

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Bioconjugate materials for the study of pigment mobility in light-harvesting systems, protein-based formulations for hydrophobic actives, and conformational changes in conjugate vaccines

### Permalink

<https://escholarship.org/uc/item/1637m4hd>

### Author

Jaffe, Jake

### Publication Date

2016

Peer reviewed|Thesis/dissertation

Bioconjugate materials for the study of  
pigment mobility in light-harvesting systems,  
protein-based formulations for hydrophobic actives,  
and conformational changes in conjugate vaccines

by

Jake Jaffe

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 Felix R. Fischer  
Professor Douglas S. Clark

Spring 2016

Bioconjugate materials for the study of  
pigment mobility in light-harvesting systems,  
protein-based formulations for hydrophobic actives,  
and conformational changes in conjugate vaccines

Copyright © 2016

By: Jake Jaffe

## **Abstract**

Bioconjugate materials for the study of pigment mobility in light-harvesting systems, protein-based formulations for hydrophobic actives, and conformational changes in conjugate vaccines

by

Jake Jaffe

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

Techniques for the preparation, purification, and characterization of protein-based materials have allowed for advances in fields ranging from medicine to materials science. While great attention has been paid to the chemistries used in the attachment of molecules of interest to proteins, the importance of the linking group is often overlooked. Each chapter of this dissertation describes the use of linkers as functional design elements in three distinct projects. In Chapter One, the development and characterization of a minimal model for investigating the role of pigment mobility in photosynthetic light-harvesting antenna systems will be discussed. In this model system, pigment-protein and pigment-pigment interactions were altered through the use of pigment-protein linkers of various lengths and rigidities. Chapter Two explores new biocompatible bond cleavage reactions for the preparation of a new class of general protein-based formulations for hydrophobic actives, wherein the linker group imparts the critical properties to the system. Chapter Three will examine conformational changes in peptide-protein conjugate vaccines. The roles of specific methods of covalent modification and linker composition in these conformational changes are emphasized.

In memory of  
Bernard "Saba" Jaffe

## Table of Contents

### Chapter 1: Linker effects on energy transfer interactions in well-defined pigment-protein complexes

1.1 Introduction .....	2
1.2 Selection and preparation of the protein scaffold .....	2
1.3 Suitability of Mth1491 as a scaffold for pigment-protein complexes .....	4
1.4 Selection and preparation of pigments.....	5
1.5 Determination of pigment mobility.....	6
1.6 Effects of pigment mobility on pigment-pigment interactions .....	9
1.7 Spectral properties of coumarin 343-Mth1491 conjugates.....	9
1.8 Two-color Mth1491 conjugates .....	11
1.9 Towards homogeneous pigment-Mth1491 conjugates.....	12
1.9.1 Synthesis of bifunctional coumarin reagent for chromatography-mediated bioconjugation .....	12
1.9.2 Preparation and characterization of pigment-protein complexes using chromatography-mediated bioconjugation.....	14
1.10 Effects of linker composition in tobacco mosaic virus-templated pigment-protein complexes .....	17
1.11 Conclusions.....	18
1.12 Materials and methods .....	18
1.12.1 General procedures and methods.....	18
1.12.2 Instrumentation and sample analysis .....	18
1.12.3 Experimental procedures.....	20
1.12.4 Small molecule synthesis .....	22
1.13 References.....	32

### Chapter 2: Development of general protein-based formulations for hydrophobic actives using hydrophobicity modulators

2.1 Introduction .....	37
2.1.1 Virus-like particles as drug delivery vehicles.....	37
2.1.2 Hydrophobicity modulation as a strategy to prepare general protein nanocarriers.....	38
2.2 Interior surface hydrophobicity modulation of MS2.....	39
2.2.1 Synthesis of thiol-reactive $\pi$ -allylpalladium substrate .....	39
2.2.2 Conjugation and hydrophobicity modulation with allyl phosphate <b>2.1</b> .....	42

2.2.3	Synthesis of a photo-induced hydrophobicity modulator.....	45
2.2.4	Hydrophobicity modulation with photocleavable substrate <b>2.15</b> .....	47
2.2.5	Characterization of surface hydrophobicity.....	49
2.2.6	Improving the biocompatibility of Norrish type II chemistry.....	52
2.3	Towards creating a hydrophobic droplet within MS2 .....	58
2.4	Towards hybrid hydrophobicity-modulating polymer-protein hybrids .....	63
2.5	Conclusions.....	66
2.6	Materials and methods .....	66
2.6.1	General procedures and methods.....	66
2.6.2	Instrumentation and sample analysis.....	67
2.6.3	Experimental procedures.....	68
2.6.4	Small molecule synthesis .....	72
2.7	References.....	88

### **Chapter 3: Elucidating the nature of conformational changes in peptide-CRM<sub>197</sub> conjugate vaccines**

3.1	Introduction .....	95
3.2	Conformational changes in CRM <sub>197</sub> conjugates .....	96
3.2.1	Effects of the conjugation step on conformational change .....	98
3.2.2	Effects of the activation step on conformational change.....	99
3.3	Alternative activation-conjugation strategies.....	102
3.3.1	Maleimide-NHS ester crosslinking .....	102
3.3.2	CuAAC and SPAAC .....	104
3.3.3	Disulfide exchange.....	104
3.4	Conclusions.....	106
3.5	Materials and methods .....	106
3.5.1	General procedures and methods.....	106
3.5.2	Instrumentation and sample analysis.....	106
3.5.3	Experimental procedures.....	107
3.5.4	Small molecule synthesis .....	110
3.6	References.....	111

## Acknowledgments

I would like to thank the following people who have supported, guided, and accompanied me through my stint as a graduate student at Berkeley. I couldn't have done it without:

Matt Francis, thank you for giving me the opportunity to work in your lab. Your constant enthusiasm and wealth of knowledge were inspiring.

Thorsten, it has been a pleasure working with you through the CARA program. Thanks for the research and career guidance, and the good times!

Anouk, thank you and the Pfizer team for the opportunity to work on such an exciting and challenging project, and for all of your insight and assistance.

