 4.5. CB6 rotaxanes probes are activated by proteases to produce a hyperCEST response. CB6 rotaxanes containing a protease cut sequence are cleaved specifically in the presence of an enzyme of interest. Upon cleavage of the rotaxane, CB6 becomes more accessible to host ^{129}Xe and produce a hyperCEST response.

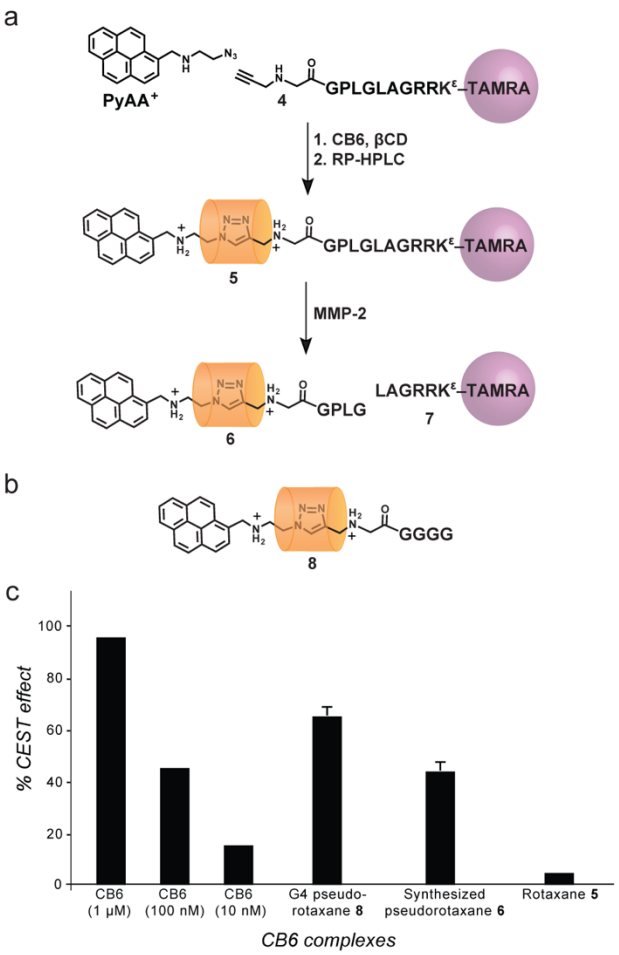


Figure 4.6. Synthesis and hyperCEST response of CB6 complexes. (a) CB6 rotaxane **5** was synthesized using CB6-catalyzed azide-alkyne click chemistry in the presence of βCD to accelerate the reaction. Cyclodextrin is removed during purification by RP-HPLC. Rotaxane **5** contains the MMP-2 recognition sequence PLG-LAG. (b) The percent CEST effect was measured for various CB6 complexes. Oligoglycine pseudorotaxane **8** and pseudorotaxane **6** were synthesized to function as positive controls. Both pseudorotaxanes **8** and **6** demonstrated significant CEST effects, whereas capped rotaxane **5** showed minimal CEST effect. All measurements were performed at 10 μM rotaxane in pH 7 phosphate buffer. Error bars represent standard instrumental error for n=4. ***The signal observed for **5** was not significant when compared to variation in background water signal.

Matrix metalloproteinases represent an interesting disease-related target for this rotaxane probe platform. Rotaxane **5** was synthesized with pyrene and (5,6)-carboxytetramethylrhodamine (TAMRA) stoppers, and an axle containing the PLG-LAG recognition sequence of matrix metalloproteinase-2 (MMP-2). MMPs are overexpressed in a wide range of metastatic tumors,[22–24] and are present in heart tissue after myocardial infarction.[25] MMP-2 has been used successfully as a target for a variety of drug delivery and imaging applications.[22–24,26] Previously, a xenon sensor for MMP was developed with a peptide recognition sequence for MMP-7 covalently attached to CryA.[27] Upon cleavage by MMP-7, a change of approximately 0.5 ppm in the ^{129}Xe@CryA chemical shift was observed. In contrast, rotaxane **5** offers the opportunity to detect MMP-2 by measuring the increase in CEST effect upon cleavage. This on/off change could allow better detection in biologically complex environments than measuring a small shift in a broad line.

Rotaxane **5** was synthesized using CB6-promoted azide-alkyne cycloaddition (Figure 4.6a). Unlike rotaxane **2**, the TAMRA and peptide moieties provided sufficient water-solubility to remove the need for βCD caps. Since CB6 bound to triazole-diammonium guests without any bulky stoppers or additional axle components reduced the hyperCEST effect when compared to free CB6 in solution, we sought to confirm that the peptide component of the rotaxane would not prevent CB6 release and subsequent hyperCEST response of the CB6 rotaxane. To test this, we synthesized two different peptide pseudorotaxanes to compare to rotaxane **5**. Oligoglycine pseudorotaxane **8** was compared to the GPLG sequence remaining after cleavage by MMP-2 (pseudorotaxane **6**), as well as rotaxane **5** to determine the percent CEST effect of these CB6 complexes (Figure 4.6c). The temperature was held constant at 25 °C, and 10 μM solutions of rotaxane or pseudorotaxane were used. A significant CEST effect was observed for pseudorotaxane **8** and a similar response was observed for pseudorotaxane **6**. This indicates that for both these constructs, CB6 is able to leave the

rotaxane axle and become available for ^{129}Xe@CB6 interactions and subsequent hyperCEST response. A hyperCEST spectrum for intact rotaxane **5** was taken as a baseline to compare the varying peptide sequences, as there was no possibility for CB6 release from the rotaxane axle. No significant CEST effect was observed for the intact rotaxane **5**, which indicates that the CB6 is held in place by the bulky end groups (Figure 4.6c). Comparative CEST responses for 1 μM, 100 nM, and 10 nM free CB6 suggest that on the order of 100 nM CB6 is being released from the oligoglycine (**8**) and GPLG (**6**) pseudorotaxanes.

To evaluate the construct for enzymatic cleavage, rotaxane **5** (100 μM) was incubated at 37 °C with MMP-2 (100 nM) and monitored by reverse phase high performance liquid chromatography (RP-HPLC) and matrix assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF MS). Approximately 60% cleavage of **5** was observed after 24 h by RP-HPLC and post-cleavage products **6** and **7** were observed by MALDI-TOF MS (Figure 4.7a). This confirmed the PLG-LAG cleavage by MMP-2. As a control, peptide **4** was subjected to identical MMP-2 cleavage conditions and was monitored by RP-HPLC and MALDI-TOF MS (Figure 4.7b). Similar rates of proteolysis were observed for peptide **4** as were observed for rotaxane **5**. This indicates that the CB6 supramolecular component does not hinder binding to, or subsequent proteolysis of the rotaxane. Rotaxane **5** cleavage by MMP-2 corresponded with an increase in hyperCEST response at 24 h (Figure 4.8). Before hyperCEST spectra were obtained, the rotaxane-enzyme solution was diluted to 5 μM rotaxane and 5 nM enzyme. After cleavage by MMP-2, xenon was able to interact strongly with CB6, producing a significant increase in hyperCEST response (15% CEST effect) and allowing sensitive detection of MMP-2 via ^{129}Xe hyperCEST.

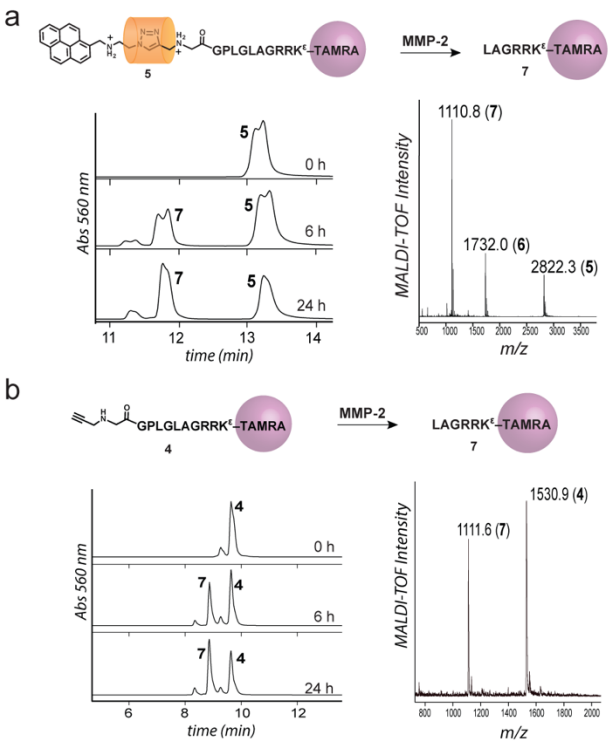


Figure 4.7. MMP-2 recognizes and cleaves CB6 rotaxane **5** at similar rate as peptide **4**. (a) HPLC traces of 100 μM 1 (9.5 min) when exposed to 100 nM MMP-2 at 37 °C. Peptide **4** contains the PLG-LAG MMP-2 recognition sequence. Approximately 60% cleavage to product **7** was observed by 24 h. Double peaks are likely due to the 5,6 isomers of the rhodamine. MALDI-TOF MS confirmed the identity of the cleavage products. (b) Rotaxane **5** contains the same MMP-2 recognition site as peptide **4**. Upon cleavage by MMP-2, CB6 is available for ^{129}Xe host-guest chemistry. HPLC traces of 100 μM **5** (13 min) when exposed to 100 nM MMP-2 at 37 °C revealed approxmiately 60% cleavage to **7** (11.7 min) by 24 h. Further degradation (11.3 min) was also observed. MALDI-TOF MS confirmed both cleavage products **6** and **7**, as well as remaining rotaxane **5**.

These results demonstrate that CB6 rotaxanes can be used as modular probes for protease detection. By tailoring the peptide axle to a protease of interest, these rotaxanes can be tuned for a wide range of disease detection applications. By cleaving the bulky end group, the CB6 can be released from the rotaxane axle and become available for hyperCEST detection. Future work to optimize the rotaxane guest for improved CB6 release upon cleavage could increase the observed

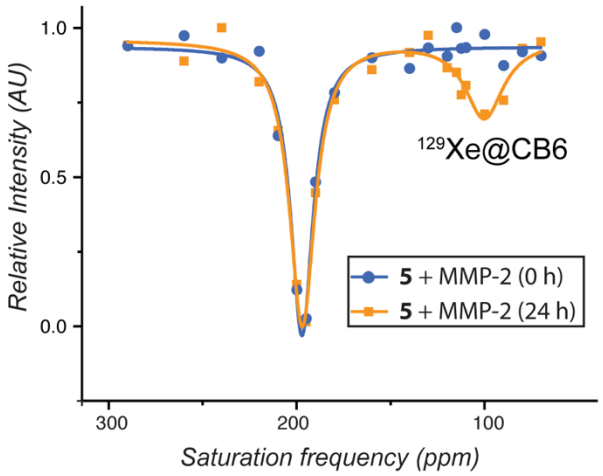


Figure 4.8. CB6 rotaxane **5** functions as a responsive probe for MMP-2 via ^{129}Xe hyperCEST activation. Rotaxane **5** was incubated with MMP-2 for 24 h at 37 °C, and was diluted to 5 μM in ddH_2O before acquiring a hyperCEST spectrum. No detectable hyperCEST response was observed for **5** with MMP-2 immediately after incubation. After 24 h, a significant ^{129}Xe hyperCEST response was observed.

CEST effect of CB6 rotaxanes, allowing the nanomolar concentration range required for enzymatic detection in biologically complex environments.

4.4 CB6 rotaxanes for dual therapeutic and diagnostic applications

Combining drug delivery vehicles with diagnostic probes could offer powerful advantages for disease treatment. These dual-functional probes are an emerging class of imaging agents that allow for the simultaneous tracking and delivery of therapeutics to targeted sites of interest.[28,29] While many of these agents have constantly "on" sensors such as PET tracers or MRI contrast agents, very few such agents contain a turn-on probe that can actively monitor drug release using a non-fluorescent reporter that can be imaged in complex biological environments.[28,29] CB6 rotaxanes can address this gap by taking advantage of the modularity of the CB6 click reaction. By installing a drug molecule as one of the bulky stoppers, and a cancer-targeting group on the other side, it is possible to create a cancer-targeted turn-on probe that is activated upon drug release. Using this rotaxane design, an array of functionality can be established in a very small molecular space.

To synthesize such a probe, epidermal growth factor (EGF) proteins were modified with pipN_3-NHS-ester (Figure 2.1) for subsequent functionalization via CB6 click chemistry (Figure 4.9a). EGF binds to the epidermal growth factor receptors (EGFR) that are overexpressed on a variety of cancer cell types.[30,31] EGF has been previously used as a tumor-targeting motif on nanocarriers for targeted cancer imaging and therapy,[30–32] and therefore functions as an ideal stopper for tumor-targeted CB6 rotaxanes. Doxorubicin (DOX) was chosen as chemotherapeutic cargo as it is a well-established small molecule therapeutic that we have demonstrated can be used in conjunction with CB6 click chemistry (Figure 2.6). DOX-propargylamine derivative **9** was synthesized with a hydrazide linker that is hydrolyzed below pH 6 to release the DOX cargo in the acidic endosome after cellular uptake. EGF-pipN_3 was incubated with 5 equiv of **9** and CB6 for 4 h, leading to the formation of EGF-CB6-DOX conjugates (Figure 4.9). DOX conjugation onto EGF caused the suppression of MS peaks; however, the near-complete disappearance of EGF-pipN_3 peaks in the mass spectrum indicate strong conversion to the EGF-CB6-DOX products. These protein-CB6-drug conjugates will be tested for hyperCEST turn-on, as well as for drug delivery to tumor cells expressing the EGF receptor.

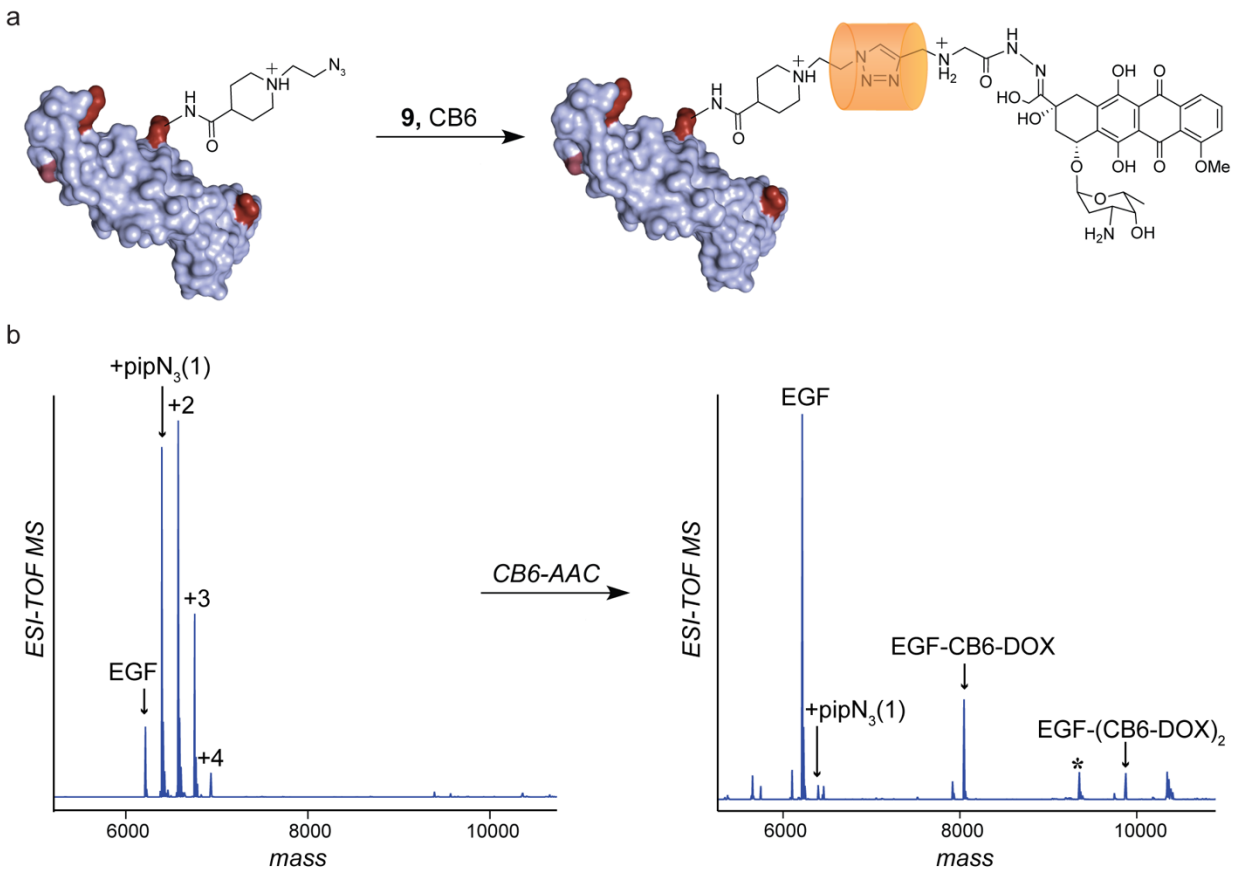


Figure 4.9. Epidermal Growth Factor (EGF) - CB6 - Doxorubicin conjugate for targeted drug delivery and diagnostics. (a) EGF was modified with piperidine-N_3 NHS ester to install click handles at the 3 lysines and N terminus of EGF. Incubation with CB6 and doxorubicin-hydrazide-propargylamine **9** led to EGF-CB6-DOX rotaxane. (b) EGF modification with pipN_3-NHS ester led to 1-4 modifications at reactive amines, as observed by ESI-TOF MS. Subsequent incubation with 10 equiv of DOX-hydrazide-propargylamine **9** and CB6 led to a complete dissapearance of EGF-pipN_3 LC/MS peaks, indicating near-quantitative conversion to the EGF-CB6-DOX products. However, MS suppression was observed for all EGF-CB6-DOX conjugates, leading to the observation of only 1 and 2 modifications. *Some hydrolysis was observed of the DOX from the EGF-$(CB6\text{-}DOX)_2$ conjugate.

4.5 Conclusions

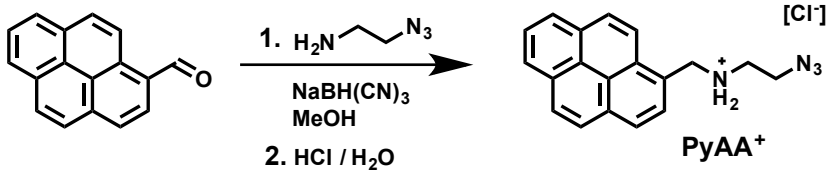
These results demonstrate that CB6 rotaxanes can be used as modular probes for disease detection. We demonstrated that ^{129}Xe@CB6 NMR signals can be completely suppressed by locking CB6 into a rotaxane mechanical bond until a specific cleavage event occurs and releases CB6 to produce a ^{129}Xe@CB6 signal. This activated ^{129}Xe NMR sensor can be easily synthesized and modulated with different cleavable linkers for tunable activation, and diverse rotaxane stoppers with varying functionalities. We demonstrated that incorporation of esters, hydrazides, and proteolytic sequences into the rotaxane axle can lead to different rotaxane probes for diverse applications. Further, we demonstrate that a variety of cargo can be attached as the bulky ends of the rotaxanes, ranging from small molecule dyes and chemotherapeutic drugs, to cancer-targeting proteins. The dense functionality within the small chemical space of the CB6 rotaxanes offer many opportunities in sensitive disease diagnostics and drug delivery.

4.6 Materials and Methods

General Methods

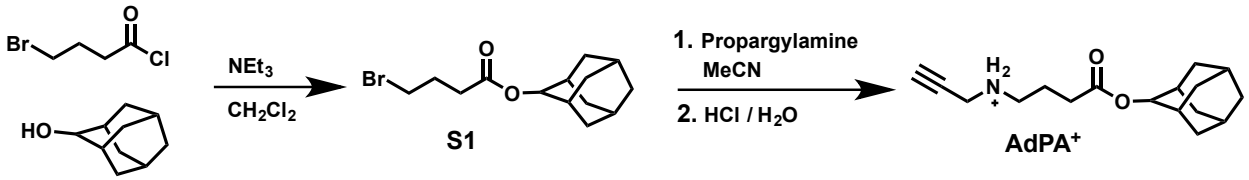
All solvents and reagents, including cucurbit[6]uril hydrate (CB6·XH_2O), β-cyclodextrin hydrate (βCD·XH_2O), and Boc-protected *N*-propargylglycine were purchased from commercial suppliers and used without further purification. **PyAA**$^+$ was prepared according to literature procedures.[33] Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} (E. Merck) and visualized under a UV lamp at 254 nm. Column chromatography was carried out on silica gel 60 (E. Merck, 230–400 mesh). A C-18 column was used for analytical and semi-preparative reverse phase high performance liquid chromatography (RP-HPLC) on an Agilent 1100 Series Capillary LC. Runs were eluted with H_2O/MeCN (0.1 % v/v TFA) and monitored using a UV-Vis detector. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker Avance 400 with working frequencies of 400 MHz for ^{1}H NMR, and 100 MHz for ^{13}C NMR, respectively. Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity, coupling constant (Hz), and integration. Data for ^{13}C NMR are reported in terms of chemical shift. Chemical shifts are referenced to the residual non-deuterated solvents for ^{1}H ($CDCl_3$: δ = 7.27 ppm, CD_3CN: δ = 1.94 ppm, $(CD_3)_2SO$: δ = 2.50 ppm) and ^{13}C ($CDCl_3$: δ = 77.0 ppm, CD_3CN: δ = 118.26 ppm, $(CD_3)_2SO$: δ = 39.52 ppm) nuclei. Matrix assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on a Voyager-DE system (PerSeptive Biosystems, USA) and data were analyzed using Data Explorer software.

Synthetic Procedures



Scheme S4.1. Synthesis of 2-azido-*N*-(pyren-1-ylmethyl)ethanaminium chloride (**PyAA**$^+$).

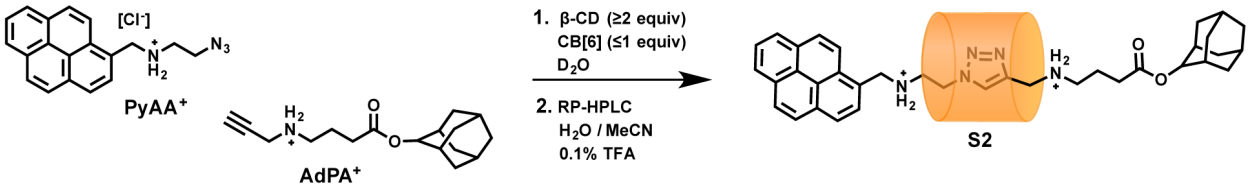
2-azido-N-(pyren-1-ylmethyl)ethanaminium chloride (PyAA$^+$). A portion of 2-Azidoethylamine (79 mg, 0.92 mmol) was added to a solution of 1-pyrenecarboxaldehyde (214 mg, 0.93 mmol) in a mixture of CH_2Cl_2 (5 mL) and MeOH (7 mL) under an atmosphere of nitrogen. After stirring at ambient temperature for 30 min, $NaBH_3CN$ (93 mg, 1.5 mmol) was added in one portion and the reaction continued stirring for 3 d. The solvent was then removed under a stream of N_2. The residue was taken up in CH_2Cl_2 (20 mL) and washed with 5 M HCl (10 mL), leading to the formation of a pale yellow precipitate, which was collected by vacuum filtration and dried under vacuum (170 mg, 54%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 293 K) δ = 9.82 (s, 2H), 8.58 (d, J = 9.0 Hz, 1H), 8.40–8.33 (m, 5H), 8.26 (d, J = 9.0 Hz, 1H), 8.22 (d, J = 9.0 Hz, 1H), 8.13 (t, J = 7.5 Hz, 1H), 4.96 (s, 2H), 3.86 (t, J = 6.0 Hz, 2H), 3.32 (t, J = 6.0 Hz, 2H). ^{13}C NMR (100 MHz, $(CD_3)_2SO$, 293 K) δ = 131.4, 130.7, 130.2, 129.4, 129.3, 128.2, 128.1, 127.3, 126.6, 125.8, 125.6, 125.5, 124.8, 123.9, 123.6, 123.3, 47.1, 46.8, 45.7. HRMS (ESI-TOF-MS): *m/z* calc'd for $C_{19}H_{17}N_4$ $[M-Cl]^+$ 301.1448, observed 301.1443.



Scheme S4.2. Synthesis of propargylammonium-functionalized adamantyl ester stopper **AdPA**$^+$.

2-Adamantyl 4-bromobutanoate (S1). A portion of 4-Bromobutyryl chloride (1.6 mL, 14 mmol) was added to a mixture of 2-adamantanol (650 mg, 4.2 mmol) and triethylamine (1.6 mL, 11 mmol) in CH_2Cl_2/THF (1:1 v/v, 30 mL). The reaction stirred at ambient temperature for 3 h, under an atmosphere of N_2. The reaction mixture was poured into 0.1 M HCl (20 mL) and extracted with CH_2Cl_2 (3 × 20 mL). The organic extracts were combined, dried ($MgSO_4$), filtered, concentrated, and subjected to flash column chromatography on SiO_2, eluting with CH_2Cl_2 to afford the product as a white solid (530 mg, 41%). ^{1}H NMR (400 MHz, $CDCl_3$, 293 K) δ = 4.95 (t, J = 3.0 Hz, 1H), 3.49, (t, J = 7.0 Hz, 2H), 2.54 (t, J = 7.0 Hz, 2H), 2.20 (q, J = 7.0 Hz, 2H), 2.04–1.97 (m, 4H), 1.88–1.82 (m, 4H), 1.80–1.72 (m, 4H), 1.60–1.54 (m, 2H). ^{13}C NMR (100 MHz, $CDCl_3$, 293 K) δ = 171.9, 77.2, 37.3, 36.3, 33.0, 32.8, 31.8, 31.8, 27.9, 27.2, 26.9. HRMS (ESI-TOF-MS): *m/z* calc'd for $C_{14}H_{22}BrO_2$ $[M + H]^+$ 301.0803, observed 301.0795.

4-(2-adamantoxy)-4-oxo-N-(prop-2-ynyl)butan-1-aminium (AdPA$^+$). Propargylamine (600 mg, 11 mmol) was added to a solution of compound **S1** (315 mg, 1.04 mmol) in MeCN (6 mL) and the mixture was stirred at 60 °C for 18 h, under an atmosphere of N_2. The solvent was removed under reduced pressure and the residue was combined with CH_2Cl_2 (4 mL), producing an orange precipitate of predominantly propargylammonium bromide. The precipitate was removed via filtration and the filtrate was diluted with CH_2Cl_2 (8 mL) and washed with 2 M HCl (10 mL). The organic layer was dried under vacuum to afford the product as a pale tan-colored solid (301 mg, 92%). ^{1}H NMR (400 MHz, $CDCl_3$, 293 K) δ = 9.97 (s, 2H), 4.91 (d, J = 3.0 Hz, 1H), 3.91 (td, J = 5.0, 2.5 Hz, 2H), 3.22 (tt, J = 6.0, 5.0 Hz, 2H), 2.64 (t, J = 2.5 Hz, 1H), 2.53 (t, J = 7.0 Hz, 2H), 2.22 (q, J = 7.0 Hz, 2H), 1.99–1.94 (m, 4H), 1.86–1.79 (m, 4H), 1.78–1.70 (m, 4H), 1.57–1.51 (m, 2H). ^{13}C NMR (100 MHz, $CDCl_3$, 293 K) δ = 171.6, 78.3, 77.6, 72.8, 45.6, 37.2, 36.2, 36.1, 31.8, 31.7, 31.7, 31.6, 27.1, 26.9. HRMS (ESI-TOF-MS): *m/z* calc'd for $C_{17}H_{26}NO_2$ $[M - Cl]^+$ 276.1964, observed 276.1954.



Scheme S4.3. Synthesis of **S2**

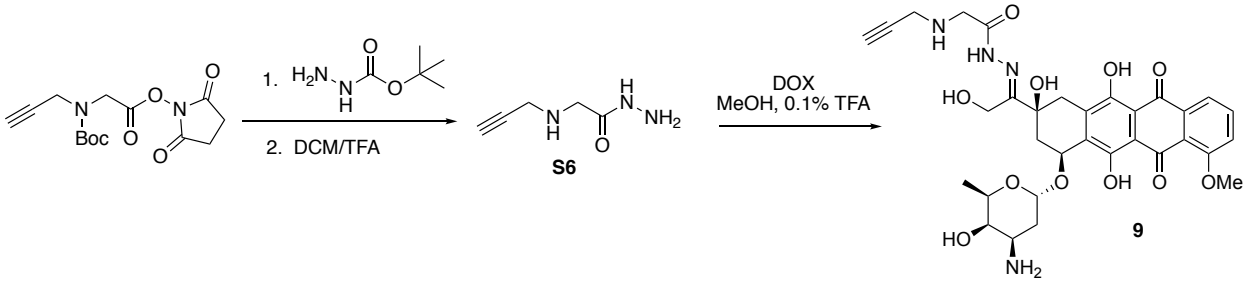
PyAA-CB6-AdPA-rotaxane (S2). Compounds **PyAA**$^+$ (3.5 mg, 10 μmol), **AdPA**$^+$ (3.2 mg, 10 μmol), βCD·XH_2O (40 mg), and CB6·XH_2O (10.8 mg) were mixed in D_2O (2 mL) and stirred at 60 °C for 1 h under an atmosphere of N_2. ^{1}H NMR spectroscopy indicated that **AdPA**$^+$ was converted quantitatively to the βCD-capped CB6-rotaxane. The solution was purified by semi-preparative RP-HPLC, which subsequently removed the βCD caps, eluting in an aqueous gradient of 5% to 95% MeCN/0.1% TFA in H_2O/0.1% TFA over 40 min at a flow rate of 3.0 mL/min. After

removal of the solvent, **S2** was obtained as an off-white powder. ^{1}H NMR (600 MHz, CD_3CN, 293 K) δ = 8.85 (d, J = 9.0 Hz, pyrene, 1H), 8.41–8.33 (m, pyrene, 5H), 8.27 (d, J = 9.0 Hz, pyrene, 1H), 8.23 (d, J = 9.0 Hz, pyrene, 1H), 8.14 (t, J = 8.0 Hz, pyrene, 1H), 8.03 (br s, $–NH_2^+–$, 2H), 7.10 (br s, $–NH_2^+–$, 2H), 6.53 (s, triazole, 1H), 5.70 (d, J = 15 Hz, CB6, 6H), 5.69 (d, J = 15 Hz, CB6, 6H), 5.31 (t, J = 5 Hz, 2H), 5.27 (s, CB6, 12H), 4.92 (t, J = 2.5 Hz, 1H), 4.26–4.23 (m, 2H), 4.06–4.02 (m, 2H), 3.87–3.81 (m, 2H), 3.31–3.25 (m, 2H), 2.74 (t, J = 7.5 Hz, 2H), 2.36 (q, J = 7.5 Hz, 2H), 2.11–2.05 (m, 2H), 2.00–1.75 (m, adamantyl, 12H), 1.59–1.54 (m, adamantyl, 2H). ^{13}C NMR (150 MHz, CD_3CN, 293 K) δ = 155.7, 132.3, 129.4, 128.6, 127.3, 126.7, 126.1, 124.9, 123.0, 82.5, 81.1, 80.2, 76.8, 51.3, 51.0, 47.3, 37.0, 35.9, 31.8, 31.5, 27.2, 27.0. HRMS (MALDI-TOF-MS): *m/z* calcd for $C_{72}H_{80}N_{29}O_{14}$ $[M+H]^+$ 1575.59, observed 1573.06.

*N-propargylglycine-GPLGLAGRRK-Rhodamine (**4**)*. Peptide **4** was synthesized using standard Fmoc solid phase peptide synthesis (SPPS) on Tentagel S-RAM resin (Rapp Polymere, Tuebingen, Germany). Side chain protecting groups used were Arg(pbf) and Lys(dde). Deprotection of the Fmoc groups was performed with two 10 min incubations in a 30% v/v piperidine in dimethylformamide (DMF) solution. Couplings were carried out using 10 equiv of amino acid with 10 equiv of 2-(6-chloro-1-*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), and 20 equiv of *N*,*N*-diisopropylethylamine (DIPEA) in DMF for 20 min. The N-terminus of the peptide was modified on resin with *N*-propargylglycine using the same coupling conditions. After *N*-propargylglycine addition, dde side-chain deprotection was accomplished using three 5 min incubations with 2% v/v hydrazine-hydrate in DMF solution. After Lys(dde) deprotection, the peptide-resin was incubated overnight with 1.13 equiv of 5-carboxy-tetramethylrhodamine *N*-succinimidyl ester (Sigma Aldrich. St. Louis, MO) and 2.26 equiv of DIPEA in DMF. Peptide **4** was cleaved from the resin after incubation with a solution of 95:2.5:2.5 TFA:TIPS:H_2O for 2 h. Excess TFA was removed under a stream of N_2 and the peptide was precipitated in cold diethyl ether. The precipitate was redissolved in H_2O and purified by RP-HPLC, eluting in an aqueous gradient 5% to 95% MeCN/0.1% TFA in H_2O/0.1% TFA over 45 min at a flow rate of 3.0 mL/min. After lyophilization, peptide **4** was obtained as a deep pink powder (3.8 mg). HRMS (MALDI-TOF-MS): *m/z* calcd for $C_{74}H_{108}N_{21}O_{15}$ $[M+H]^+$ 1530.83, observed 1529.86.

*N-propargylglycine-GGGG (**S3**) and N-propargylglycine-GPLG (**S4**)*. Peptides **S3** and **S4** were synthesized using standard Fmoc solid phase peptide synthesis (SPPS) on Gly-Wang resin (Anaspec, Fremont, CA). Deprotection of the Fmoc groups was performed with two 10 min incubations in a 30% v/v piperidine in dimethylformamide (DMF) solution. Couplings were carried out using 10 equiv of amino acid with 10 equiv of 2-(6-chloro-1-*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), and 20 equiv of *N*,*N*-diisopropylethylamine (DIPEA) in DMF for 20 min. The N-terminus of the peptide was modified on resin with *N*-propargylglycine using the same coupling conditions. Peptides were cleaved from the resin after incubation with a solution of 95:2.5:2.5 TFA:TIPS:H_2O for 2-3 h. Excess TFA was removed under a stream of N_2 and the peptide was precipitated in cold diethyl ether. The precipitate was redissolved in H_2O and purified by RP-HPLC, eluting in an aqueous gradient 5% to 95% MeCN/0.1% TFA in H_2O/0.1% TFA over 45 min at a flow rate of 3.0 mL/min. After lyophilization, peptides were obtained as off-white powders (**S3,** 3.5 mg; **S4,** 3.8 mg). HRMS (MALDI-TOF-MS): *m/z* calcd for **S3** $C_{13}H_{20}N_5O_6$ $[M+H]^+$ 342.13, observed 342.4. *m/z* calcd for **S4** $C_{20}H_{32}N_5O_6$ $[M+H]^+$ 438.50, observed 438.66.

*PyAA-CB6-peptide pseudorotaxanes **6** and **8**, and rotaxane **5***. Compounds **PyAA**$^+$ (1 equiv, 2-10 μmol) and βCD·XH_2O (1.3 equiv) were mixed in H_2O and stirred at 60 °C to allow for complexation. Once **PyAA**$^+$ dissolved, the solution was added to one of peptides **4, S4,** or **S5** (1 equiv), and the solution was then transferred to an Eppendorf charged with CB6·XH_2O (1 equiv). The solution was stirred at 60 °C for 6 h removed from light, and the reaction was monitored by MALDI-TOF MS. The solution was then filtered with a 0.2 μm spin filter to remove unreacted CB6, and the product was purified by semi-preparative RP-HPLC, eluting in an aqueous gradient 10% to 90% MeCN/0.1% TFA in H_2O/0.1% TFA over 35 min at a flow rate of 3.0 mL/min. RP-HPLC removed βCD from the pyrene stopper of the rotaxanes. After lyophilization, rotaxanes **8** (2.8 mg, 18% yield) and **6** (3.1 mg (31% yield) were obtained as off-white powders and **5** (4.9 mg, 84% yield) was obtained as a deep pink powder. HRMS (MALDI-TOF-MS): **5.** *m/z* calcd for $C_{129}H_{159}N_{49}O_{27}$ $[M+H]^+$ 2825.97, observed 2825.38. **6.** *m/z* calcd for $C_{75}H_{84}N_{33}O_{18}$ $[M+H]^+$ 1734.64, observed 1734.04. **8.** *m/z* calcd for $C_{68}H_{72}N_{33}O_{18}$ $[M+H]^+$ 1638.27, observed 1638.65.



Scheme S4.4. Synthesis of **9.**

*Propargylglycine-hydrazide **S6.*** Propargyl-glycine-NHS ester was synthesized as described in section 2.9. t-Butyl carbazate (61 mg, 0.46 mmol) was dissolved in 2 mL of DMF. To that solution was added triethylamine (92.9 mg, 0.92 mmol). NHS ester (72 mg, 0.23 mmol) was added to the solution of t-butyl carbazate and the solution was stirred at ambient temperature overnight. The solvent was diluted with ddH_2O (10 mL) and extracted 3x with EtOAc. The organic layers were collected and washed with ddH_2O (2x) and brine (1x). The organic layers were then dried with Na_2SO_4 and the solvent was reduced *in vacuo.* The product was purified by column chromatography using a Biotage flash purification system. The solvent was then evaporated under reduced pressure and the product was resuspended in 50:50 DCM/TFA (3 mL) and stirred at room temperature for 1 h. The solvent was then evaporated and the deprotected product was resuspended in ddH2O and lyophilized to yield product **S6** as a tan oil (60 mg, 80%). ^{1}H NMR (600 MHz, $CDCl_3$) 4.00 (d, J = 2.6 Hz, 2H), 3.91 (s, 2H), 3.27 (t, J = 2.6 Hz, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ = 164.33, 78.27, 72.67, 45.25, 35.83.

*Doxorubicin-hydrazide-propargylamine **9.*** Doxorubicin-HCl (13.3 mg, 0.023 mmol) was dissolved in 2 mL of anhydrous MeOH and was combined with **S6**•TFA (16.6 mg, 0.069 mmol) dissolved in 1 mL of MeOH. To the solution was added 2 drops of TFA. The solution was stirred protected from light for 24 h. The solution as then evaporated under reduced pressure and the product was precipitated with EtOAc (3x 20 mL). The product was dissolved in ddH_2O and

lyophilized to yield **9** as a dark red powder (3 mg). HRMS (MALDI-TOF-MS): **9.** *m/z* calcd for $C_{32}H_{36}N_4O_{11}$ $[M+H]^+$ 653.7, observed 675.2 (+Na).

Protease-induced cleavage procedures

*MMP-2-induced cleavage of peptide **4**, monitored by RP-HPLC and MALDI-TOF MS.* To a solution of 100 μM **4** in 60 μL MMP-2 buffer (40 mM TRIS-Cl, 8 mM Zn^{2+}, 8 mM Ca^{2+}, 8 mM $Na_2PO_4/NaPO_4$, 0.04% Brij-35, pH 7) was added MMP-2 (active catalytic domain, Enzo Life Sciences, Farmingdale, NY) as a 0.1 mg/mL solution for a final concentration of 100 nM MMP-2, and the solution was incubated at 37 °C. Aliquots (15 μL) were taken at each timepoint of interest and analyzed by RP-HPLC eluting with an aqueous gradient of 10% to 90% MeCN/0.1% TFA in H_2O/0.1% TFA over 20 min at a flow rate of 1.0 mL/min, monitoring at the Abs_{560} of rhodamine. Cleavage product **7** was confirmed by MALDI-TOF MS. *m/z* calcd for **7** $C_{54}H_{79}N_{16}O_{10}$ $[M+H]^+$ 1112.3, observed 1111.6.