The Francis lab members who welcomed me as a first year and set me up for success in the world of bioconjugation: Abby, Allie, Amy, Chris, Dan, Gary, Jeff, Jelly, Kareem, Katherine, Kristen, Leah, Michel, Michelle, Mike, Stacy, Troy, Wesley, and Zac. A special thanks to Allie and Michel for teaching me everything I know about molecular biology.

A special shout out to the "original" members of 748: Dan, Katherine, Richard, and Stacy, you made coming in to lab worth it! Dan, thanks for mentoring me on the light-harvesting project. Katherine, I don't think I would've been able to stay caffeinated without you. Stacy, your dedication to your work meant that I was never alone in the lab. Richard, although I enjoyed getting close to you, I'm glad we parted and got our own desks. Josh, while you were only a temporary resident of 748, you were a great help and became quite the expert on the Feely-Beavers reaction.

My cohort, Ioana, Jenna, Jim, Kanwal, and Richard, we did it! Ioana, your professionalism and to-the-point demeanor were greatly appreciated. Jenna, your mixology really took our group gatherings to the next level. Jim, thanks to you I always kept on top of the lab gossip. Kanwal, you've been an excellent sounding board and a great friend to me and Danielle.

The flood of postdocs, Adel, Alex, Ariel, Ayo, Carson, Christian, Henrik, Meera, and Rafi, and new students, Adam, Am, Emily, Jess, Jing, Joel, Kristin, Marco, Matt Smith, Rachel, and Sarah, who joined after me have reenergized the lab, have taught me so much, and will surely keep research in the Francis group going strong. Jing, I hope I've set you up well to successfully harvest light. I know the project is in good hands. Kristin, I'm positive you will do great things with CRM. Best of luck to all of you!

Of course, I never would have had the opportunity to attend grad school without the lifetime of love, support, and encouragement from my family. Especially from my parents, Vicki and Jeff, and my grandparents, Bubbie and Zaidie, and Safta and Saba. I am incredibly thankful for all that you have done!

While I never would have made it to grad school without the support of my family, I owe the fact that I made it through to my partner and best friend, Danielle. I look forward to seeing what the two of us achieve together in the coming years.



## Chapter 1

### Linker effects on energy transfer interactions in well-defined pigment-protein complexes

#### **Abstract**

Natural photosynthetic antenna systems maintain strict orientational control of scaffolded pigments, a functional design element overlooked in synthetic analogs. This chapter describes the development of a minimal model system for the study of pigment-protein linker composition on pigment-protein and pigment-pigment interactions in light-harvesting antenna systems. Using a series of cysteine-reactive coumarin dyes, a collection of well-defined pigment-protein conjugates were prepared from a trimeric protein. Pigment mobility was determined in these constructs through fluorescence anisotropy studies and its effects on energy transfer processes were investigated. Through these studies, we have demonstrated the importance of linker composition as a design element for artificial light-harvesting systems.

## 1.1 Introduction

Photosynthetic organisms rely on highly-ordered chromophore arrays to capture and transfer solar energy to reaction centers for conversion into chemical potential energy.<sup>1</sup> These arrays are generally pigment-protein complexes and are capable of efficiently harvesting broad-spectrum light with a limited variety of chromophore types—chlorins, bacteriochlorins, carotenoids, and bilins—dependent on the specific class of phototroph.<sup>2</sup> Different groups of photosynthetic organisms have evolved protein-scaffolded antenna systems of vastly different sizes, varied patterns of pigment arrays, and specific pigment subtypes.<sup>3</sup> However, within all of these pigment-protein complexes, the chromophores are frozen translationally and rotationally, giving rise to what are thought to be nearly static inter-pigment orientations. This ordering of chromophores results in specific pigment-protein and pigment-pigment interactions, altering the photophysical properties of the individual chromophores to utilize the solar spectrum effectively and to funnel the collected energy through Förster resonance energy transfer (FRET) towards reaction centers.

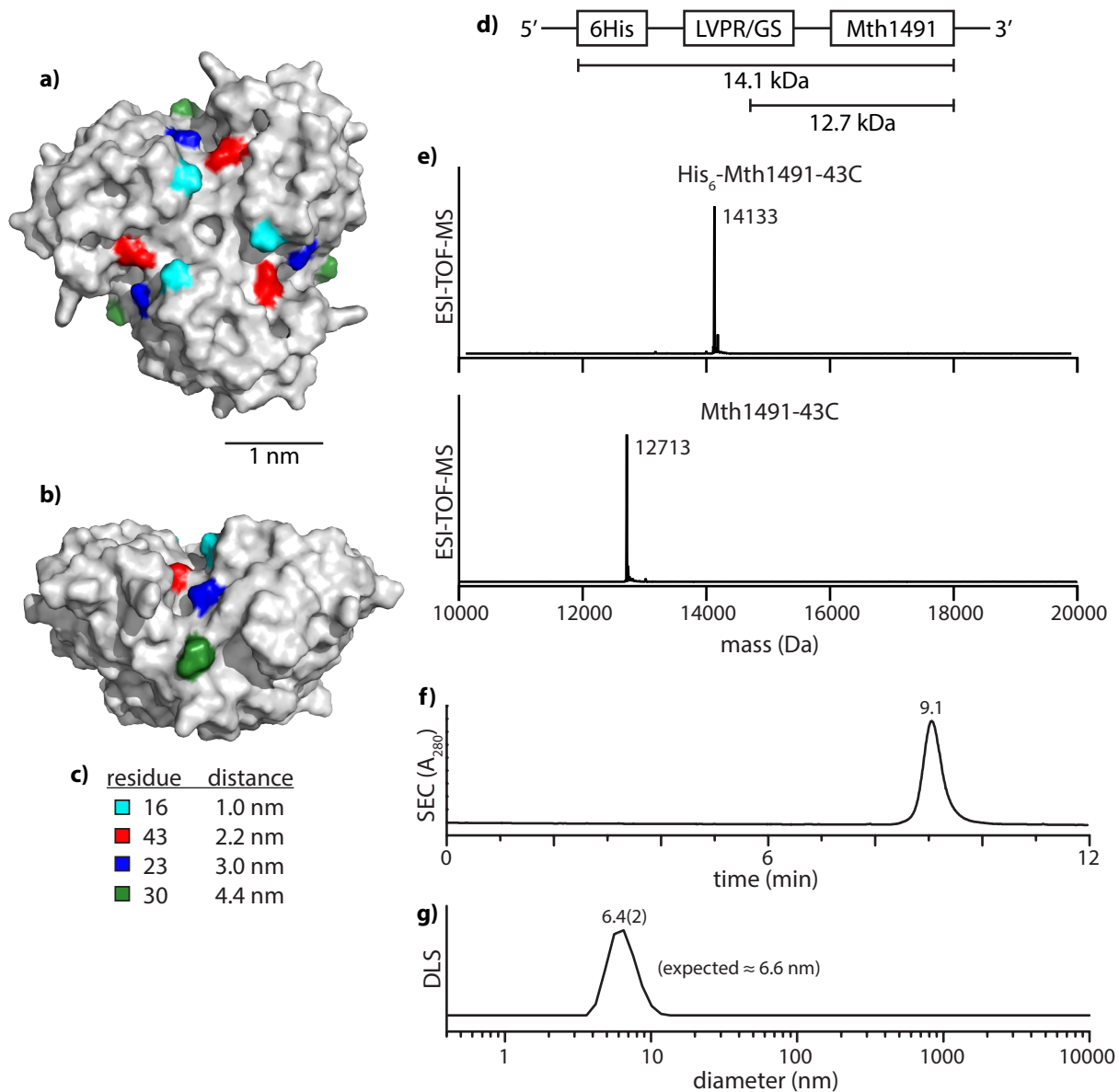
To study light-harvesting systems, many groups have devised synthetic mimics of these systems.<sup>4-6</sup> By doing so, the chromophore component can be altered to meet desired spectral input characteristics. Particularly successful systems have been based on dendrimers,<sup>7-9</sup> polymers,<sup>10,11</sup> nanotubes,<sup>12</sup> and surfaces.<sup>13,14</sup> The Francis group has explored tobacco mosaic virus coat proteins (TMVcp) as a means to establish the interchromophore distances, as the rigidity inherent in these systems allows for greater positional specificity.<sup>15-19</sup> Other protein systems have also proven useful in this regard.<sup>20-22</sup>