*MMP-2-induced cleavage of rotaxane **5**, monitored by RP-HPLC, MALDI-TOF MS, and ^{129}Xe hyperCEST.* To a solution of 100 μM **5** in 150 μL MMP-2 buffer (40 mM TRIS-Cl, 8 mM Zn^{2+}, 8 mM Ca^{2+}, 8 mM $Na_2PO_4/NaPO_4$, 0.04% Brij-35, pH 7) was added MMP-2 (active catalytic domain, Enzo Life Sciences, Farmingdale, NY) as a 0.1 mg/mL solution for a final concentration of 100 nM MMP-2, and the solution was incubated at 37 °C. Aliquots (25 μL) were taken at each timepoint of interest and analyzed by RP-HPLC eluting with an aqueous gradient of 10% to 90% MeCN/0.1% TFA in H_2O/0.1% TFA over 20 min at a flow rate of 0.5 mL/min, monitoring at the Abs_{560} of rhodamine. Cleavage products **6** and **7** were confirmed by MALDI-TOF MS. At 24 h, the solution was diluted in ddH_2O to 5 μM **5** and 5 nM MMP-2, and a ^{129}Xe hyperCEST spectrum was obtained. HRMS (MALDI-TOF-MS): *m/z* calcd for **6** $C_{75}H_{84}N_{33}O_{18}$ $[M+H]^+$ 1734.64, observed 1732.0. HRMS (MALDI-TOF-MS): *m/z* calcd for **7** $C_{54}H_{79}N_{16}O_{10}$ $[M+H]^+$ 1112.3, observed 1110.8.

Xenon hyperCEST NMR.

Xenon polarization was achieved using a home built spin-exchange optical-pumping setup resulting in a 10% polarization of a xenon gas mixture (2 % Xe, 10% N_2, 88% He). The hyperpolarized gas was bubbled directly into a 5 mm phantom containing the solution of interest for 20 s then left to settle for 2 s. The sample was held at 3.4 atm and 25 °C throughout. A 9.4 T (400 MHz) Varian VNMRS console was used for all hyperCEST experiments with optimized saturation power and duration for each sample. A standard hyperCEST pulse sequence was used sweeping the saturation frequency, at 20 dB, in 1000 Hz increments 7000 Hz to 29000 Hz, also including a measurement at 11250 and 11500 Hz where CB6 is typically observed, and at 19500 Hz, corresponding to the center of the $^{129}Xe@H_2O$ peak covering a 200 ppm range in total. The saturation pulse length was 4 s for three cycles at 20 dB. For % CEST effect data, an on/off saturation experiment was conducted saturating at the center of the CB6 peak then switching to 29000 Hz on the other side of the $^{129}Xe@H_2O$ peak alternating 8 times for a total of 4 on saturation and 4 off saturation values. Data processing was carried out using MATLAB. FIDs were zero-filled to 16384 points, baseline was corrected, apodized with an 11 Hz exponential, and a Fourier

transform was performed. Each $^{129}Xe@H_2O$ areas in the spectra were integrated and the contrast of each spectrum was compared between the maximum and minimum area in each data series. Each profile was fitted with Lorentz profile using ORIGINLAB.

The percent CEST effect is obtained by alternating the on and off resonance saturation frequencies and observing the integrated area of the dissolved xenon peak. For CB6 the on resonance saturation is around 11250 Hz so the off resonance saturation is placed equidistant on the other side of the water peak at 27750 Hz. The placement of the saturation peaks is such that if the saturation pulse for the on resonance peak has any overlap with the dissolved water peak, it will have the same overlap during the off resonance saturation. For this study the on/off cycle was repeated 4 times for each rotaxane and pseudorotaxane then the % CEST effect was determined bynormalizing the spectra and taking the signal from the on resonance and dividing it by the off resonance and subtracting from one, %CEST effect = $(1 - S_{on}/S_{off})$, for each cycle. The signal for xenon-NMR experiments is generated external to the sample of interest. This causes fluctuations in the amount of polarized xenon dissolved in water for each scan. By taking 25 off resonance scans the fluctuation is signal due to xenon polarization and bubbling was determined to be +/- 3.5% of the average $Xe@H_2O$ signal so a percent CEST effect must be greater than 7% for it to be considered significant. For the locked rotaxane the percent CEST effect was determined to be less then 4% and thus can be concluded that no measurable CEST is taking place.

4.7 References

(1) Major, J. L.; Meade, T. J. Bioresponsive, Cell-Penetrating, and Multimeric MR Contrast Agents. *Acc. Chem. Res.* **2009**, *42* (7), 893–903.

(2) Que, E. L.; New, E. J.; Chang, C. J. A Cell-Permeable Gadolinium Contrast Agent for Magnetic Resonance Imaging of Copper in a Menkes Disease Model. *Chem. Sci.* **2012**, *3* (6), 1829.

(3) Sowers, M. A.; McCombs, J. R.; Wang, Y.; Paletta, J. T.; Morton, S. W.; Dreaden, E. C.; Boska, M. D.; Ottaviani, M. F.; Hammond, P. T.; Rajca, A.; et al. Redox-Responsive Branched-Bottlebrush Polymers for in Vivo MRI and Fluorescence Imaging. *Nat. Commun.* **2014**, *5*, 5460.

(4) Albert, M. S.; Cates, G. D.; Driehuys, B.; Happer, W.; Saam, B.; Springer, C. S.; Wishnia, A. Biological Magnetic Resonance Imaging Using Laser-Polarized 129Xe. *Nature* **1994**, *370* (6486), 199–201.

(5) Oros, A.-M.; Shah, N. J. Hyperpolarized Xenon in NMR and MRI. *Phys. Med. Biol.* **2004**, *49* (20), 105–153.

(6) Spence, M. M.; Rubin, S. M.; Dimitrov, I. E.; Ruiz, E. J.; Wemmer, D. E.; Pines, A.; Yao, S. Q.; Tian, F.; Schultz, P. G. Functionalized Xenon as a Biosensor. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (19), 10654–10657.

(7) Lowery, T. J.; Garcia, S.; Chavez, L.; Ruiz, E. J.; Wu, T.; Brotin, T.; Dutasta, J.-P.; King, D. S.; Schultz, P. G.; Pines, A.; et al. Optimization of Xenon Biosensors for Detection of Protein Interactions. *ChemBioChem* **2006**, *7* (1), 65–73.

(8) Schroder, L.; Lowery, T. J.; Hilty, C.; Wemmer, D. E.; Pines, a. Molecular Imaging Using a Targeted Magnetic Resonance Hyperpolarized Biosensor. *Science (80-.).* **2006**, *314* (5798), 446.

(9) Palaniappan, K. K.; Ramirez, R. M.; Bajaj, V. S.; Wemmer, D. E.; Pines, A.; Francis, M. B. Molecular Imaging of Cancer Cells Using a Bacteriophage-Based 129Xe NMR Biosensor. *Angew. Chemie Int. Ed.* **2013**, *52* (18), 4849–4853.

(10) Shapiro, M. G.; Ramirez, R. M.; Sperling, L. J.; Sun, G.; Sun, J.; Pines, A.; Schaffer, D. V.; Bajaj, V. S. Genetically Encoded Reporters for Hyperpolarized Xenon Magnetic Resonance Imaging. *Nat. Chem.* **2014**, *6* (April), 629–634.

(11) Rose, H. M.; Witte, C.; Rossella, F.; Klippel, S.; Freund, C.; Schröder, L. Development of an Antibody-Based, Modular Biosensor for 129 Xe NMR Molecular Imaging of Cells at Nanomolar Concentrations. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (32), 11697–11702.
(12) Garimella, P. D.; Meldrum, T.; Witus, L. S.; Smith, M.; Bajaj, V. S.; Wemmer, D. E.; Francis, M. B.; Pines, A. Hyperpolarized Xenon-Based Molecular Sensors for Label-Free Detection of Analytes. *J. Am. Chem. Soc.* **2014**, *136* (1), 164–168.
(13) Witte, C.; Martos, V.; Rose, H. M.; Reinke, S.; Klippel, S.; Schröder, L.; Hackenberger, C. P. R. Live-Cell MRI with Xenon Hyper-CEST Biosensors Targeted to Metabolically Labeled Cell-Surface Glycans. *Angew. Chemie Int. Ed.* **2015**, *54* (9), 2806–2810.
(14) Masson, E.; Ling, X.; Joseph, R.; Kyeremeh-Mensah, L.; Lu, X. Cucurbituril Chemistry: A Tale of Supramolecular Success. *RSC Adv.* **2012**, *2*, 1213–1247.
(15) Wang, Y.; Dmochowski, I. J. Cucurbit[6]uril Is an Ultrasensitive 129 Xe NMR Contrast Agent. *Chem. Commun.* **2015**, *51*, 8982–8985.
(16) Kunth, M.; Witte, C.; Hennig, A.; Schröder, L. Identification, Classification, and Signal Amplification Capabilities of High-Turnover Gas Binding Hosts in Ultra-Sensitive NMR. *Chem. Sci.* **2015**, *6*, 6069–6075.
(17) Ayhan, M. M.; Karoui, H.; Hardy, M.; Rockenbauer, A.; Charles, L.; Rosas, R.; Udachin, K.; Tordo, P.; Bardelang, D.; Ouari, O. Comprehensive Synthesis of Monohydroxy–Cucurbit[*N*]Urils (*N* = 5, 6, 7, 8): High Purity and High Conversions. *J. Am. Chem. Soc.* **2015**, *137* (32), 10238–10245.
(18) Schnurr, M.; Sloniec-Myszk, J.; Döpfert, J.; Schröder, L.; Hennig, A. Supramolecular Assays for Mapping Enzyme Activity by Displacement-Triggered Change in Hyperpolarized 129 Xe Magnetization Transfer NMR Spectroscopy. *Angew. Chemie Int. Ed.* **2015**, *54* (45), 13444–13447.
(19) Hou, X.; Ke, C.; Fraser Stoddart, J. Cooperative Capture Synthesis: Yet Another Playground for Copper-Free Click Chemistry. *Chem. Soc. Rev.* **2016**, *45* (14), 3766–3780.
(20) Mock, W. L.; Irra, T. A.; Wepsiec, J. P.; Adhya, M. Catalysis by Cucurbituril. The Significance of Bound-Substrate Destabilization for Induced Triazole Formation. *J. Org. Chem.* **1989**, *54* (22), 5302–5308.
(21) Wang, Y.; Roose, B. W.; Philbin, J. P.; Doman, J. L.; Dmochowski, I. J. Programming A Molecular Relay for Ultrasensitive Biodetection through ^{129} Xe NMR. *Angew. Chemie Int. Ed.* **2016**, *55* (5), 1733–1736.
(22) Jiang, T.; Olson, E. S.; Nguyen, Q. T.; Roy, M.; Jennings, P. A.; Tsien, R. Y. Tumor Imaging by Means of Proteolytic Activation of Cell-Penetrating Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (51), 17867–17872.
(23) Vartak, D. G.; Gemeinhart, R. A. Matrix Metalloproteases: Underutilized Targets for Drug Delivery. *J. Drug Target.* **2007**, *15* (1), 1–20.
(24) Aguilera, T. A.; Olson, E. S.; Timmers, M. M.; Jiang, T.; Tsien, R. Y. Systemic in Vivo Distribution of Activatable Cell Penetrating Peptides Is Superior to that of Cell Penetrating Peptides. *Integr. Biol.* **2009**, *1* (5–6), 371–381.
(25) Nguyen, M. M.; Carlini, A. S.; Chien, M.-P.; Sonnenberg, S.; Luo, C.; Braden, R. L.; Osborn, K. G.; Li, Y.; Gianneschi, N. C.; Christman, K. L. Enzyme-Responsive Nanoparticles for Targeted Accumulation and Prolonged Retention in Heart Tissue after Myocardial Infarction. *Adv. Mater.* **2015**, *27* (37), 5547–5552.
(26) Olson, E. S.; Jiang, T.; Aguilera, T. A.; Nguyen, Q. T.; Ellies, L. G.; Scadeng, M.; Tsien, R. Y. Activatable Cell Penetrating Peptides Linked to Nanoparticles as Dual Probes for in Vivo Fluorescence and MR Imaging of Proteases. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (9), 4311–4316.
(27) Wei, Q.; Seward, G. K.; Aru Hill, P.; Patton, B.; Dimitrov, I. E.; Kuzma, N. N.; Dmochowski, I. J. Designing 129 Xe NMR Biosensors for Matrix Metalloproteinase Detection. *J. Am. Chem. Soc.* **2006**, *128* (40), 13274–13283.
(28) Lammers, T.; Aime, S.; Hennink, W. E.; Storm, G.; Kiessling, F. Theranostic Nanomedicine. *Acc. Chem. Res.* **2011**, *44* (10), 1029–1038.

(29) Kelkar, S. S.; Reineke, T. M. Theranostics: Combining Imaging and Therapy. *Bioconjug. Chem.* **2011**, *22* (10), 1879–1903.
(30) Master, A. M.; Sen Gupta, A. EGF Receptor-Targeted Nanocarriers for Enhanced Cancer Treatment. *Nanomedicine* **2012**, *7* (12), 1895–1906.
(31) Bertrand, N.; Wu, J.; Xu, X.; Kamaly, N.; Farokhzad, O. C. Cancer Nanotechnology: The Impact of Passive and Active Targeting in the Era of Modern Cancer Biology. *Adv. Drug Deliv. Rev.* **2014**, *66*, 2–25.
(32) ElSohly, A. M.; MacDonald, J. I.; Hentzen, N. B.; Aanei, I. L.; El Muslemany, K. M.; Francis, M. B. *Ortho* -Methoxyphenols as Convenient Oxidative Bioconjugation Reagents with Application to Site-Selective Heterobifunctional Cross-Linkers. *J. Am. Chem. Soc.* **2017**, *139* (10), 3767–3773.
(33) Hou, X.; Ke, C.; Bruns, C. J.; McGonigal, P. R.; Pettman, R. B.; Stoddart, J. F. Tunable Solid-State Fluorescent Materials for Supramolecular Encryption. *Nat. Commun.* **2015**, *6*, 6884.

4.8 Additional Figures

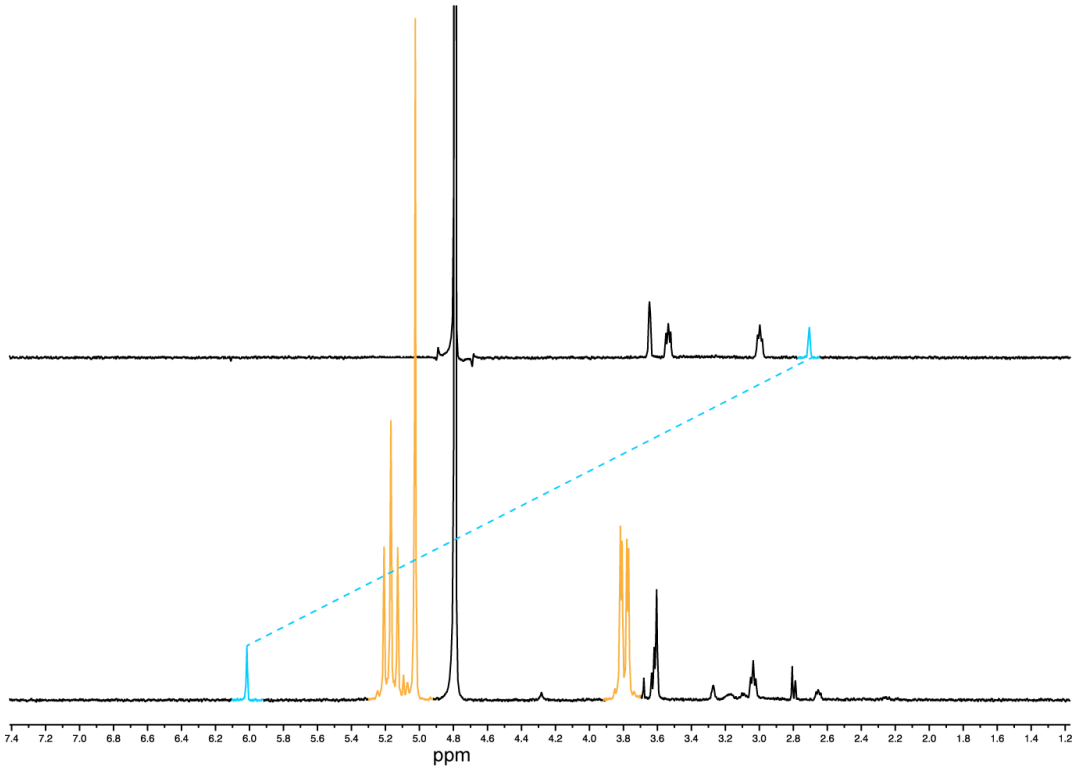


Figure S1. ^{1}H NMR of co-incubated alkyne and azide (top) and complex **1** product (bottom). Blue indicates alkyne (top) and triazole (bottom) chemical shifts, and orange indicates CB6 chemical shifts.

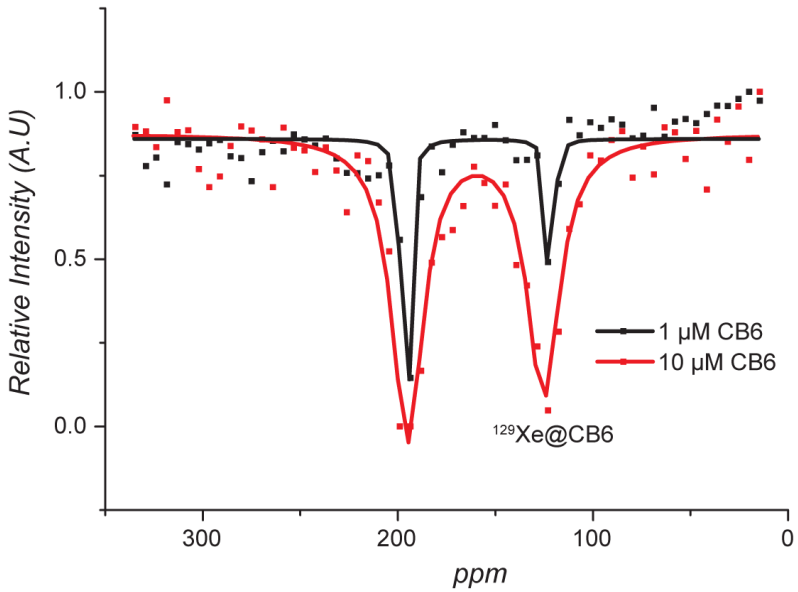


Figure S2. ^{129}Xe hyperCEST NMR spectra of 1 μM (black) and 10 μM (red) CB6 in water. Profiles were fit with Lorentz profile using ORIGINLAB.

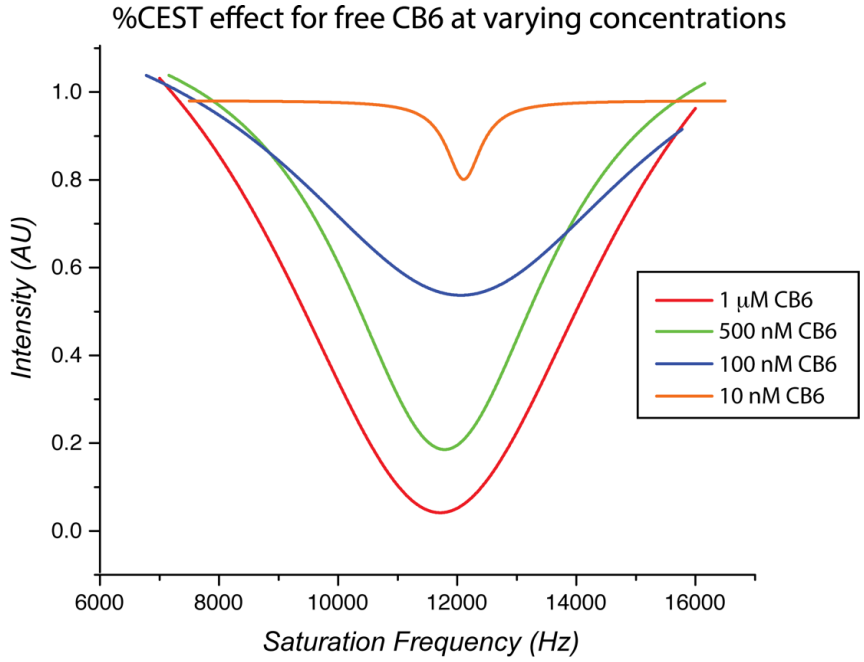


Figure S3. ^{129}Xe hyperCEST spectra of free CB6 at 1000, 500, 100, and 10 nM in NaPhos pH 7.

Chapter 5

Morphological effects of protein-based nanomaterials for drug delivery

ABSTRACT: Current approaches to nanoscale therapeutic delivery rely on the attachment of a drug of interest to a nanomaterial scaffold that is capable of releasing the drug selectively in a tumor environment. One class of nanocarriers receiving significant attention is protein nanomaterials, which are biodegradable, homogenous in morphology, and can be equipped with multiple functional handles for drug attachment. While most protein-based nanocarriers are spherical in morphology, recent research has revealed that non-spherical nanomaterials may have favorable tumor uptake when compared to their spherical counterparts. It is therefore important to expand the number of non-spherical protein-based nanocarriers that are available. Herein, we report the development of a self-assembling nanoscale disk derived from a double arginine mutant of recombinantly expressed tobacco mosaic virus coat protein (RR-TMV). RR-TMV disks display highly stable double disk assembly states. These RR-TMV disks were functionalized with the chemotherapy drug doxorubicin (DOX) and further modified with polyethylene glycol (PEG) for improved solubility. RR-TMV$_{DOX\text{-}PEG}$ displayed cytotoxic properties similar to DOX alone when incubated with U87MG glioblastoma cells, but unmodified RR-TMV did not cause any cytotoxicity. RR-TMV disks were further tested *in vivo* as part of a panel of protein-based nanomaterials of varying shape and size to determine how the morphology of protein-based nanomaterials affect their tumor penetration and retention in the context of drug delivery and cancer imaging.

Portions of the work described in this chapter have been reported in the following publication:
JA Finbloom, K Han, IL Aanei, EC Hartman, DT Finley, MT Dedeo, M Fishman, KH Downing, MB Francis. Stable disk assemblies of a tobacco mosaic virus mutant as nanoscale scaffolds for applications in drug delivery. *Bioconjugate Chem.* 2016. 27(10); 2480-85

5.1 Protein-based nanomaterials for drug delivery

Nanomaterials are attracting significant attention for their applications in drug delivery, disease imaging, and diagnostics.[1] Ideal properties for such agents include uniform sizes and shapes, biodegradability, and multiple sets of functional handles for chemical manipulation.[1–3] Self-assembling protein nanomaterials derived from viruses meet many of these qualifications. As such, intact virus nanoparticles have been functionalized for applications in drug delivery *in vivo*, typically through the covalent attachment of drugs or imaging agents onto the coat proteins.[4–9] However, the injection of replication-competent viruses into subjects may limit their clinical appeal. For that reason, recombinantly expressed viral capsids that do not contain genetic material for propagation could be uniquely suited for advanced drug delivery and imaging applications.

Many labs have used virus-like particles for both drug delivery and the targeted imaging of cancer and other diseases.[10] These efforts have frequently relied on spherical virus-like particles, such as the MS2 bacteriophage capsid,[8,9] the cowpea mosaic virus,[11,12] the P22 virus,[13] and others.[5,7] However, several reports have demonstrated that spherical nanomaterials had lower degrees of tumor accumulation when compared to non-spherical nanomaterials.[10,12,14–18] As important counterparts, rod-shaped virus nanoparticles have been developed based on the tobacco mosaic virus (TMV),[19–21] filamentous phage,[22,23] potato virus X,[12] and others; and differential tumor accumulations have indeed been observed.[12,17] Further expanding the shape library of protein nanomaterials would allow additional understanding of the morphological considerations that are necessary for successful nanomaterial-based approaches to drug delivery and imaging.

Developing new morphologies for these applications would be particularly beneficial if the protein nanomaterial were recombinantly expressed. Along these lines, Baker and coworkers have developed computational approaches to the *de novo* design of a variety of self-assembling protein nanomaterials, including a recent report on a hyperstable icosahedron that is amenable to subsequent protein engineering.[24] However, there are often challenges when non-spherical virus-based protein nanomaterials are expressed without the genetic material that functions as a backbone to the assembly architecture. They are frequently polydisperse in size and shape, and change their assembly behavior in response to different conditions such as pH and ionic strength.[20,25] These assembly dynamics are especially true for variants of the tobacco mosaic virus (TMV). Replicating TMV nanoparticles have previously been used for drug delivery and imaging applications;[19,26] however, when recombinantly expressed, TMV proteins can assemble into a variety of states, including single monomers, 20S double disks, and helical rods that can be microns in length.[20] It would be desirable to develop a new and stable morphology to complement the spherical and rod-shaped protein nanomaterials that already exist.

Disk-shaped nanomaterials are a relatively unexplored but potentially interesting morphology for drug delivery applications. Disks could offer advantages over other morphologies, as they are ideally suited to pack into small spaces[27] and could allow enhanced penetration into tumor tissues. Preliminary studies with disk-shaped polymer nanoparticles in the size regime of hundreds of nanometers have indicated favorable tumor-homing and uptake properties.[16,28–30] However, very little research has been conducted with disk-shaped nanoparticles that are within the smaller size regime required to take advantage of the enhanced permeability and retention (EPR) effect,[1] nor has there been significant research on disk-shaped protein nanomaterials in the context of drug delivery. This is principally due to the lack of well-defined discoidal objects of this size.

Our lab reported one of the few examples of a self-assembled disk from protein monomers using a circular permutant of TMV.[25] While these cpTMV disks did have increased stability over

typical recombinantly expressed TMV, disassembly of the disks into monomers and further assembly into rods were achieved under certain conditions. Herein, we report a double Arg mutant of TMV (RR-TMV) that selectively assembles into double disks that remain uniform despite variations in environmental conditions. These RR-TMV disks were site-specifically functionalized at a non-native cysteine with doxorubicin and at the N terminus with polyethylene glycol for drug delivery applications, and the new agents were characterized for their cell uptake, toxicity potential, and *in vivo* drug delivery potential as part of a panel of protein-based nanomaterials.

5.2 Expression and characterization of RR-TMV disks

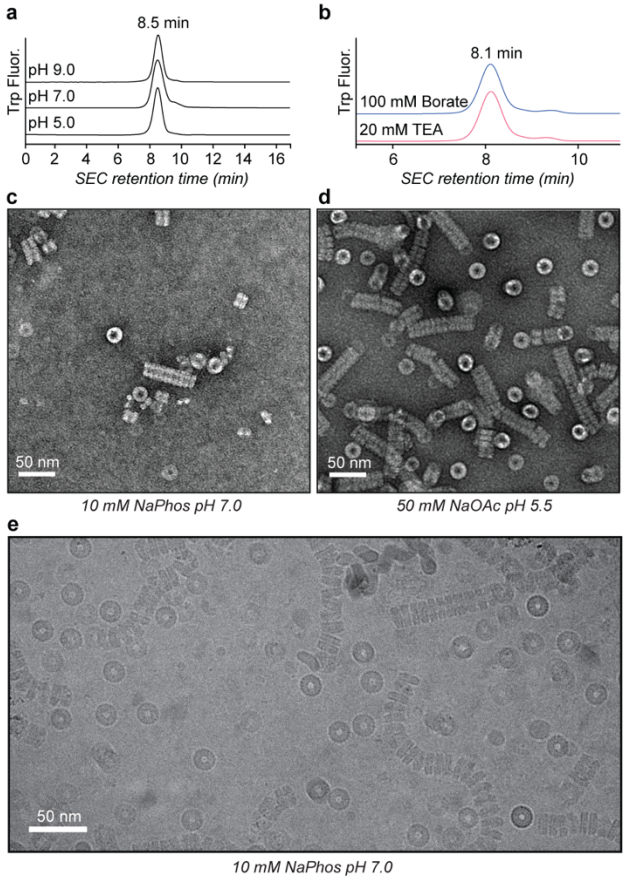


Figure 5.1. Assembly characterization of RR-TMV disks. (a) RR-TMV disks were subjected to HPLC-SEC at varying pH values (all in 10 mM sodium phosphate buffer). Disk assemblies (8-9 min retention time) were observed for all pH conditions tested. (b) HPLC-SEC traces of RR-TMV in borate and TEA buffers at pH 8. Disk assemblies were the dominant species observed. TEM images of RR-TMV in (c) NaPhos and (d) NaOAc buffers. (e) Cryo-EM image of RR-TMV in NaPhos. Disks and short stacks of disks were observed under both conditions. Larger versions of these images appear as Figures S5.1-S5.3.

A mutated variant of the TMV coat protein was expressed and purified in which the two native lysines at positions K53 and K68 were changed to arginines, and a non-native lysine was introduced at position T104. This mutation was originally introduced to allow the site-specific modification of TMV proteins on the T104K position with activated ester chemistry. The RR-TMV mutant was expressed in *E. coli* and purified by anion exchange chromatography to provide high yields (>100 mg/L broth) of pure protein. Upon characterization of these RR-TMV mutants, it was apparent that the standard assembly state of recombinantly expressed TMV was perturbed such that only disks were visible in the RR-TMV population (Figure 5.1). This is in contrast to the combination of rods, disks, and monomers typically observed for recombinantly expressed TMV.[25]

RR-TMV was characterized by size exclusion chromatography (HPLC-SEC) under varying pH (Figure 5.1a) and buffer (Figure 5.1b) conditions. Under all conditions tested, only SEC elution times of 8-9 min were observed, corresponding to the known double disk assembly of TMV.[20,25] TEA and borate buffers (both at pH 8) were evaluated, as they have been previously shown to promote disassembly of TMV disks into monomers;[25] however, disks were predominantly observed by SEC for both buffers. Transmission electron

microscopy (TEM) was conducted on RR-TMV in both phosphate buffer and 50 mM NaOAc pH 5.5, as NaOAc has been reported to drive rod assembly for TMV proteins.[20,25] Under both conditions, only RR-TMV disks and short stacks of disks were observed (Figure 5.1c,d). The stacks of disks did not contain the helical assembly typical of TMV rods, indicating that while there might be some attraction between double disks to form these stacks, the predominant assembly is a variant of the literature-reported double disk assembly state, comprising 34 TMV monomers. Dynamic light scattering indicated a 18.2 nm diameter for the RR-TMV disks (Figure S5.4), closely matching the known diameter of TMV assemblies.[20,25] No rods were evident in solution using this analysis method.

There are two distinct ways in which RR-TMV double disks can be formed: one in which both disks face the same direction, and a C_2-symmetric arrangement in which common faces are in contact. Canonical wtTMV assembly is presumed to involve the former arrangement, but the crystal structures of the wtTMV disk assemblies consist of quad disks exhibiting both types of interfaces.[31] Previously, our group reported a circularly permuted monomer (cpTMV) that only formed C_2-symmetric assemblies.[25] All current data, including TEM and assembly behaviors are consistent with a C_2 arrangement for RR-TMV as well. To elucidate some of the potential reasons for this presumed assembly behavior, UCSF Chimera models[32] of the RR-TMV double disk based on the known crystal structure of wild type TMV were constructed. The unique assembly characteristics of the RR-TMV may be due to increased salt-bridges that form between K53R and nearby acidic residues on the adjacent disk (Figure S5.5, S5.6). With K53R, these salt bridges were predicted for both possible double disk stacking interfaces, but were only predicted for the unidirectional interface with native K53. Future work will focus on obtaining a high-resolution crystal structure of the RR-TMV disks to further characterize this assembly phenomenon.

When protein-based nanomaterials are developed for drug delivery and imaging applications, they are commonly conjugated to polyethylene glycol (PEG) polymers to reduce immunogenicity and increase serum stability.[10,17] As such, it was important to determine how effectively the RR-TMV

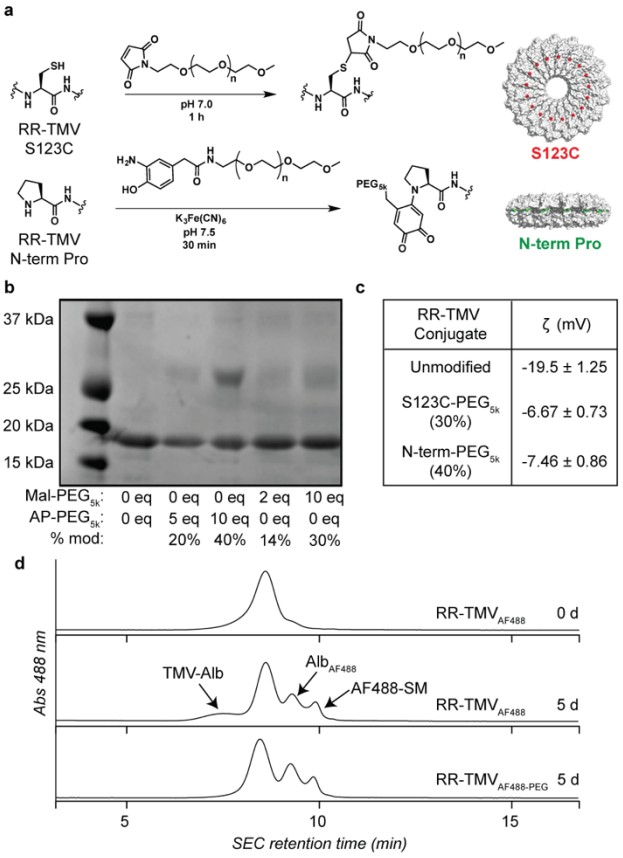


Figure 5.2. PEGylation of RR-TMV disks. (a) RR-TMV disks were expressed to contain a cysteine (S123C) for maleimide modification and a proline-extended N terminus for oxidative coupling with aminophenol-PEG_{5k}. (b) PEG5k was conjugated to either site at varying equivalents to achieve modifications ranging from 14 to 40%. (c) The zeta potentials of the more extensively modified RR-TMVPEG derivatives were measured to determine the effects of charge-shielding by the PEG. (d) The serum stabilities of RR-TMV disks were tested. RR-TMV was modified with AF488 and incubated in 10% FBS solution at 37 °C for 5 d. Stability was monitored at 488 nm absorbance by HPLC-SEC.

disks could be PEGylated for such applications. The RR-TMV disks were expressed to contain two orthogonal handles for bioconjugation: cysteine S123C for maleimide conjugation and an N-terminal proline for an oxidative coupling strategy previously reported by our lab.[33] RR-TMV disks were PEGylated with PEG_{5k} to varying degrees at either the S123C or the N-terminal Pro position (Figure 5.2a,b). The zeta potentials of RR-TMV disks that were PEGylated at 30% for S123C and 40% for the N-terminal Pro positions were measured (Figure 5.2c). Unmodified RR-TMV disks displayed a zeta potential of -19.5 ± 1.25 mV, in close agreement with literature values.[25,31] PEGylation at the S123C position and at the N-terminal Pro shielded the charge of the RR-TMV, with zeta potential values of -6.67 ± 0.73 and -7.46 ± 0.86 mV, respectively. Therefore, both modification sites allow for suitable PEGylation and charge shielding of the RR-TMV disks.

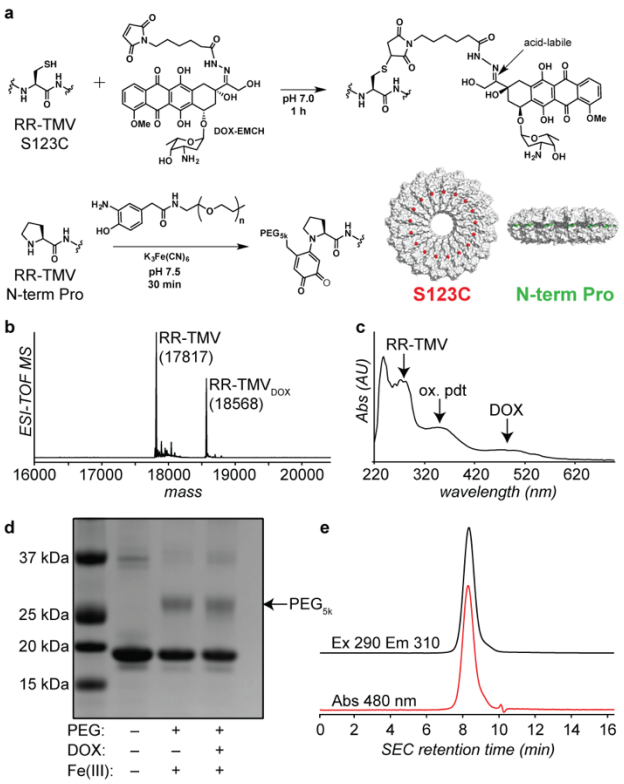


Figure 5.3. Characterization of $RR\text{-}TMV_{DOX\text{-}PEG}$ conjugates. (a) RR-TMV disks were modified at their S123C positions with EMCH-DOX and at their N-terminal prolines with aminophenol-PEG5k. (b) ESI-TOF MS of RR-TMV modified with DOX-EMCH showed ~40% modification using 1 equiv. (c) A UV-vis spectrum of $RR\text{-}TMV_{DOX\text{-}PEG}$ conjugates confirmed 40% modification, as determined by the ratio of absorbance at 280 nm (TMV) to that at 480 nm (DOX). Absorbance at 330 nm corresponds to the oxidative coupling product. (d) RR-TMV PEGylation with aminophenol-PEG_{5k} was confirmed by gel electrophoresis. (e) HPLC-SEC traces of $RR\text{-}TMV_{DOX\text{-}PEG}$ conjugates demonstrate that the disk assembly state is maintained after bioconjugation.

The serum stability of the RR-TMV disks was monitored by HPLC-SEC. RR-TMV disks were labeled with Alexa Fluor 488 maleimide (Figure S5.7) in order to detect at a wavelength that is not effected by background absorbance from the vast excess of serum proteins, and incubated in 10% fetal bovine serum (FBS) solution at 37 °C. The SEC trace at 8.5 min remained the dominant observed peak after 5 d of incubation; however, new peaks at 9.3 and 9.9 min appeared (Figure 5.2d). These peaks are likely due to maleimide-thiol exchange between the AF488-maleimide and thiols present in serum solution. Protein-maleimide conjugates have been observed to undergo thiol exchange on this timescale at 37 °C and to exchange with BSA and thiol-bearing small molecules present in serum.[34] To confirm this, we collected fractions at 8.5, 9.3, and 9.9 min. No proteins were observed in the 9.9 min fraction, and an injection of AF488-maleimide alone eluted at ~10 min (Figure S5.8), indicating that the 9.9 min fraction likely corresponds to thiol exchange between small molecules in serum and the AF488-maleimide (AF488-SM). LC/MS analysis of the 8.5 min peak revealed RR-TMV and $RR\text{-}TMV_{AF488}$ (Figure S5.9). Only BSA and a small peak corresponding to BSA_{AF488} were observed in the 9.3 min peak by LC/MS (Figure S5.9). Additionally, RR-TMV (2 μM) was incubated in 10% FBS to determine if lower concentrations of RR-TMV could be detected by mass

spectrometry in the presence of serum. A small but detectable peak corresponding to the mass of RR-TMV was observed (Figure S5.9d). No such RR-TMV peak was observed in the 9.3 minute SEC fraction, indicating that the disk assembly of the RR-TMV remains intact in serum, while AF488 underwent thiol exchange, as has been previously reported in the literature.[34]

A small peak at 7.5 min appeared in the SEC traces during serum incubation. This is likely BSA bound to the RR-TMV disks. To determine if we could minimize any serum interactions with the RR-TMV, RR-TMV$_{AF488}$ was PEGylated with aminophenol-PEG$_{5k}$ (Figure S5.10) and incubated in 10% FBS solution for 5 d (Figure 5.2d). The emergence of 9.3 and 9.9 min peaks were similarly observed, corresponding to the AF488 thiol exchange products, but no peak at 7.5 min was observed even after 5 d of incubation. Additionally, no disassembly was observed for RR-TMV that was incubated in PBS at 37 °C for 5 d (Figure S5.11). We therefore concluded that the assembly state of RR-TMV remains largely intact in serum solution up to 5 d, and peripheral serum-disk interactions can be minimized by PEGylation.

5.3 Development of RR-TMV disks for drug delivery applications

With these stable RR-TMV disks in hand, we sought to use them for drug delivery applications. Doxorubicin (DOX), a widely used chemotherapy drug for nanoscale drug delivery,[2,4,35,36] was condensed onto the hydrazide-maleimide linker EMCH, and was conjugated via maleimide chemistry to the RR-TMV disks at the S123C position (Figure 5.3a). EMCH allows for acid-sensitive release of the DOX upon endocytosis of the RR-TMV disks by cancer cells.