However, the chromophores used in virtually all synthetic light-harvesting systems are generally able to rotate freely relative to the scaffold surface and the other chromophores. This is in stark contrast to the natural systems discussed above. Although critical to the function of natural photosynthetic antenna systems, the functional design element of constrained chromophore mobility has not been explored systematically in synthetic systems.

In this work, a series of pigment-protein bioconjugates was designed specifically to probe the effects of covalently-bound pigment mobility on pigment-protein and pigment-pigment interactions. In order to carry out this study in a systematic manner, the development of a new minimal protein scaffold for preparing pigment-protein conjugates, methods for determining chromophore mobility, and a method for preparing conjugates with controlled levels of modification were required and will be discussed in this chapter.

## 1.2 Selection and preparation of the protein scaffold

To study the effects of pigment mobility on pigment-protein and pigment-pigment interactions in artificial protein-based light-harvesting systems,<sup>15-19</sup> a new protein scaffold was required. While we have extensive experience using assemblies of TMV capsid proteins for studying light-harvesting, these systems are too large and complex for a systematic study of pigment mobility. The ideal scaffold for these studies would be a small (relative to viral capsid assemblies) homomultimeric protein with known structure, high stability, and high processability. An initial keyword-based search of the nearly 120,000 crystal structures in the Protein Data Bank (PDB) resulted in excessively large sets of potential scaffolds.<sup>23</sup> To narrow down potential scaffolds, the online 3D Complex hierarchical classification tool was used.<sup>24</sup> Using 3D Complex, a search for homomultimeric proteins with 3–10 subunits was performed, narrowing down the number of



**Figure 1.1.** Preparation of Mth1491 for use as a scaffold in pigment-protein conjugates. a) Top and b) side views of the Mth1491 trimer with 16C, 43C, 23C, and 30C mutations shown in cyan, red, blue, and green, respectively. c) Inter-thiol distances were determined from the crystal structure. d) A plasmid was prepared containing the Mth1491 gene with an N-terminal His<sub>6</sub>-tag separated by a thrombin cleavage site. Recombinantly expressed Mth1491 was characterized by ESI-TOF-MS e) before and f) after cleavage of the His<sub>6</sub>-tag. Mth1491 was found to exist exclusively in the trimeric state by f) SEC and g) DLS.

potential protein scaffolds to just over 4000. In an attempt to limit this set to proteins of high stability and processability, this collection of trimers to decamers was further minimized by selecting proteins originating from thermophilic and hyperthermophilic organisms. This final set consisted of only 234 structures, which could be manually parsed. Ultimately, the trimeric hypothetical protein Mth1491 from the thermophilic archaea *Methanobacterium thermoautorophicum* was chosen for these studies (PDB ID: 1L1S, **Figure 1.1**).<sup>25</sup> Gel permeation chromatography experiments have shown that this multimeric protein does not dissociate to a significant extent ( $K_d < 4 \times 10^{-19} \text{ M}^2$ ) and

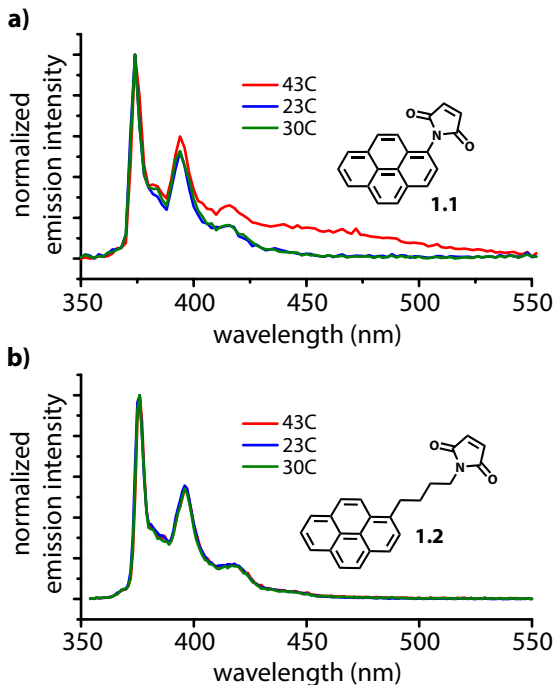
thus provides an appropriate scaffold for a model pigment-protein.<sup>26</sup> Mth1491 does not contain any endogenous solvent-accessible cysteines, allowing for incorporation of a reactive cysteine at a location of interest on each monomer using site-specific genetic mutation without having to first generate cysteine-knockout mutants. Additionally, as there are no lysines in Mth1491, Cys-Lys double mutants could be envisioned for preparing multi-dye protein complexes. However, the features that ultimately led to the choice of Mth1491 as a scaffold in these studies are the three deep radial grooves along its top surface. These surface features could provide pockets in which the motion of pendant chromophores could be constrained at roughly 120° angles relative to one another.