Incubation with 1 equiv of DOX-EMCH afforded ~40% modification as observed by ESI-TOF MS and UV-vis spectroscopy (Figure 5.3b,c). RR-TMV disks were further modified at their N-terminal proline positions with aminophenol-PEG$_{5k}$ via oxidative coupling. A common oxidizing agent used in bioconjugate techniques, $NaIO_4$, can cause the oxidation of DOX, at either the 1,2-aminoalcohol of the sugar moiety, or at the 1,2-ketoalcohol moiety,[37,38] and can lead to further degradation. A

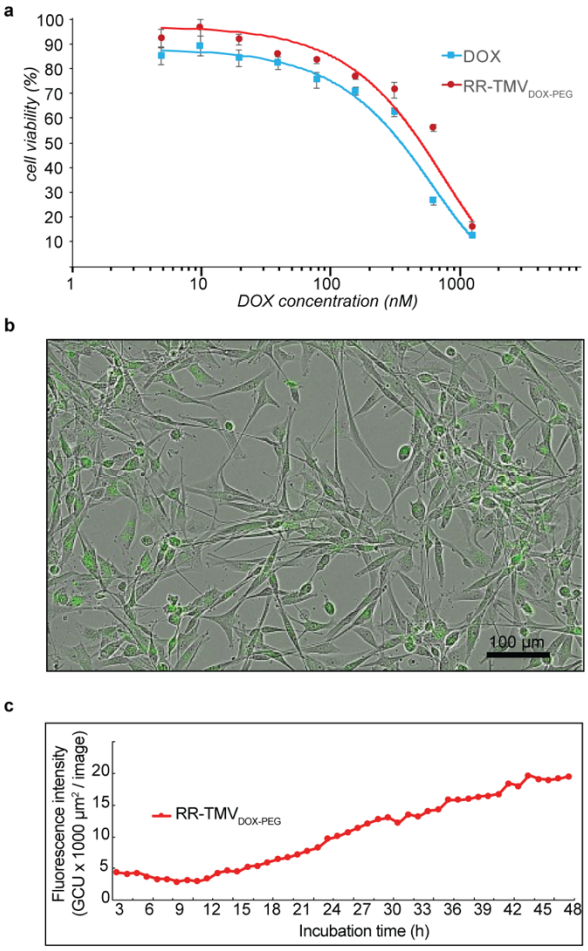


Figure 5.4. Cellular drug delivery profile of RR-TMV$_{DOX-PEG}$ conjugates. (a) U87MG glioblastoma cells (5000 cells/well) were treated with varying concentrations of RR-TMV$_{DOX-PEG}$ conjugates. Significant cell death was observed after 72 h of incubation. Cells incubated with RR-TMV alone displayed 100.9 ± 0.2% viability at TMV concentrations equal to the highest concentration of RR-TMV$_{DOX-PEG}$ samples. (b) U87MG cell uptake of RR-TMV$_{DOX-PEG}$ conjugates at 48 h. (c) Uptake kinetics were monitored in the green channel to detect DOX fluorescence.

milder oxidant, $K_3Fe(CN)_6$, was used to avoid any decomposition and loss of functionality of DOX. In feasibility experiments, treatment of RR-TMV_{DOX} with $NaIO_4$ led to a mass loss of 129 Da, corresponding to the cleavage of the sugar moiety (Figure S5.12c). However, no such changes were observed upon treatment with $K_3Fe(CN)_6$ (Figure S5.12b). Conjugation of RR-TMV_{DOX} with aminophenol-PEG_{5k} in the presence of $K_3Fe(CN)_6$ resulted in 30% PEGylation, as observed by gel electrophoresis (Figure 5.3d). SEC traces of RR-$TMV_{DOX\text{-}PEG}$ conjugates revealed that the disk assembly state was maintained after conjugation (Figure 5.3e).

The drug delivery potential of the RR-$TMV_{DOX\text{-}PEG}$ conjugates was next evaluated *in vitro*. U87MG glioblastoma cells (5000 cells per well) were incubated with RR-$TMV_{DOX\text{-}PEG}$ conjugates at varying concentrations, and cell viability was monitored via an MTS assay. After 72 h of incubation, significant cell death was observed for both DOX-treated and RR-$TMV_{DOX\text{-}PEG}$-treated cells when compared to cell controls (Figure 5.4a). Cells incubated with RR-TMV alone displayed $100.9 \pm 0.2\%$ viability at TMV concentrations equal to the highest concentration of RR-$TMV_{DOX\text{-}PEG}$ samples. The difference in cell viability curves of DOX alone and RR-$TMV_{DOX\text{-}PEG}$ is most likely due to the slow release profile of the hydrazone linkage between the DOX and EMCH.[4,35]

The cellular uptake of RR-$TMV_{DOX\text{-}PEG}$ was monitored by fluorescence microscopy, taking advantage of the fluorescent properties of DOX. Cellular internalization was observed after 48 h of incubation, with fluorescence observed throughout the cytoplasm and in the nuclei, where the DOX can take effect (Figure 5.4b). Uptake kinetics were monitored via time lapsed live cell fluorescent microscopy. Significant uptake was observed starting after 12 h of incubation, with a linear increase in uptake observed up to 48 h (Figure 5.4c).

5.4 Evaluation of RR-TMV disks for glioblastoma accumulation

To determine the effects of shape and size on tumor accumulation and retention, a panel of protein-based nanomaterials were evaluated for their drug delivery potentials. This panel consisted of the MS2 sphere, the RR-TMV disks, as well as a nanoscale filamentous phage[39] (Figure 5.5a). While the MS2 had previously been evaluated for *in vivo* tumor accumulation in breast models,[8,40] neither the RR-TMV disks nor the nanophage were previously tested *in vivo*. Glioblastoma models were chosen for these studies, as glioblastoma is a particularly deadly form of cancer, where one of the primary barriers to treatment is in the penetration and accumulation of drugs to the tumor tissue.[1] It is therefore important to determine the effects that nanomaterial morphology has on the penetration into and subsequent retention of the nanocarriers in the glioma microenvironment.

Two approaches were taken in evaluating these nanomaterials for the drug delivery potential. We first sought to determine if the morphology of the nanocarriers affected their passive accumulation into the tumor environment. All three protein-based nanomaterials were modified with maleimide-NOTA-^{64}Cu for detection via positron emission tomography (PET), as well as with PEG_{5k} polymers to enhance their biodistribution. The radiochemical yield for MS2 was 87% with a specific activity (radioactivity per mol) of 479 mCi/nmol or 86 mCi/nmol capsid, while the radiochemical yield of TMV was 65% with a specific activity of 12.1 mCi/nmol capsid, and the radiochemical yield of the nanophage was 13% with a specific activity of 2.5 mCI/nmol capsid. These discrepancies between nanocarriers was accounted for in the PET analyses, and each nanocarrier presented sufficient radioactivity for sensitive PET detection.

Orthotopic glioblastoma mouse models were established using U87MG-Luciferase cell lines, allowing for tumor size monitoring via luciferase activity. When tumors reached a significant size

at 26 days post-implantation, the nanocarriers were injected intravenously at a dose of 50-200 mCi per injection. At 5 h post-treatment, a scan was performed and images were recorded and analyzed after decay correction and normalization to the injected dose (Figure 5.5b). After 5 h, the mice were sacrificed, and the radioactivity of the organs was measured to determine the percent injected dose per gram (%IG/g). For all three of the nanocarriers, the majority of the agents accumulated into clearance organs such as the liver and spleen. Interestingly, a significant amount of both the MS2 spheres and the RR-TMV disks remained in the blood pool, suggesting that longer time point analyses may be desirable to increase tumor accumulation. Approximately 2-5% ID/g accumulation in the tumor was observed for each carrier, with the nanophage displaying higher tumor accumulation than the MS2 and RR-TMV agents (Figure 5.5c). This was also observed in the PET image taken at 5 h, where the nanophage was clearly visible in the tumor tissue, while the RR-TMV and MS2 were not as visibly present. Importantly, there was negligible accumulation of all three nanocarriers into the non-cancerous brain tissue, indicating that the tumor-associated disruption of the blood-brain barrier led to accumulation of the protein-based nanomaterials into the tumor environment. The overall levels of protein-based nanomaterial accumulation into the tumor environment is in agreement with literature values of other synthetic materials such as gold nanoparticles.[41] While the enhanced accumulation of the nanophage into the tumor tissue was not statistically significant when compared to the RR-TMV or MS2 accumulation, future studies with greater replicates and more consistent protein modification across capsids may yield statistically significant comparisons between nanomaterials.

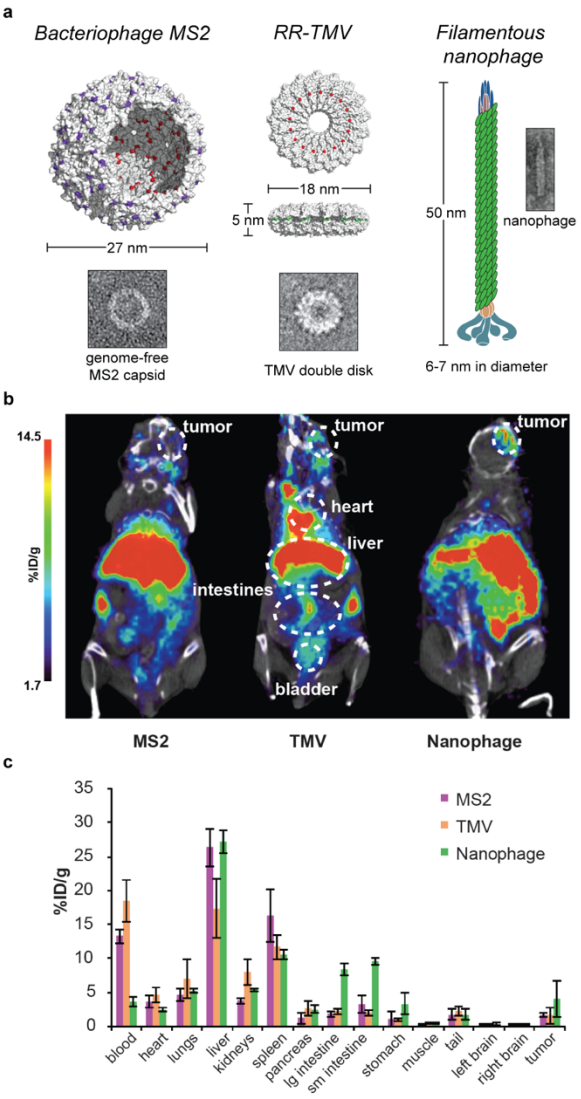


Figure 5.5. Evaluation of protein-based nanomaterials for glioblastoma accumulation. (a) Three different nanocarriers were tested to determine morphological effects on tumor penetration. (b) PET-CT image of nanocarrier accumulation at 5 h post-injection. Nanophage and RR-TMV were detected in the tumor region, with nanophage accumulation observably stronger, while MS2 was not visible in the tumor. (c) Biodistributions of the tested nanocarriers as determined by percent injected dose per gram (%ID/g). All three nanocarriers accumulated at 2-5% ID/g with no statistical significance observed between the three.

5.5 Evaluation of RR-TMV disks for glioblastoma treatment via CED injection

While intravenous injection is the standard method of chemotherapy delivery in the clinic, other injection methods may have greater chemotherapeutic potency. One such method is convection enhanced delivery (CED), whereby chemotherapeutic agents are administered directly to the targeted brain tissue with a microinfusion pump.[42] A constant pressure is maintained during the injection that creates a fluid convection to facilitate a homogenous diffusion of the chemotherapeutic throughout the targeted area. After CED infusion, it is important for the chemotherapeutic agent to remain in the tumor tissue. We hypothesized that the delivery of chemotherapeutic agents within protein-based nanomaterials via CED may help facilitate tumor retention of the payload owing to the EPR effect.

Each of the three protein-based nanocarriers were modified with doxorubicin and polyethylene glycol using similar methods as described in section 5.3. While the payload of DOX per nanocarrier was variable (Table S5.1), each glioma-bearing mouse was injected with a standardized DOX payload of 20 µg/kg, and tumor size was monitored via luminescence. While the tumor growth of the RR-TMV-treated mice does suggest some tumor inhibition, (Figure 5.6a) the large variability in tumor growth within groups led to no measured statistical difference between RR-TMV-treated and PBS control groups at day 27, the last day with enough surviving mice in the PBS control to provide statistical analysis. The Kaplan-Meier survival curve also suggests improved tumor inhibition from the RR-TMV treatment, as three out of eight mice survived past 40 days, whereas for the PBS-treated group, no mouse survived past 34 days (Figure 5.6b). While these results are particularly promising, upon log-rank analysis of the two survival curves, a P value of 0.052 was determined, preventing us from claiming statistical significance between the two groups. Future work will focus on reducing the high standard deviation present within cohorts to better achieve statistical significance between treatments.

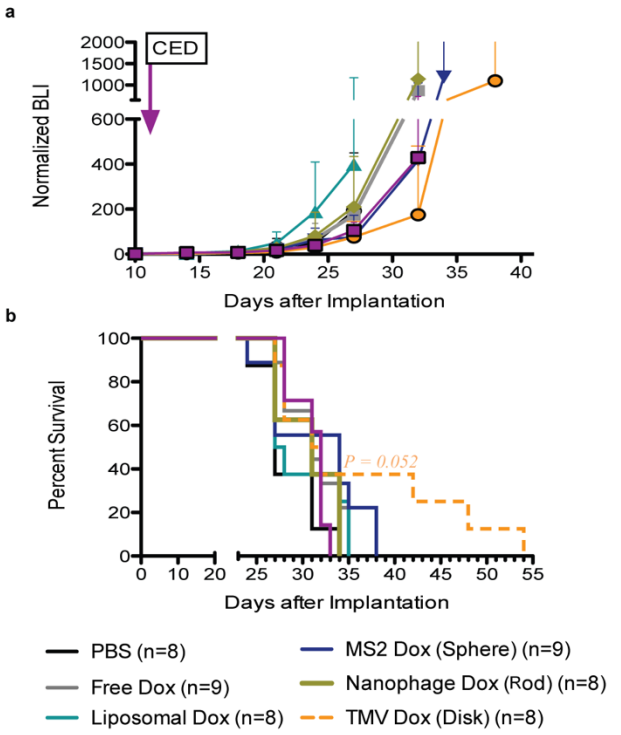
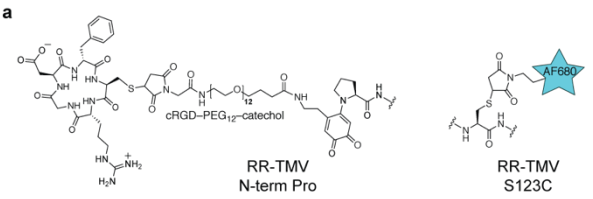


Figure 5.6. Evaluation of protein-based nanomaterials for glioblastoma treatment. MS2 spheres, TMV disks, and nanophage rods bearing DOX payloads of 20 µg/kg were injected via CED infusion into U87MG glioma-bearing mice. (a) Tumor growth analysis suggested improved tumor inhibition from TMV-treated groups, although no statistical significance was determined between groups. (b) Kaplan-Meier survival curve of mice with U87MG glioma. While the TMV-treated mice showed enhanced survival over all other groups, no statistical significance (P = 0.052) between PBS-treated and TMV-treated mice was observed.

5.6 Adapting the RR-TMV disks for active tumor targeting via integrin receptors

To enhance the accumulation and retention of the protein-based nanomaterials into the tumor environment, an active targeting approach was taken. While passive targeting does provide some tumor accumulation via the EPR effect, attaching receptor-targeting motifs onto nanocarriers has shown promise in increasing tumor accumulation[1] and in increasing tumor retention by binding to receptors present on the cancer cells. Cyclo-RGD (cRGD) peptides, which specifically bind to integrin receptors present on cancer cells,[43] were attached to the surface of the RR-TMV disks, as these nanocarriers showed the most promise in the CED trial.



b

Agent:	No Treatment	cRGD-AF680	TMV-AF680	TMV-AF680-cRGD
Mean Uptake:	*79*	*2155*	*3561*	*5577*

Figure 5.7. Development of protein-based nanomaterials for active tumor targeting. (a) cRGD targeting groups were attached to the N terminus of RR-TMV disks and Alexa Fluor 680 dyes were attached to reactive cysteines at position 123. (b) RR-TMV$_{PEG\text{-}cRGD}$ nanocarriers demonstrated 1.5x improved uptake in U87MG glioblastoma cells as observed by flow cytometry when compared to RR-TMV without cRGD groups.

Catechol-PEG$_{12}$-cRGD was synthesized and attached to the RR-TMV capsids via oxidative coupling to the proline N terminus of RR-TMV (Figure 5.7a). Alexa Fluor 680 dyes were attached at reactive cysteines in order to monitor active targeting to cancer cells via flow cytometry. Flow cytometry revealed a 1.5x enhancement in binding for the RR-TMV$_{cRGD}$ capsids when compared to unmodified RR-TMV disks. These findings indicate that the RR-TMV disk is a promising nanocarrier for receptor-targeted drug delivery.

5.7 Supramolecular-induced morphology changes of TMV nanocarriers

We have demonstrated that there are different distribution behaviors for VLPs of varying morphologies. It would therefore be beneficial to develop a VLP that can alter its morphology upon external stimulus. (Figure 5.8). This could allow for responsive biodistribution and tumor uptake behaviors, as well as alternating the display patterns of receptor-targeting moieties conjugated onto the VLP surface. Several examples of shape-adaptable nanomaterials have been reported, although primarily with polymeric scaffolds that undergo swilling or shrinking in response to a stimulus.[44] We sought to create a shape-adaptable nanomaterial that is composed of VLPs and can undergo large morphological rearrangements in response to an external stimulus. To address this, cucurbit[8]uril (CB8) host-guest chemistry was adapted to trigger TMV assembly into higher ordered structures (Figure 5.8). CB8 host-guest chemistry has been previously used to create protein assemblies,[45] although the type of higher-ordered structure demonstrated herein had yet to be achieved. We chose to use methylviologen and azobenzene guests for induction of supramolecular TMV assemblies. Methylviologen and azobenzene derivatives are known to

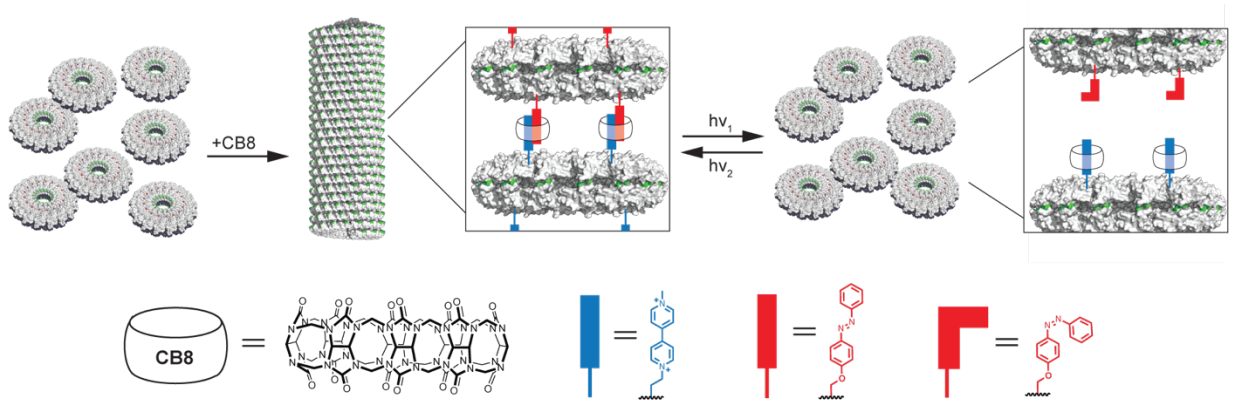


Figure 5.8. Cucurbit[8]uril-triggered TMV assemblies. TMV disks were modified with either methylviologen (blue) or azobenzene (red) maleimide derivatives. These derivatives can undergo a 1:1:1 complex within the cucurbit[8]uril (CB8) cavity. This complex would then trigger the assembly of a TMV rod from the conjugated disks. Upon stimulation with UV light, the azobenzene trans-cis isomerization would force the ejection of the azobenzene-labeled TMV from the CB8 cavity and therefore trigger the disassembly of the TMV rods back into disks. Subsequent stimulation with a blue light (>400 nm) can switch the isomerization and lead to a reversible morphological shift.

assemble into heteroternary complexes in the CB8 cavity.[46,47] Interestingly, while methylviologen (MV) is able to bind to CB8 in a 1:1 complex, azobenzene is unable to bind to CB8 until methylviologen is already bound, at which point the 1:1:1 complex between CB8, azobenzene, and methylviologen is formed, as the 1:1:1 complex is thermodynamically favored over the 1:1 MV:CB8 complex.[46] Upon stimulation by UV-light, the *trans* to *cis* conformational change of azobenzene stimulates the expulsion of the azobenzene derivative from the CB8 cavity, causing the disassembly of the heteroternary complex.[47] For this study, a previously reported circular permutant TMV (cpTMV) disk was chosen as it has been shown to have a C_2 axis of symmetry[25] that is necessary for higher order CB8-triggered assembly. Methylviologen-maleimide and azobenzene (Az) derivatives were synthesized and conjugated onto two separate populations of TMV (Figure S5.14). MV-TMV conjugates were incubated with CB8, followed by Az-TMV in order to trigger the assembly. After 2 d, higher order assemblies were observed by dynamic light scattering (DLS). A diameter increase from 18.3 nm to 33.2 nm was observed by DLS (Figure 5.9). As the cpTMV undergoes a morphological change from disk to rod, the DLS shifts in measuring the diameter of the flat disk, to the length of the newly assembled rod. Therefore, this change in diameter would

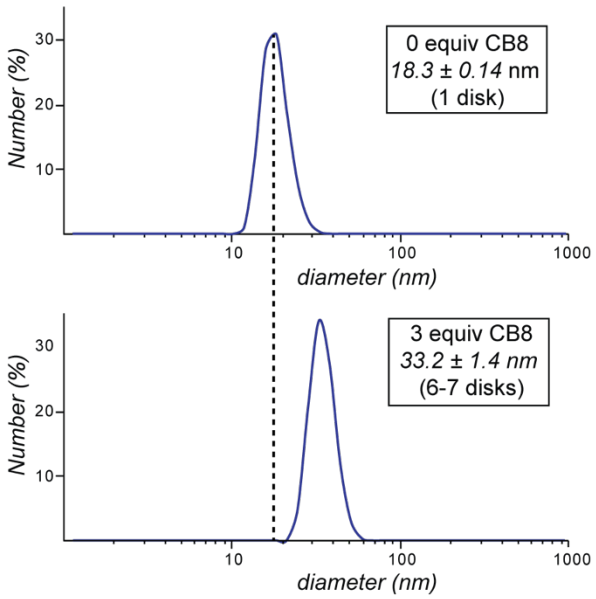


Figure 5.9. Dynamic light scattering (DLS) of CB8-triggered TMV assemblies. Incubation of MV-TMV and Az-TMV conjugates with 3 equiv of CB8 led to an increase in diameter from 18 to 33 nm, as observed by DLS. This increase in diameter corresponds to a newly formed TMV assembly consisting of 6-7 stacked disks.

correspond to an increase from a single cpTMV disk of 5 nm in height and 18 nm in diameter, to 6-7 disks assembled via CB8 host-guest chemistry. Further studies confirming the higher order assemblies via electron microscopy and X-ray scattering will be conducted. Additionally, cell uptake studies of CB8-assembled TMV rods and TMV disks will be performed to determine if the morphology of the disks will affect their cellular uptake behaviors.

5.8 Conclusions

We present here a stable protein nanomaterial with a novel morphology for drug delivery applications. The characterization of the RR-TMV disks revealed that these proteins self-assemble into stable nanoscale disks that retained their morphology under various physiologically relevant conditions. RR-TMV disks were modified with both doxorubicin and PEG_{5k} via orthogonal bioconjugation strategies, and were successfully used to deliver doxorubicin both *in vitro* and *in vivo*. *In vivo* studies indicate that RR-TMV disks may have improved tumor retention and therefore improved CED chemotherapeutic potential over other shaped protein-based nanomaterials. RR-TMV disks were further modified with integrin-targeting cRGD moieties and demonstrated enhanced binding to U87MG glioblastoma cells. RR-TMV disks therefore provide a promising new nanocarrier that will expand the shape library of protein-based nanomaterials for targeted drug delivery and imaging. Lastly, preliminary findings indicate that supramolecular interactions could lead to higher order assemblies from the TMV disks, which could lead to triggered morphological changes for drug delivery applications. Future work will focus on further characterizing the unique assembly state of the RR-TMV disks, as well as exploring the tumor accumulation and drug delivery via active integrin targeting to cancer cells *in vivo*.

5.9 Materials and methods

All solvents and reagents, including doxorucibin-HCl, were purchased from commercial suppliers and used without further purification. DOX-EMCH was prepared according to literature procedures.[36] Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} (E. Merck) and visualized under a UV lamp at 254 nm. A C18 column was used for analytical and semi-preparative reverse phase high performance liquid chromatography (RP-HPLC) on an Agilent 1100 Series Capillary LC. Runs were eluted with H_2O/MeCN (0.1 % v/v TFA) and monitored using a UV-Vis detector. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker Avance 400 with working frequencies of 400 MHz for ^{1}H NMR, and 100 MHz for ^{13}C NMR, respectively. UV-Vis spectra were obtained using a NanoDrop (Thermo Scientific).

Protein expression and purification. The starting point for the RR-TMV protein was a gene for the coat protein of the TMV U1 strain optimized for the codon usage of *E. coli* (Genscript, Piscataway, NJ). Site-directed mutagenesis was performed using QuikChange mutagenesis (Stratagene, Santa Clara, CA). For N-terminal oxidative coupling, the N terminus of the RR-TMV was extended from SYS to PAGSYS using the primers:

sense: 5’ – GAAGGAGATATACATATGCCTGCCGGCAGCTATAGCATTACC – 3’
antisense: 5’ – TGCTATAGCTGCCGGCAGGCATATGTATATCTCCTTCTTAAG – 3’

The RR-TMV coat protein was expressed to contain K53R, K68R, T104K, S123C, and PAGSYS N-terminal mutations. BL21 DE3 RIL Codon+ cells were transformed with the vector described above, and cultured in Terrific Broth with 100 μg/L ampicillin at 37 °C. When cultures reached optical densities of 0.6 to 0.8, IPTG was added to a final concentration of 30 μM. Cultures were grown 24 h at 30 °C, harvested by centrifugation, and stored at -80 °C. Cells (from a 1 L expression batch) were thawed, resuspended in 20 mL of 20 mM TEA pH 8, and lysed by sonicating with a 2 s on, 4 s off cycle for a total of 30 min using a standard disruptor horn at 90% amplitude (Fischer Scientific). The resulting lysate was cleared by ultracentrifugation for 30 min at 45,000 rpm using a Beckman 45 Ti rotor in an Optima L-80 XP (Beckman Coulter). The clarified lysate was decanted, warmed to room temperature, and stirred while adding a saturated solution of ammonium sulfate dropwise to a final concentration of 50% (v/v). After 5 min, the white ppt that formed was pelleted by ultracentrifugation (30 min at 45,000 rpm in a Beckman 45 Ti rotor), and resuspended in 20 mM TEA pH 7.2. The resulting protein solution was next loaded onto a diethylaminoethanol (DEAE) Sepharose column and eluted with a 0 – 300 mM NaCl gradient. Purity was confirmed by SDS-PAGE and HPLC. This preparation provided pure RR-TMV in yields up to 100 mg/L culture.

Size Exclusion Chromatography (SEC). Analytical size exclusion was performed on an Agilent 1100 series HPLC equiped with a PolySep-GFC-P 5000 column (Phenomenex, Torrance, CA), at a flow rate of 1 mL/min.

Transmission Electron Microscopy (TEM). Samples were prepared for TEM analysis by applying an analyte solution (approximately 0.2 mg/mL in RR-TMV protein) to carbon-coated copper grids for 2 min, followed by triple rinsing with dd-H_2O. The grids were then exposed to a 1.6% aqueous solution of uranyl acetate for 2 min as a negative stain. Images were obtained at the Berkeley Electron Microscope Lab using a FEI Tecnai 12 transmission electron microscope with 120 kV accelerating voltage.

Dynamic light scattering (DLS) and zeta potential studies. DLS and zeta potential measurements were obtained using a Malvern Instruments Zetasizer Nano ZS. Data plots and standard deviations were calculated from an average of three measurements, each of which consisted of 10 runs of 45 s each. Measurement data are presented as a number distribution.

Mass spectrometry. Protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an Agilent 6224 Time-of Flight (TOF) LC/MS system equipped with a Turbospray ion source.

Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a Mini-Protean apparatus (Bio-Rad, Hercules, CA), using a 10-20% precast linear gradient polyacrylamide gel (Bio-Rad). All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 75-90 min at 120 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). ImageJ was used to determine the level of modification by optical densitometry.

Synthesis of cRGD-PEG$_{12}$-catechol. To a solution of commercially available maleimide-PEG$_{12}$-NHS (9.9 mg, 0.011 mmol) in 1 mL DMF was added 6.3 equiv of dopamine (10.7 mg, 0.07 mmol) and 6 equiv of DIPEA. The reaction was stirred for several hours and monitored via LC/MS. To the crude mixture was added cRGDFC (1 mg, 0.0018 mmol) in 1 mL of 100 mM KPhos pH 7.5. The reaction was stirred at room temperature overnight. The reaction mixture was then lyophilized and resuspended in 1:1 ACN:H_2O with 0.1% TFA and purified via RP-HPLC to yield cRGD-PEG$_{12}$-catechol. HRMS (LC/MS): *m/z* calc'd for $C_{66}H_{103}O_{25}N_{11}S$. $[M + H]^+$ 1482.66, observed 1482.24.

General procedure for Alexa Fluor 488 conjugation to RR-TMV. To a solution of RR-TMV (PAGSYS N-terminus, S123C, T104K, 100 μM) in 10 mM pH 7 NaPhos buffer, 1.5 equiv of AF488 (22.6 μL of 1 mM stock solution) was added. The solution was incubated for 1 h, protected from light at room temperature. The solution was then spin concentrated 3-5 times into 10 mM pH 7 NaPhos buffer.

General procedure for aminophenol-PEG$_{5k}$ conjugation to RR-TMV. To a solution of RR-TMV (PAGSYS N-terminus, S123C, T104K, 100 μM) in 20 mM pH 7.5 NaPhos buffer, 2 equiv of *N*-ethyl maleimide (0.8 μL of a 100 mM stock solution) was added. The reaction was incubated for 1 h at room temperature. An aliquot of the mixture was analyzed by ESI-TOF MS to confirm the formation of the RR-TMV$_{NEM}$ complex. After 1 h of reaction time, 5 equiv of aminophenol-PEG$_{5k}$ (133 μL of 3 mM stock solution) was added. Following this, 50 equiv of $K_3Fe(CN)_6$ (80 μL of a 3 mM stock solution) was added. After 30 min of incubation at room temperature, the solution was spin concentrated 3-5 times into 10 mM pH 7 NaPhos buffer with a 30k molecular weight cutoff (MWCO) spin concentrator (Millipore).

General procedure for maleimide-PEG$_{5k}$ conjugation to RR-TMV. To a solution of RR-TMV (PAGSYS N-terminus, S123C, T104K, 100μM) in 20 mM NaPhos buffer pH 7.5, 10 equiv of maleimide-PEG$_{5k}$ (8 μL of 50 mM stock solution) was added. After 2 h of incubation at room temperature, the solution was spin concentrated 3-5 times into 10 mM NaPhos buffer pH 7 with a 30k MWCO spin concentrator.

General procedure for doxorubicin and aminophenol-PEG$_{5k}$ conjugation to RR-TMV. To a solution of RR-TMV (PAGSYS N-terminus, S123C, T104K, 100 μM) in 20 mM pH 7.5 NaPhos buffer, 1 equiv of DOX-EMCH (2 μL of a 100 mM stock solution in DMSO) was added. The solution was incubated for 1 h, protected from light at room temperature. An aliquot of the mixture was analyzed by ESI-TOF MS to confirm the formation of the RR-TMV$_{DOX}$ complex. Then, 5 equiv of PEG$_{5k}$-aminophenol (400 μL of a 2.5 mM stock) was added to the solution. Following this, 50 equiv of $K_3Fe(CN)_6$, (200 μL of a 50 mM stock solution) was added. After 30 min of incubation at room temperature, the solution was spin concentrated 3-5 times into PBS with a 30k MWCO spin concentrator.

Cell culture. U87MG human glioblastoma cells were obtained from the UC Berkeley Cell Culture Facility. Cells were cultured in DMEM containing phenol red (ATCC, Manassas, VA) or DMEM without phenol red (Thermo, Waltham, MA) with 10% fetal bovine serum (Omega Scientific, Tarzana, CA) and 1% penicillin/streptomycin (Thermo) at 37 °C and 5% CO_2.

MTS cell viability assay. A 3-(4,5-Dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium (MTS) cell viability assay was purchased from Promega and used according to the manufacturer's instructions. U87MG cells were cultured in high glucose DMEM containing 10% fetal bovine serum and penicillin and streptomycin. Cells were plated in 96-well microtiter plates at 5000 cells per well and incubated overnight to allow adhesion. After overnight incubation, the media was removed and exchanged with either DOX, RR-TMV, or RR-$TMV_{DOX-PEG}$ stock solutions prepared in cell media at the appropriate concentration. Cell viability was measured after 72 h using an MTS cell viability assay and was used according to the supplier's instructions. The cell media was replaced after 72 h with a stock of 20% MTS solution in cell media. The plate was incubated for 1−3 h, and the absorbance was read using a microplate reader (490 nm). Cell viability was calculated as an absorbance percent relative to the untreated cell control. Cells were treated in triplicate.

Cell uptake studies. U87MG cells were seeded in a 96 well plate as 200 µL solutions of 50,000 cells/mL. The cells were allowed to grow and adhere to the plate for 48 h in DMEM media with 10% FBS and 1% P/S at 37 °C and 5% CO_2. The media was removed and replaced with 200 µL DMEM without phenol red containing RR-$TMV_{DOX-PEG}$ conjugates at a final concentration of 1 µM DOX (3 µM RR-TMV coat protein) or 2 µM unmodified RR-TMV coat protein. The cells were incubated with the agents for an additional 48 h. Incucyte Zoom Live-Cell Analysis System (EssenBio, Ann Arbor, MI) was used to collect images every hour post incubation. Phase and green fluorescence images at 20x magnification were collected, capitalizing on the intrinsic fluorescence of doxorubicin. The images were processed using Incucyte Zoom proprietary software v.2016A, and the Top-Hat background subtraction algorithm (radius 10 µm, threshold 0.5% of green calibration dye signal, GCU) was used to define the boundaries of the cells (green objects). The mean green object fluorescence intensities of 4 images taken in each well were plotted against the incubation time.

Flow cytometry studies. U87MG cells were washed with DPBS, removed with trypsin-free cell dissociation buffer, and then transferred into a 15 mL Falcon tube. Cells were centrifuged at 180 rcf for 3 min and washed once into 1 mL DPBS. Cells were counted using a hemocytometer and aliquoted to $3x10^5$ cells per Eppendorf tube. Cells were then washed twice with cold FACS buffer. After a third resuspension, the cells were pelleted (180 rcf for 3 min), and then resuspended in 100 µL of either cold FACS buffer, a solution of 1 µM protein-based nanomaterial (based on monomer) in cold FACS buffer. Cell solutions were mixed and incubated for 45 min on ice. Following incubation on ice, the cells were pelleted (180 rcf for 3 min), and washed twice with cold FACS buffer. After the last round of centrifugation, cells were resuspended into 200 µL of cold FACS buffer and transferred to a 5 mL Falcon tube containing a 2 mL insert. The cells were then observed by flow cytometry (FACSCalibur, BD Biosciences) by gating the live cells and quantifying their fluorescence in the FL1 channel with a histogram plot to determine degree of fluorescence. The degree of fluorescence could then be correlated to the binding efficiencies of the protein samples.

PET/CT biodistribution studies. All procedures were performed as previously reported,[39,40] and according to a protocol approved by the UCSF Institutional Animal Care and Use Committee (IACUC). Five-week old athymic (nude) female mice weighing 18-23 g were purchased from Simonsen Labs (Gilroy, CA). For tumor inoculation, $3x10^5$ U87MG-Luc glioblastoma cells with luciferase reporter gene were implanted intracranially. The imaging and biodistribution

experiments were started 26 days following implantation. Tumor-bearing mice in sets of 3 animals per study group were injected in the tail vein (intravenous, IV) with 50-200 μCi of 64Cu-labeled conjugates in 100 μL of sterile saline. One animal from each group was selected for imaging with microPET/CT (Inveon microPET docked with microCT, Siemens, Washington, D.C.). Dynamic imaging was performed from the time of injection to 1 h post injection, followed by a 20 min static scan at 5 h. CT scans were performed after each PET scan to provide anatomical localization of radionuclide data, as well as photon attenuation map for attenuation-corrected PET reconstruction. Images were reconstructed using the AMIDE software v.1.0.4. More information 95 regarding the circulation profile of the agents based on the PET data is presented in the Appendix. Mice were sacrificed at 5 h post-injection and the blood, tumor, and major organs were harvested and weighed. The radioactivity present in each sample was measured using a Wizard gamma-counter (Perkin Elmer, Waltham, MA). All values were decay corrected, and the percentage injected dose per gram (% ID/g) was calculated for each tissue sample. Means and standard deviations were calculated for each group. Using Excel software (Microsoft, Redmond, WA), an unpaired t-test with equal variance and a two-tailed p value was performed for organs from different data sets. A result was considered statistically significant if it occurred at the $p < 0.05$ level.

Tumor shrinkage studies in glioblastoma models. All procedures were performed as previously reported.[39,40] Mice bearing U87MG-Luc intracranial tumors in sets of 9 animals per study group were injected via CED with 10 μL of protein-Dox conjugates in sterile saline, at day 10 post-implantation. Tumor size was monitored using the Luciferase reporter system and measuring bioluminescence on an IVIS 50 Lumina system.

5.10 References

(1) Peer, D.; Karp, J. M.; Hong, S.; Farokhzad, O. C.; Margalit, R.; Langer, R. Nanocarriers as an Emerging Platform for Cancer Therapy. *Nat. Nanotechnol.* **2007**, *2* (12), 751–760.

(2) Ulbrich, K.; Etrych, T.; Chytil, P.; Pechar, M.; Jelinkova, M.; Rihova, B. Polymeric Anticancer Drugs with pH-Controlled Activation. In *International Journal of Pharmaceutics*; 2004; Vol. 277, pp 63–72.

(3) Fox, M. E.; Szoka, F. C.; Fréchet, J. M. J. Soluble Polymer Carriers for the Treatment of Cancer: The Importance of Molecular Architecture. *Acc. Chem. Res.* **2009**, *42* (8), 1141–1151.

(4) Toita, R.; Murata, M.; Abe, K.; Narahara, S.; Piao, J. S.; Kang, J.-H.; Hashizume, M. A Nanocarrier Based on a Genetically Engineered Protein Cage to Deliver Doxorubicin to Human Hepatocellular Carcinoma Cells. *Chem. Commun.* **2013**, *49* (67), 7442–7444.

(5) Dedeo, M. T.; Finley, D. T.; Francis, M. B. Viral Capsids as Self-Assembling Templates for New Materials. *Prog. Mol. Biol. Transl. Sci.* **2011**, *103*, 353–392.

(6) Obermeyer, A. C.; Capehart, S. L.; Jarman, J. B.; Francis, M. B. Multivalent Viral Capsids with Internal Cargo for Fibrin Imaging. *PLoS One* **2014**, *9* (6), e100678.

(7) Rhee, J.-K.; Baksh, M.; Nycholat, C.; Paulson, J. C.; Kitagishi, H.; Finn, M. G. Glycan-Targeted Virus-like Nanoparticles for Photodynamic Therapy. *Biomacromolecules* **2012**, *13* (8), 2333–2338.

(8) Farkas, M. E.; Aanei, I. L.; Behrens, C. R.; Tong, G. J.; Murphy, S. T.; O'Neil, J. P.; Francis, M. B. PET Imaging and Biodistribution of Chemically Modified Bacteriophage MS2. *Mol. Pharm.* **2013**, *10* (1), 69–76.