The presence of a solvent-accessible cysteine allows for the incorporation of appropriately modified chromophores through reliable thiol-maleimide chemistry. A plasmid containing a gene for Mth1491 with an N-terminal hexahistidine tag, separated by a thrombin cleavage site, was generated (**Figure 1.1d**). Four mono-Cys mutants of Mth1491 were prepared through site-selective mutagenesis and expressed recombinantly in *E. coli*. Each mutant possessed one solvent-exposed cysteine per monomer: 1) the A30C mutant has cysteines at the shallow distal end of the groove with an inter-cysteine distance of approximately 4.4 nm; 2) the S23C mutant displays cysteines deep in the middle of the radial groove with an inter-cysteine spacing of 3.0 nm; 3) the S43C mutant has cysteines situated 2.2 nm apart deep in the groove near the proximal end; and 4) the S16C mutant has cysteines situated 1.4 nm apart near the central axis of the complex. These mutants were chosen, as they provide varied topographies with which pendant chromophores could interact sterically. Additionally, they provide a range of inter-site distances to modulate the degree of FRET between chromophores. Expression of wt-Mth1491 and the 23C, 30C, and 43C mutants was successful, providing 20–40 mg of purified protein per liter of culture. However, Mth1491 16C was found to express at much lower yields (5 mg/L) and, as was later discovered, the cysteine in each monomer was found to be generally unreactive towards maleimides. As such, the 16C mutant was not used further in this study.

To ensure that the three mutants maintained the same quaternary structure as wt-Mth1491, samples of each were analyzed by size exclusion chromatography (SEC) and dynamic light scattering (DLS). The SEC elution profile of each mutant matched that of wt-Mth1491 exactly, all showing a single peak corresponding to the trimer. DLS revealed a single particle size with a hydrodynamic radius of 6.4(2) nm for all mutants, which is consistent with the longest atom-to-atom distance in the crystal structure (6.6 nm).

### 1.3 Suitability of Mth1491 as a scaffold for pigment-protein complexes

To ensure that Mth1491-bound chromophores would not suffer from contact quenching, pyrene labeling studies were performed. When in close proximity (subnanometer separation), pyrene molecules are known to form excimers with very distinct emission properties.<sup>27</sup> In contrast to the multiple sharp emission peaks around 360–450 nm exhibited by monomeric pyrene, the pyrene-pyrene excimer has an emission spectrum characterized by a broad peak red-shifted closer to 500 nm. Two pyrene-based probes with varying linker lengths were used to examine the potential for pigment-pigment contact on the Mth1491 mutants. First, commercially available *N*-(1-pyrenyl)maleimide **1.1**, a zero-length cysteine-reactive pyrene probe, was conjugated to the Mth1491 mutants. High levels of modification were observed for the three Cys mutants, whereas wt-Mth1491 was unmodified. Characterization of the emission spectra of the three conjugates revealed no excimer formation for the 30C or 23C mutants, though a weak signal corresponding



**Figure 1.2.** Characterization of Mth1491 mutants modified with pyrene maleimides **1.1** and **1.2**. a) Fluorescence emission spectra of **1.1**-Mth1491 conjugates did not reveal excimer formation, except perhaps in the 43C mutant. b) No excimer formation was observed for **1.2**-Mth1491 conjugates. Samples were excited at 344 nm.

signal is only seen for the probe with the much more rigid zero-length linker, perhaps the signal is due to excitonic coupling as a result of the rigidity provided by the Mth1491 scaffold along with the relatively close proximity of the bound pyrenes, rather than excimer formation. Taken together, these findings provided sufficient evidence to suggest that these Mth1491 mutant scaffolds could be used to study pigment-pigment interactions in the FRET regime.

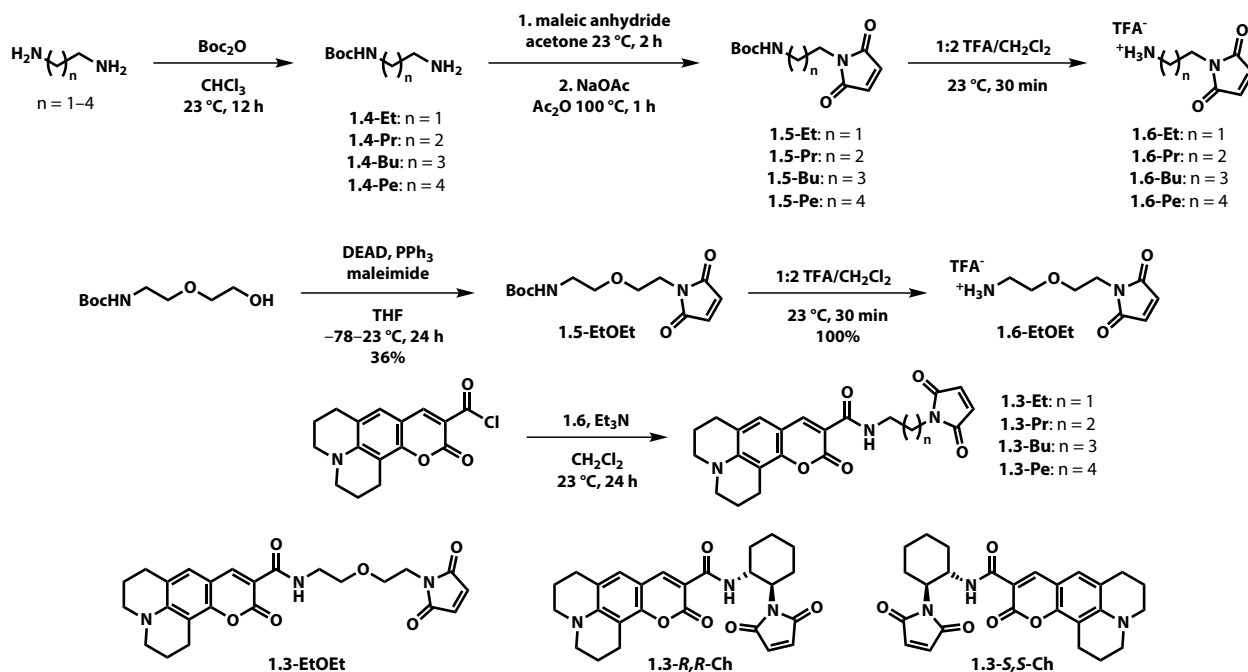
## 1.4 Selection and preparation of pigments

For this study of pigment mobility in pigment-protein complexes, coumarin 343 was chosen as the primary chromophore of interest. A number of properties were considered in this selection. This chromophore is ideal for use in a model system of photosynthetic antenna complexes as it has been studied extensively and its properties are well known. Its major electronic transitions are characterized and its transition dipole is well-defined.<sup>29</sup> Coumarin 343 is highly solvatochromatic, exhibiting a range of absorption and emission maxima, dependent on the dielectric constant of the media.<sup>30</sup> This solvatochromism could be used to probe the local environment of the protein-bound pigments. In addition to the photophysical properties of coumarin 343, it is highly tractable, synthetically, compared to other chromophores commonly used in bioconjugation (e.g., xanthene and cyanine dyes).

A collection of coumarin 343 dyes with pendant maleimides (**1.3**) attached through varying linkers was designed and synthesized (**Figure 1.3**). Coumarin 343 maleimides were synthesized

to excimer formation was observed in the 43C mutant (**Figure 1.2a**).

As the pigments that will be used for studying pigment mobility in this system have non-zero-length linkers (See Section 1.4), a pyrene maleimide probe with an extended flexible butyl linker, *N*-(4-(pyren-1-yl)butyl)maleimide **1.2**, was prepared from 1-pyrenebutanol using Mitsunobu chemistry.<sup>28</sup> Mth1491 conjugates were prepared using butylpyrene probe **1.2**, and were characterized by spectrofluorometry (**Figure 1.2b**). Unlike the pyrene probe **1.1** with a zero-length linker, conjugates prepared with butylpyrene **1.2** showed no evidence of excimer formation for any of the mutants. While it was clear that no pigment-pigment contact was occurring in the 23C and 30C mutants, excimer formation may be possible in conjugates of the 43C mutant. The finding that the 43C-**1.2** conjugate did not form excimers is strong evidence that pyrene-pyrene contact is not possible in this system. However, the small excimer signal seen for the 43C-**1.1** conjugate suggests otherwise. Since the excimer



**Figure 1.3.** Synthesis of coumarin 343 maleimides **1.3**. Cyclohexyl linkers **1.6-R,R-Ch** and **1.6-S,S-Ch** were prepared following the same synthetic route as the linear alkyl linkers shown.

starting from commercially available diamines. The diamines were first mono-Boc protected to afford **1.4**. Amines **1.4** were then reacted with maleic anhydride to form the respective maleamic acids, and cyclizing condensation was achieved in hot acetic anhydride with sodium acetate, to afford the desired maleimide functionality (**1.5**). Deprotection of Boc-protected amine **1.5** to *N*-aminoalkylmaleimide **1.6** followed by amide formation with coumarin 343 acid chloride afforded the desired alkyl-linked coumarin 343 maleimides (**1.3**). A series of alkyl linkers, ranging from two to five carbons, was incorporated into the **1.3** series. As the linker length increases, the chromophore would be expected to gain mobility through increasing degrees of conformational freedom and from potentially being able to rise above the walls of the protein groove. In addition to the series of linear linkers, a pair of enantiomeric 1,2-diaminocyclohexane-based linkers was prepared. These chiral linkers were designed both to act as highly rigid linkers and to create a pair of diastereomeric products when attached to the enantiopure protein surface, leading to distinct steric interactions and chromophore orientations without augmenting pigment-protein separation.<sup>31</sup> To control for effects stemming from strong hydrophobic interactions between the linkers and the protein scaffold, a five-atom ether-linked coumarin (**1.3-EtOEt**) was prepared as a control.

## 1.5 Determination of pigment mobility

In order to determine the effects of chromophore mobility on energy transfer, the degree of mobility was first determined for each coumarin maleimide conjugated to each of the three Mth1491 cysteine mutants. Fluorescence anisotropy provides a means of quantifying the relative mobility of protein-bound fluorophores through the use of linearly polarized excitation. Comparison of the polarization of the emitted signal that is parallel to and perpendicular to the excitation polarization provides a degree of anisotropy (*r*) corresponding to the offset in transition moments,