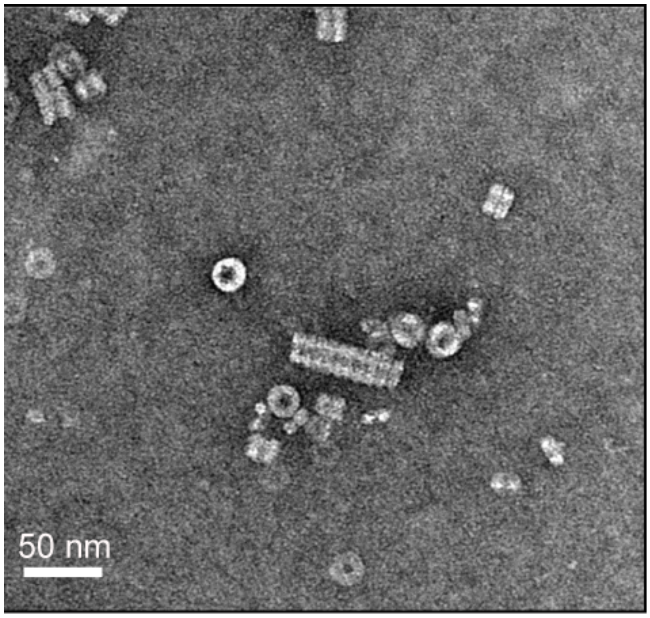
(9) Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery. *Angew. Chemie - Int. Ed.* **2009**, *48* (50), 9493–9497.

(10) Wen, A. M.; Steinmetz, N. F. Design of Virus-Based Nanomaterials for Medicine, Biotechnology,

and Energy. *Chem. Soc. Rev.* **2016**.
(11) Singh, P.; Prasuhn, D.; Yeh, R. M.; Destito, G.; Rae, C. S.; Osborn, K.; Finn, M. G.; Manchester, M. Bio-Distribution, Toxicity and Pathology of Cowpea Mosaic Virus Nanoparticles in Vivo. *J. Control. Release* **2007**, *120* (1), 41–50.
(12) Shukla, S.; Ablack, A. L.; Wen, A. M.; Lee, K. L.; Lewis, J. D.; Steinmetz, N. F. Increased Tumor Homing and Tissue Penetration of the Filamentous Plant Viral Nanoparticle Potato Virus X. *Mol. Pharm.* **2013**, *10* (1), 33–42.
(13) Qazi, S.; Liepold, L. O.; Abedin, M. J.; Johnson, B.; Prevelige, P.; Frank, J. A.; Douglas, T. P22 Viral Capsids as Nanocomposite High-Relaxivity MRI Contrast Agents. *Mol. Pharm.* **2013**, *10* (1), 11–17.
(14) Geng, Y.; Dalhaimer, P.; Cai, S.; Tsai, R.; Tewari, M.; Minko, T.; Discher, D. E. Shape Effects of Filaments versus Spherical Particles in Flow and Drug Delivery. *Nat. Nanotechnol.* **2007**, *2* (4), 249–255.
(15) Decuzzi, P.; Pasqualini, R.; Arap, W.; Ferrari, M. Intravascular Delivery of Particulate Systems: Does Geometry Really Matter? *Pharmaceutical Research*. 2009, pp 235–243.
(16) Gratton, S. E. a; Ropp, P. a; Pohlhaus, P. D.; Luft, J. C.; Madden, V. J.; Napier, M. E.; DeSimone, J. M. The Effect of Particle Design on Cellular Internalization Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (33), 11613–11618.
(17) Shukla, S.; Eber, F. J.; Nagarajan, A. S.; DiFranco, N. A.; Schmidt, N.; Wen, A. M.; Eiben, S.; Twyman, R. M.; Wege, C.; Steinmetz, N. F. The Impact of Aspect Ratio on the Biodistribution and Tumor Homing of Rigid Soft-Matter Nanorods. *Adv. Healthc. Mater.* **2015**, *4* (6), 874–882.
(18) Moyer, T. J.; Finbloom, J. A.; Chen, F.; Toft, D. J.; Cryns, V. L.; Stupp, S. I. pH and Amphiphilic Structure Direct Supramolecular Behavior in Biofunctional Assemblies. *J. Am. Chem. Soc.* **2014**, *136* (42), 14746–14752.
(19) Czapar, A. E.; Zheng, Y.-R.; Riddell, I. A.; Shukla, S.; Awuah, S. G.; Lippard, S. J.; Steinmetz, N. F. Tobacco Mosaic Virus Delivery of Phenanthriplatin for Cancer Therapy. *ACS Nano* **2016**, *10* (4), 4119–4126.
(20) Miller, R. A.; Presley, A. D.; Francis, M. B. Self-Assembling Light-Harvesting Systems from Synthetically Modified Tobacco Mosaic Virus Coat Proteins. *J. Am. Chem. Soc.* **2007**, *129* (11), 3104–3109.
(21) Zhao, X.; Chen, L.; Luckanagul, J. A.; Zhang, X.; Lin, Y.; Wang, Q. Enhancing Antibody Response against Small Molecular Hapten with Tobacco Mosaic Virus as a Polyvalent Carrier. *ChemBioChem* **2015**, *16* (9), 1279–1283.
(22) Carrico, Z. M.; Farkas, M. E.; Zhou, Y.; Hsiao, S. C.; Marks, J. D.; Chokhawala, H.; Clark, D. S.; Francis, M. B. N-Terminal Labeling of Filamentous Phage to Create Cancer Marker Imaging Agents. *ACS Nano* **2012**, *6* (8), 6675–6680.
(23) Chen, L.; Wu, Y.; Lin, Y.; Wang, Q. Virus-Templated FRET Platform for the Rational Design of Ratiometric Fluorescent Nanosensors. *Chem. Commun.* **2015**, *51* (50), 10190–10193.
(24) Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C.; Huang, P.-S.; Ravichandran, R.; et al. Design of a Hyperstable 60-Subunit Protein Icosahedron. *Nature* **2016**, *535* (7610), 136–139.
(25) Dedeo, M. T.; Duderstadt, K. E.; Berger, J. M.; Francis, M. B. Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. *Nano Lett.* **2010**, *10* (1), 181–186.
(26) Chariou, P. L.; Lee, K. L.; Pokorski, J. K.; Saidel, G. M.; Steinmetz, N. F. Diffusion and Uptake of Tobacco Mosaic Virus as Therapeutic Carrier in Tumor Tissue: Effect of Nanoparticle Aspect Ratio. *J. Phys. Chem. B* **2016**.
(27) Donev, A.; Cisse, I.; Sachs, D.; Variano, E. A.; Stillinger, F. H.; Connelly, R.; Torquato, S.; Chaikin, P. M. Improving the Density of Jammed Disordered Packings Using Ellipsoids. *Science* **2004**, *303* (5660), 990–993.
(28) Adriani, G.; de Tullio, M. D.; Ferrari, M.; Hussain, F.; Pascazio, G.; Liu, X.; Decuzzi, P. The Preferential Targeting of the Diseased Microvasculature by Disk-like Particles. *Biomaterials* **2012**,

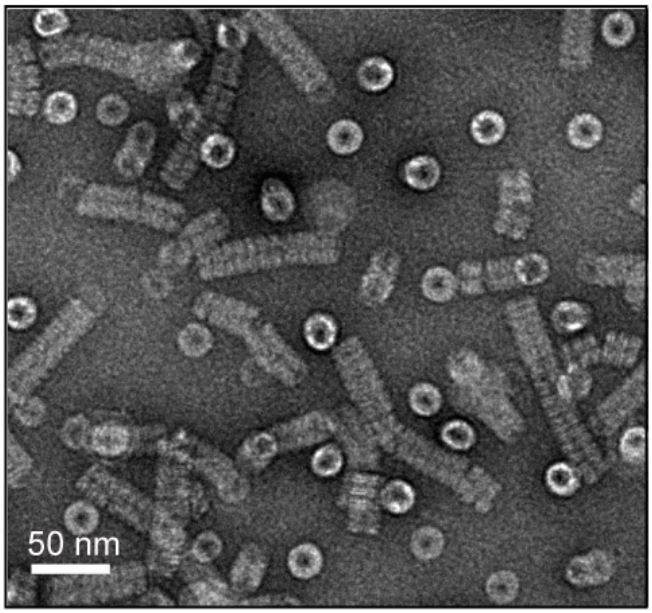
33 (22), 5504–5513.
(29) Agarwal, R.; Singh, V.; Jurney, P.; Shi, L.; Sreenivasan, S. V.; Roy, K. Mammalian Cells Preferentially Internalize Hydrogel Nanodiscs over Nanorods and Use Shape-Specific Uptake Mechanisms. *Proc. Natl. Acad. Sci.* **2013**, *110* (43), 17247–17252.
(30) Kinnear, C.; Moore, T. L.; Rodriguez-Lorenzo, L.; Rothen-Rutishauser, B.; Petri-Fink, A. Form Follows Function: Nanoparticle Shape and Its Implications for Nanomedicine. *Chem. Rev.* **2017**, *117* (17), 11476–11521.
(31) Leberman, R.; Finch, J. T.; Gilbert, P. F. C.; Witz, J.; Klug, A. TMV Crystal Structure. *J. Mol. Biol.* **1974**, *86*, 179–180.
(32) Petterson, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.
(33) Obermeyer, A. C.; Jarman, J. B.; Francis, M. B. N-Terminal Modification of Proteins with O-Aminophenols. *J. Am. Chem. Soc.* **2014**, *136* (27), 9572–9579.
(34) Shen, B.-Q.; Xu, K.; Liu, L.; Raab, H.; Bhakta, S.; Kenrick, M.; Parsons-Reponte, K. L.; Tien, J.; Yu, S.-F.; Mai, E.; et al. Conjugation Site Modulates the in Vivo Stability and Therapeutic Activity of Antibody-Drug Conjugates. *Nat. Biotechnol.* **2012**, *30* (2), 184–189.
(35) Gillies, E. R.; Fréchet, J. M. J. pH-Responsive Copolymer Assemblies for Controlled Release of Doxorubicin. *Bioconjug. Chem.* **2005**, *16* (2), 361–368.
(36) Cao, J.; Guenther, R. H.; Sit, T. L.; Opperman, C. H.; Lommel, S. A.; Willoughby, J. A. Loading and Release Mechanism of Red Clover Necrotic Mosaic Virus Derived Plant Viral Nanoparticles for Drug Delivery of Doxorubicin. *Small* **2014**, *10* (24), 5126–5136.
(37) Tong, G.; Lee, W. W.; Black, D. R.; Henry, D. W. Adriamycin Analogs. Periodate Oxidation of Adriamycin. *J. Med. Chem.* **1976**, *19* (3), 395–398.
(38) Sastry, C. S. P.; Lingeswara Rao, J. S. V. M. Determination of Doxorubicin Hydrochloride by Visible Spectrophotometry. *Talanta* **1996**, *43* (11), 1827–1835.
(39) Bernard, J. M. L. High-Aspect Ratio Protein-Based Carriers for Delivery Applications, University of California Berkeley, 2016.
(40) Aanei, I. L. Protein-Based Nanoparticles for Imaging and Drug Delivery Applications, University of California Berkeley, 2016.
(41) Jensen, S. A.; Day, E. S.; Ko, C. H.; Hurley, L. A.; Luciano, J. P.; Kouri, F. M.; Merkel, T. J.; Luthi, A. J.; Patel, P. C.; Cutler, J. I.; et al. Spherical Nucleic Acid Nanoparticle Conjugates as an RNAi-Based Therapy for Glioblastoma. *Sci. Transl. Med.* **2013**, *5* (209), 209ra152.
(42) Allard, E.; Passirani, C.; Benoit, J.-P. Convection-Enhanced Delivery of Nanocarriers for the Treatment of Brain Tumors. *Biomaterials* **2009**, *30* (12), 2302–2318.
(43) Zhu, Y.; Zhang, J.; Meng, F.; Deng, C.; Cheng, R.; Feijen, J.; Zhong, Z. cRGD-Functionalized Reduction-Sensitive Shell-Sheddable Biodegradable Micelles Mediate Enhanced Doxorubicin Delivery to Human Glioma Xenografts in Vivo. *J. Control. Release* **2016**, *233*, 29–38.
(44) Kozlovskaya, V.; Xue, B.; Kharlampieva, E. Shape-Adaptable Polymeric Particles for Controlled Delivery. *Macromolecules* **2016**, *49* (22), 8373–8386.
(45) van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. Supramolecular Chemistry Targeting Proteins. *J. Am. Chem. Soc.* **2017**, *139* (40), 13960–13968.
(46) Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115* (22), 12320–12406.
(47) Tian, F.; Jiao, D.; Biedermann, F.; Scherman, O. A. Orthogonal Switching of a Single Supramolecular Complex. *Nat. Commun.* **2012**, *3* (1), 1207.

5.11 Additional Figures



10 mM NaPhos pH 7.0

Figure S5.1. Transmission electron microscopy of RR-TMV disks in 10 mM pH 7.0 NaPhos buffer.



50 mM NaOAc pH 5.5

Figure S5.2. Transmission electron microscopy of RR-TMV disks in 50 mM pH 5.5 NaOAc buffer.

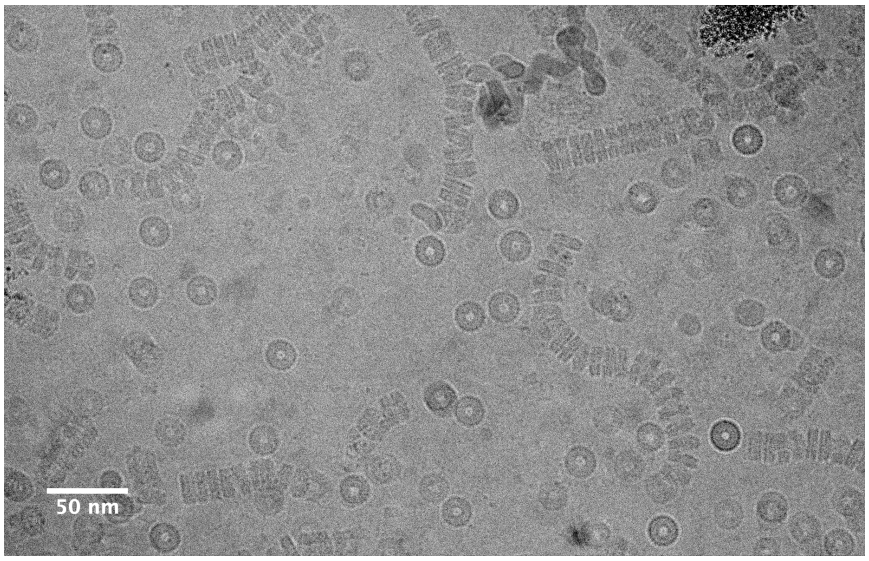


Figure S5.3. Cryogenic transmission electron microscopy of RR-TMV disks in 10 mM pH 7.0 NaPhos buffer.

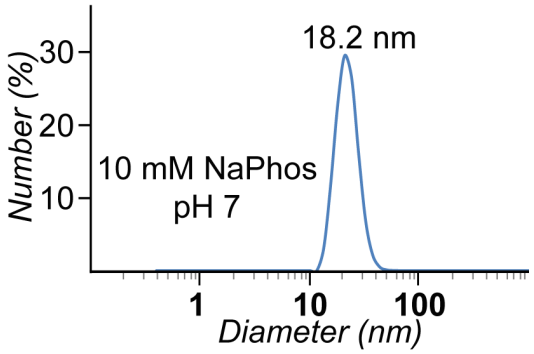


Figure S5.4. Dynamic light scattering (DLS) indicated a 18.2 nm diameter of RR-TMV disks.

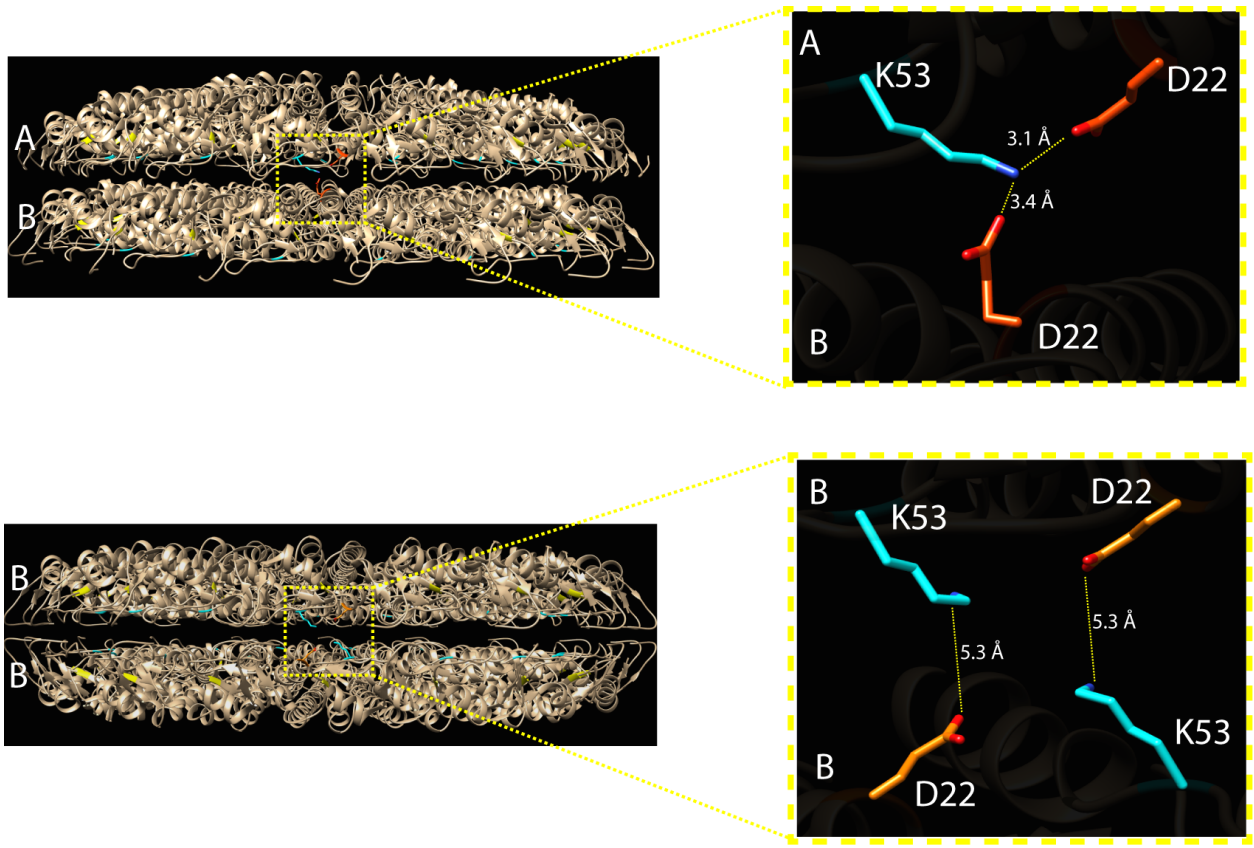


Figure S5.5. Modeling of wtTMV salt bridges. UCSF Chimera[S2] was used to locate any natural salt bridges present in wtTMV at position K53. AB disk stacking revealed potential salt bridges, while BB stacking did not contain any salt bridges within the literature reported salt bridge distance of 4 Å.[S3] PDB file 1EI7 (wtTMV) was used for modelling. No salt bridges were observed at K68.