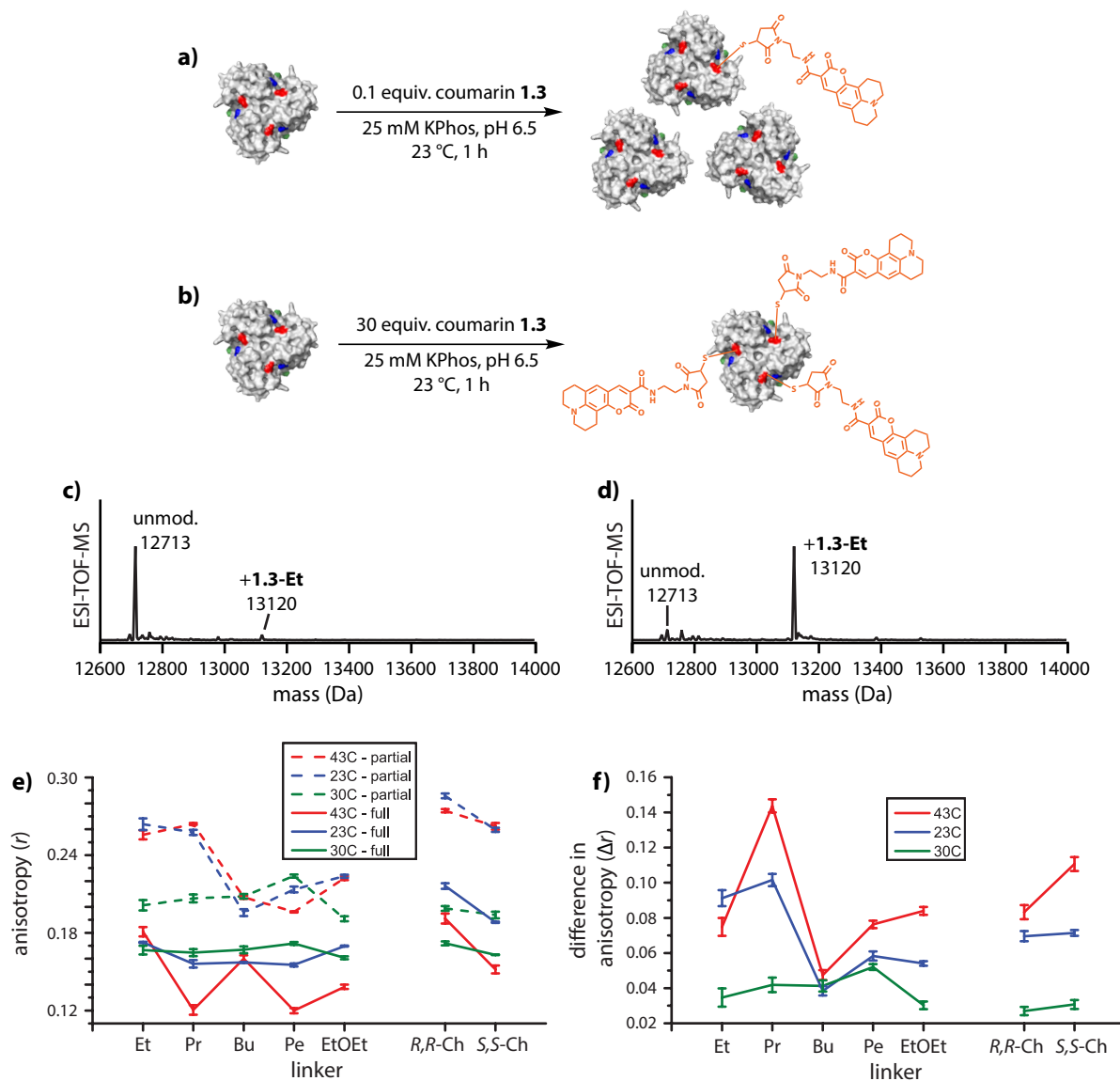
in addition to tumbling or rotation of the fluorophore.<sup>32</sup> Fluorescence anisotropy values for chromophores in solution range from  $-0.2$  to  $0.4$ , where  $0.4$  corresponds to a dilute vitrified solution of a fluorophore with parallel transition moments. In addition to molecular tumbling or rotation, FRET between fluorophores will decrease the measured anisotropy of a sample.<sup>33</sup> Thus, fluorescence anisotropy provides both a method for determining the relative mobility of an isolated fluorophore on a protein surface and the degree of energy transfer between fluorophores within a pigment-protein complex containing multiple pigments.

Relative to natural light-harvesting systems, a particular advantage of synthetic pigment-protein complexes is the ability to conduct measurements at both high and low degrees of substitution. This allows both parameters to be evaluated independently. As the three cysteines on the Mth1491 trimer are equivalent, a chemically controlled method of producing a singly-labeled trimer was not feasible.<sup>34</sup> However, by using 0.1 equivalents of coumarin maleimide **1.3** relative to Mth1491 monomer, a statistical mixture of unmodified, singly-, doubly-, and triply-modified trimer was produced (*Figure 1.4*). Mass spectrometry revealed that samples prepared in this manner were modified to 5–10%, which corresponded to over 90% of trimers bearing  $\leq 1$  dye with a 10:1 to 20:1 ratio of singly-modified to doubly-modified trimers, assuming a binomial distribution. As the majority of the population of proteins was either unmodified or singly labeled, these samples were used to determine the relative mobility of the coumarin dyes attached with the series of linkers to three Mth1491 mutants.

The results of the chromophore mobility experiments are shown in *Figure 1.4e* (dashed lines). Compared to free coumarin **1.3** in solution,  $r = 0.0154(2)$ , all partially-modified Mth1491-coumarin samples displayed significantly higher anisotropy values. Thus, while the free chromophore tumbles at a rate comparable to or faster than its rate of emission, the roughly 38 kDa coumarin-labeled protein trimer tumbles at a much slower rate. As the small chromophore label on the protein should not significantly perturb the hydrodynamic properties of the complex, the differences in fluorescence anisotropy observed between the different partially-labeled samples can be ascribed to differential fluorophore mobility on the protein surface.

As the groove at the 30C position is shallow, linker composition was not expected to have a substantial effect on chromophore mobility. The anisotropy data showed similar mobility between the ethyl, propyl, and butyl linkers and, surprisingly, slightly increased rigidity for the pentyl linker. The increased mobility of the five-atom ethereal linker relative to the pentyl linker suggests the presence of stabilizing interactions between the alkyl linker and hydrophobic residues on the protein surface. Without any significant topographic features in close proximity, the rigid enantiomeric cyclohexyl linkers exhibited comparable levels of chromophore mobility to the linear series of linkers. These results suggest that all linkers, with the exception of the pentyl linker, allowed for free rotation of the chromophore on the protein surface.