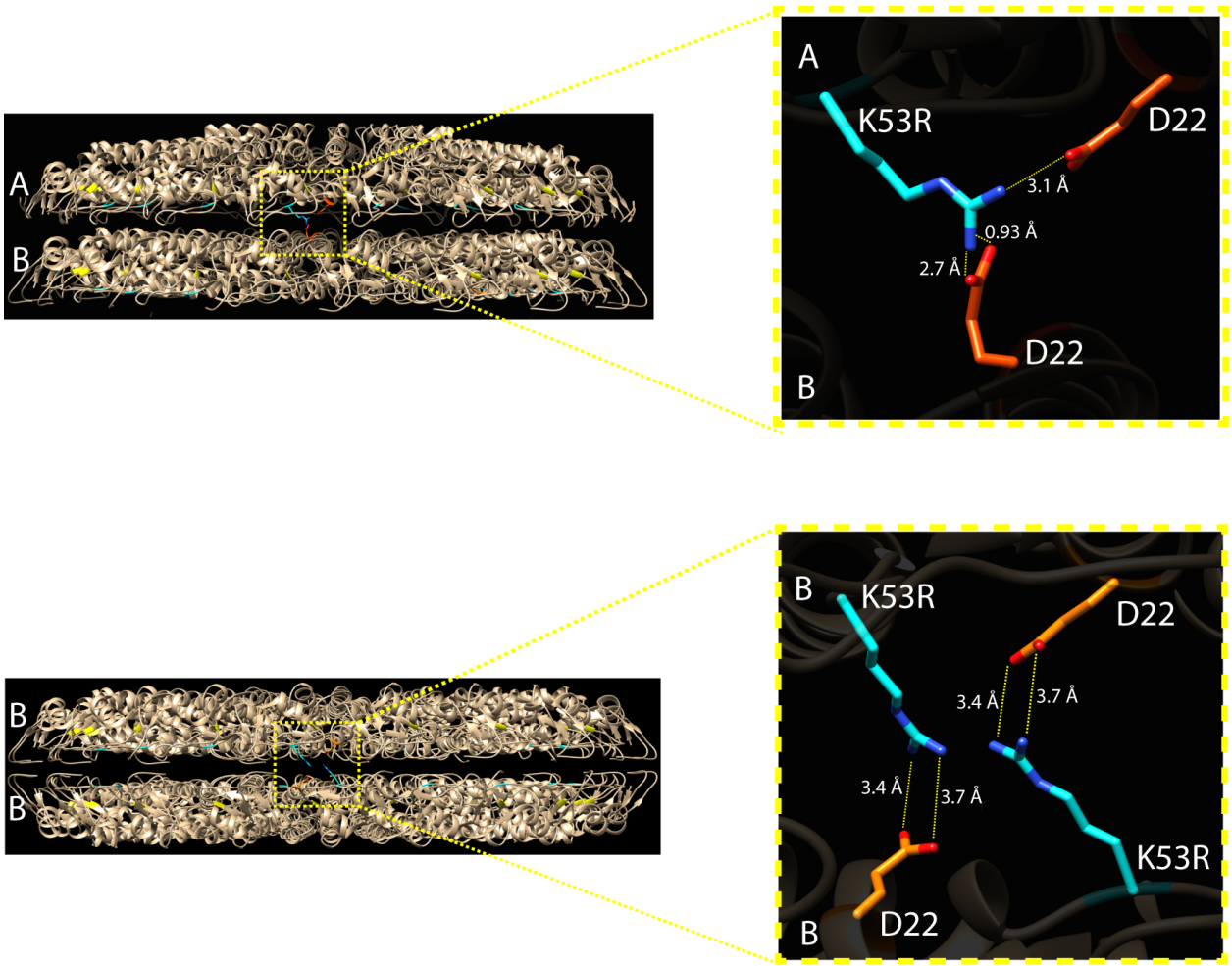


Figure S5.6. Modeling of RR-TMV salt bridges. UCSF Chimera[S2] was used to locate any salt bridges present in RR-TMV at position K53R. AB and BB disk stacking revealed potential salt bridges in accordance with the literature reported salt bridge distance of 4 Å.[S3] PDB file 1EI7 (wtTMV) mutated with K53R and K68R was used for modelling. No salt bridges were observed at K68R.

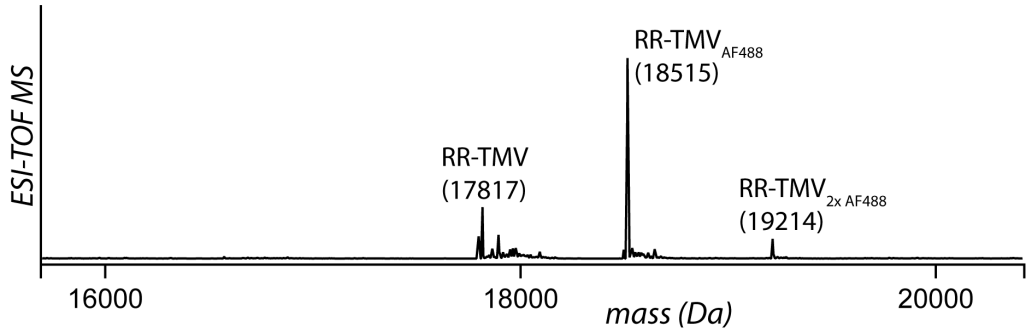


Figure S5.7. AF488 modification of RR-TMV. RR-TMV (expected mass of 17817 Da) was modified with AF488-maleimide (18515 Da). A slight second addition of AF488 (19214 Da) was also observed.

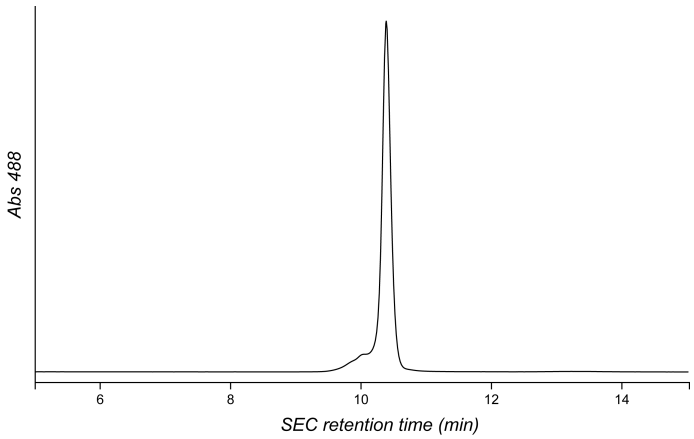


Figure S5.8. SEC analysis of AF488-maleimide**.** An HPLC-SEC analysis of AF488-maleimide revealed a retention time of ~10.2 min.

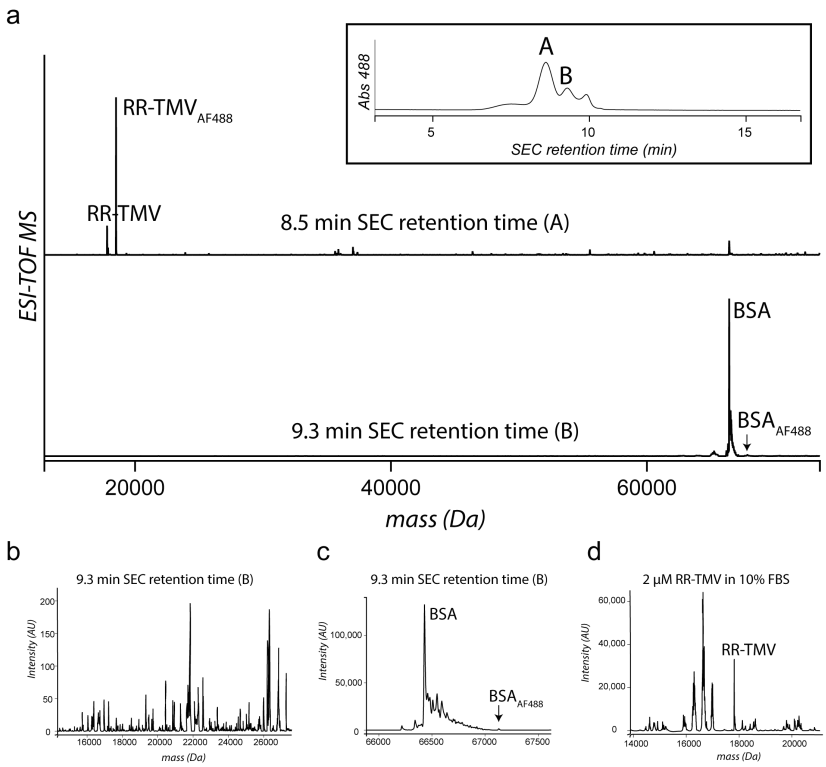


Figure S5.9. ESI-TOF MS analysis of serum stability samples. **(a)** Fractions at 8.5 and 9.3 min SEC retention times were collected and subjected to ESI-TOF MS analysis. RR-TMV and RR-TMV$_{AF488}$ were observed in the 8.5 min fraction (A). Small traces of BSA were also observed. In the 9.3 min fraction (B), only BSA was observed, with a very small BSA$_{AF488}$ peak at 67130 Da. No RR-TMV was observed in the 9.3 min fraction. **(b)** In-depth analysis of TMV MS range for the 9.3 min fraction. No MS peak corresponding to RR-TMV was observed, even in the noise intensity range. **(c)** Zoomed-in analysis of BSA MS range for the 9.3 min fraction. A small peak corresponding to BSA$_{AF488}$ (67130 Da) was observed with a ~1500 intensity (AU). **(d)** RR-TMV (2 µM) was incubated in 10% FBS to determine if lower concentrations of RR-TMV could be detected by MS in the presence of serum. A small but detectable peak corresponding to the mass of RR-TMV was observed. No such peak was observed in the 9.3 minute SEC fraction.

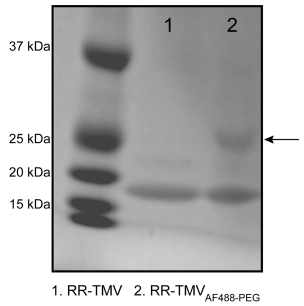


Figure S5.10. Gel electrophoresis analysis of $TMV_{AF488-PEG}$. $RR\text{-}TMV_{AF488}$ was PEGylated at the N-terminus with aminophenol-PEG_{5k}. Approximately 30% PEGylation was observed.

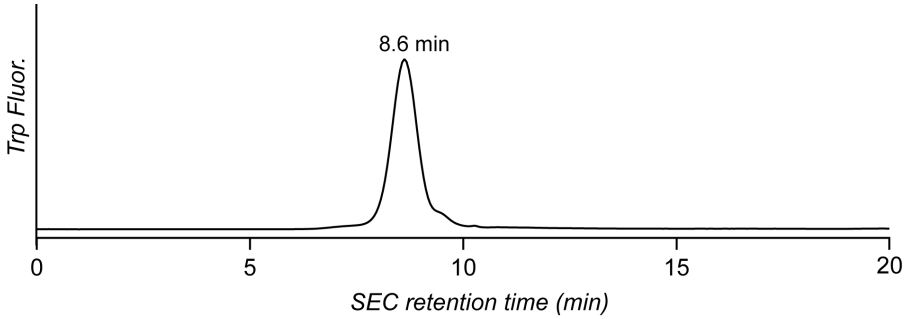


Figure S5.11. Stability of RR-TMV disks in PBS after 5 d. RR-TMV was incubated at 37 °C for 5 d and analyzed by HPLC-SEC. Disks were the only assembly state observed.

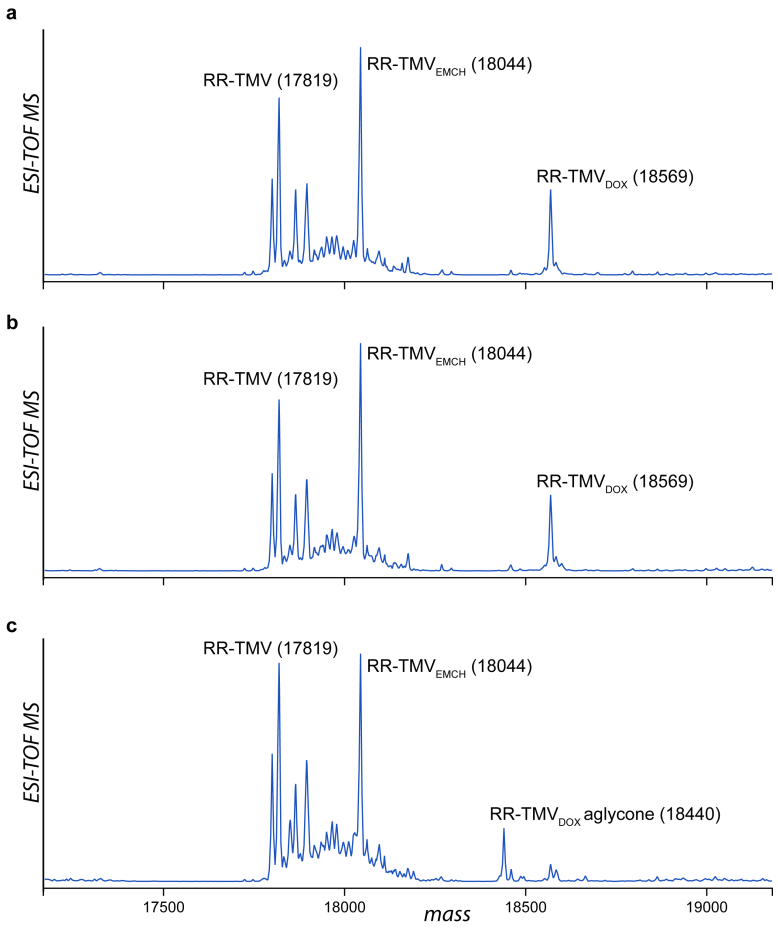


Figure S5.12. Treatment of RR-TMV$_{DOX}$ with $NaIO_4$ and $K_3Fe(CN)_6$. (**a**) ESI-TOF mass spectrum of RR-TMV$_{DOX}$. The RR-TMV$_{EMCH}$ peak is likely due to DOX release in the 0.1% formic acid mobile phase. Additional peaks were observed at 17865 and 17895 Da. (**b**) Treatment with $K_3Fe(CN)_6$ did not cause any observable mass changes by ESI-TOF MS analysis. (**c**) Treatment with $NaIO_4$ caused a mass loss of 129 Da, corresponding with the glycolysis of doxorubicin.

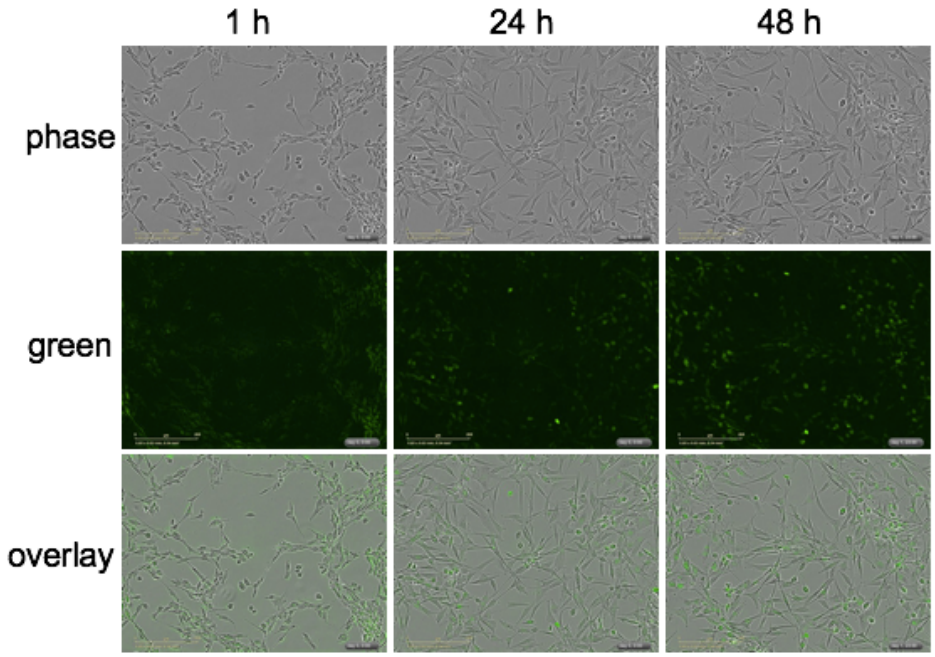
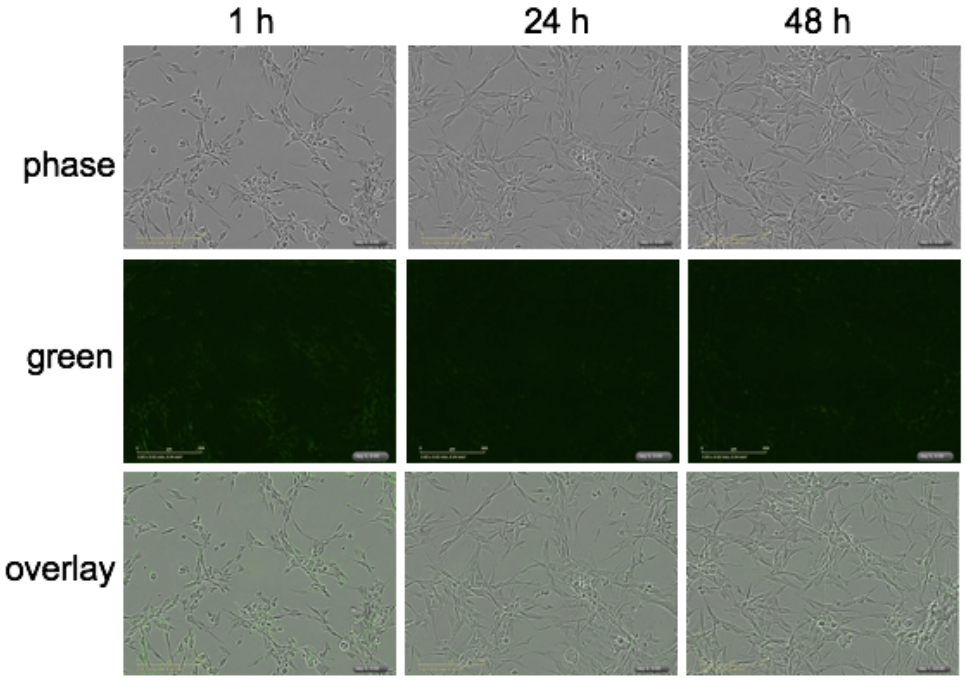


Figure S5.12. Cell uptake of RR-TMV$_{DOX\text{-}PEG}$**.** Fluorescence microscopy images of U87MG cells incubated with RR-TMV$_{DOX\text{-}PEG}$ at 1, 24, and 48 h incubation times.



Figures S5.13. Cell uptake of RR-TMV to check for background fluorescence and confirm normal cell morphology. Fluorescence microscopy images of U87MG cells incubated with RR-TMV at 1, 24, and 48 h incubation times.

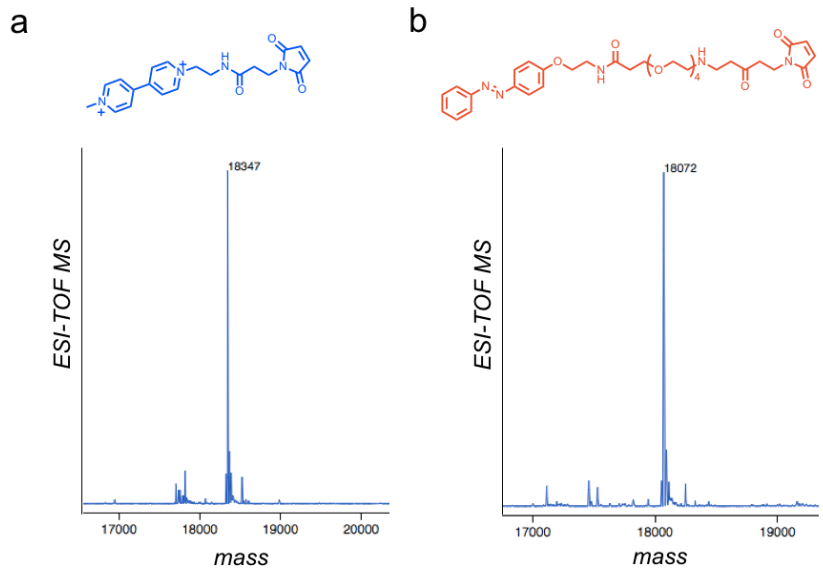


Figure S5.14. Modification of cpTMV with CB8 guests. Modification of cpTMV with (a) methylviologen-maleimide and (b) azobenzene maleimide led to complete conversion to the MV-TMV and Az-TMV conjugates, respectively.

Nanocarrier	**DOX attachment site**	**DOX modification**	**PEG attachment site**	**PEG modification %**
MS2 spheres	*87C (interior)*	*35% (63 per capsid)*	*T19pAF (exterior)*	*70% (126 per capsid)*
RR-TMV disks	*123C (exterior face)*	*30% (10 per capsid)*	*N-term Pro (exterior edge)*	*50% (17 per capsid)*
Nanophage rods	*Lys (exterior)*	*20% (20 per capsid)*	*Lys (exterior)*	*15% (15 per capsid)*

Table S5.1. Attachment strategies and modification percentages for protein-based nanomaterials