The cysteines in both the 43C and 23C mutants are located deeper in the groove where rotational space available to a covalently-bound chromophore should be significantly reduced relative to that of the 30C mutant. Indeed, the ethyl- and propyl-linked coumarins showed much higher emission anisotropy at these positions than at the 30C position. When the pigment-protein distance is increased to the four-atom butyl linker, the anisotropy decreased to a level comparable to the 30C samples. This suggests that with the butyl linker, the coumarin is able to gain mobility by rising above the walls of the protein groove. The rigidity of the 23C pentyl species and the ethereal linked samples for the 43C and 23C mutants is slightly increased, suggesting stabilizing hydrophobic interactions between linker and specific residues. At both the 43C and 23C positions, the two



**Figure 1.4.** Effect of pigment-protein linker on pigment mobility and homo-FRET efficiency. a) Partially- and b) fully-labeled **1.3**-Mth1491 mutants were prepared through cysteine-maleimide coupling. Representative deconvoluted mass spectra of **1.3-Et**-Mth1491-43C show b) <10% modification for partially-labeled trimers and c) >95% modification for fully-modified trimers. d) Fluorescence anisotropy ( $r$ ) of partially- (dashed lines) and fully-labeled (solid lines) samples. Higher values of  $r$  correspond to lower pigment mobilities in partially-labelled samples. f) The difference in anisotropy ( $\Delta r$ ) between partially- and fully-modified samples correlates to homo-FRET efficiency. Higher values of  $\Delta r$  correspond to greater homo-FRET efficiencies. Error bars represent SD of  $\geq 9$  measurements.

enantiomeric cyclohexyl linkers demonstrated rigidity similar to, if not slightly higher than, the ethyl and propyl linkers. In both constructs, the **1.3-*R,R*-Ch** proved to be slightly more rigid than its enantiomer. These results confirmed that through fine linker adjustments; the rotational space accessible to a chromophore, and in turn its average orientation, can be tuned.



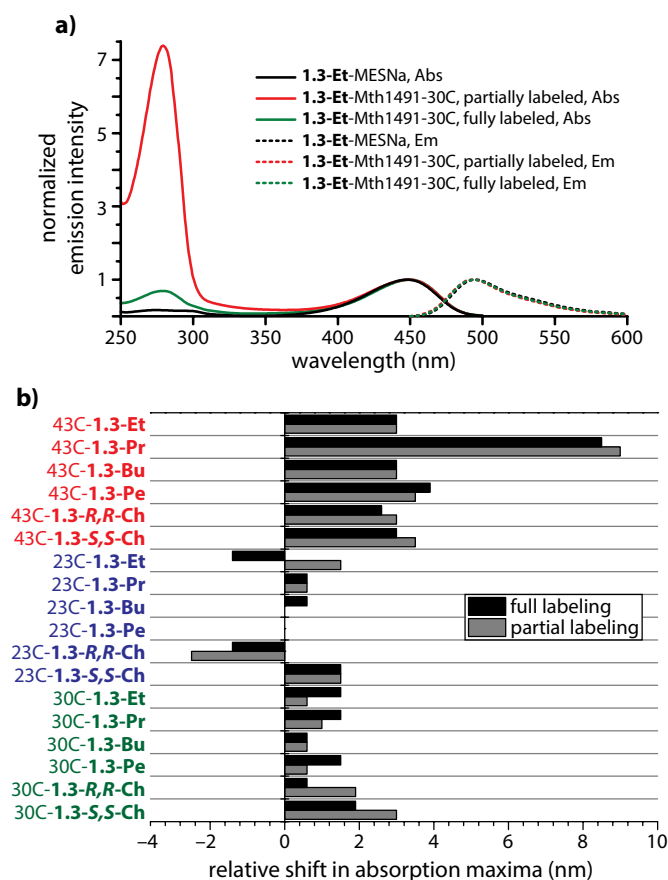
## 1.6 Effects of pigment mobility on pigment-pigment interactions

To examine the effects of the orientational constraints imposed by the combination of linker identity and local scaffold topography on homo-FRET between identical chromophores, fully-labeled coumarin 1.3-Mth1491 conjugates were prepared using the three mutants and the full linker series. Full labeling was achieved using a three-fold excess of coumarin 343 maleimide relative to protein monomer (*Figure 1.4b*). As mentioned above, fluorescence anisotropy can provide a measure of energy transfer due to the orientational differences between donor and acceptor dyes. In the case of systems containing only a single type of chromophore capable of transferring energy in the FRET regime, fluorescence anisotropy, either steady-state or time-resolved, is the only method capable of measuring this energy transfer efficiency.<sup>36</sup> As FRET efficiency is strongly influenced by inter-chromophore distance and orientation, the specific orientation or orientations imposed on the chromophore by the linker and interactions with the surrounding protein surface should give rise to differential changes in anisotropy relative to the baseline provided by the partially-modified samples. In these fully-labeled systems, lower anisotropy values correspond to increased homo-FRET efficiency. However, since the baseline anisotropy of the partially-modified samples is not constant between mutants or linkers, consideration of the absolute difference between the two anisotropy values can more clearly reveal trends.

Examination of the emission anisotropy of the fully-modified samples in *Figure 1.4e* (solid lines) revealed some interesting trends. For the full series of linear linkers, both the 30C and 23C mutants display similar anisotropy values of  $r = 0.16$ – $0.18$ . However, when considering the difference between the observed anisotropy for partially-modified samples and that for fully-modified samples, clear differences emerge (*Figure 1.4f*). Whereas all of the 30C samples, including those with cyclohexyl linkers, shared a fairly constant degree of energy transfer, the ethyl and propyl linkers on the 23C scaffold showed a marked increase in energy transfer. Although the enantiomeric cyclohexyl linkers at the 23C position imparted similar mobility constraints to those of the ethyl and propyl linkers, the ability of the chromophores to interact was significantly lower in comparison. In contrast to the fairly clear pattern of short linear linkers at the 23C position leading to high degrees of energy transfer and longer linear linkers leading to low degrees of energy transfer, the situation at the 43C position is more complicated. At the 43C position, coumarins linked through ethyl, pentyl, ethoxyethyl, and the *R,R*-cyclohexyl linkers all demonstrated similar degrees of energy transfer ability. Notably, the propyl linker on the 43C mutant presented with the highest degree of FRET efficiency despite having nearly identical rigidity as the ethyl linker at the same position. The *S,S*-cyclohexyl linker at this position presented significantly increased homo-FRET ability versus its enantiomeric linker, supporting the hypothesis that the two diastereomeric pigment-proteins would exhibit different pigment-protein and pigment-pigment interactions. In concert, these results suggest that the linkers not only limit the accessible rotational space of the chromophore, but are also able to maintain the chromophores in different orientations relative to one another.

## 1.7 Spectral properties of coumarin 343-Mth1491 conjugates

As demonstrated in Section 1.5, chromophore mobility in pigment-protein complexes can be adjusted through the use of different pigment-protein linking groups. In turn, this altered mobility, in addition to protein-specific factors, results in differential energy transfer efficiencies (Section 1.6). Natural light-harvesting antenna systems, such as peripheral light-harvesting complex 2



**Figure 1.5.** Spectroscopic characteristics of coumarin 1.3-Mth1491 conjugates. a) Representative absorption and emission spectra comparing coumarin 1.3-Et treated with sodium 2-mercaptoethanesulfonate (MESNa) to 1.3-Et-Mth1491-30C, partially or fully labeled with coumarin 343-ethyl-maleimide, normalized to the coumarin signal at 448 nm. All conjugates exhibited similar absorption and emission peak shapes as well as consistent emission maxima. b) Shifts in  $\lambda_{\max}$  relative to that of the free dye (448 nm) were observed between constructs.

the PET pathway. Comparison of fluorescence spectra of the pigment-protein conjugates with the MESNa conjugate of 1.3-Et revealed essentially no change. The peak shape of the emission spectra remained constant, and emission maxima varied by less than 1 nm.

UV-Vis absorption spectra, conversely, showed significant variation between samples. Peak shapes were not found to vary significantly between samples. However, absorption maxima ranged from 445.5 nm (2.5 nm blue-shifted from the free dye) to 457 nm (9 nm red-shifted from the free dye). Shifts in absorption maxima did not correlate with chromophore mobility, as determined by fluorescence anisotropy. Between fully-labeled and partially-labeled trimers, shifts in absorption maxima were generally small (<1 nm) and varied in directionality. While no clear trends could be discerned, this data set does speak to the ability of small adjustments to the pigment-protein linker to tune the photophysical properties of pigment-protein complexes.

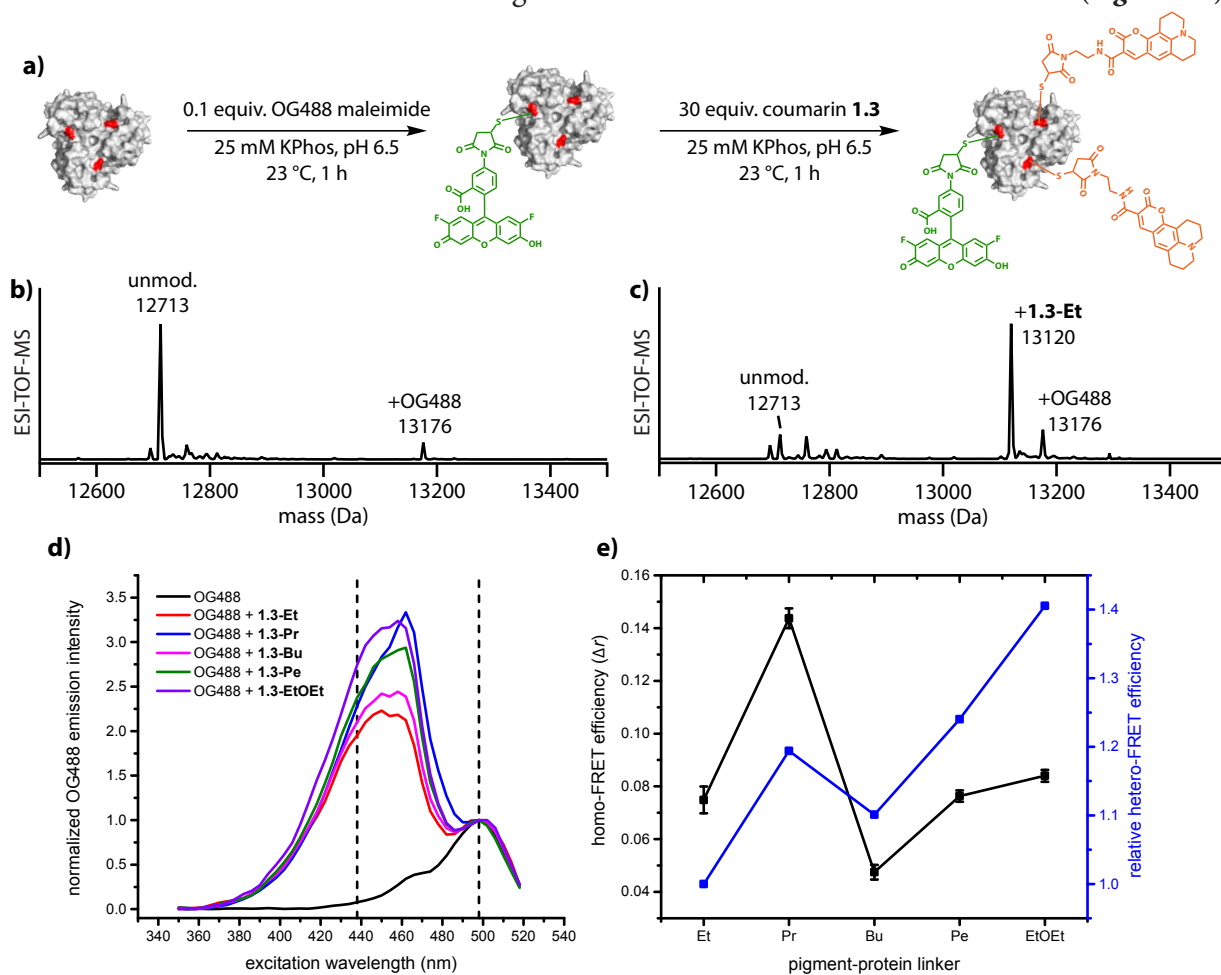
(LH2) found in purple photosynthetic bacteria, have evolved to harness pigment immobilization to adjust energy transfer but also to alter the spectral properties of their pigments.<sup>35</sup> A combination of pigment orientation, inter-pigment spacing, and specific pigment interactions with nearby residues, leads to the formation of distinct groupings of identical pigments that, together, exhibit altered spectral properties (e.g., the B850 and B800 rings of bacteriochlorophyll a pigments in LH2).<sup>37</sup> Due to the variable degree of chromophore mobility and orientation afforded by the linking groups used in the present study, spectral changes between partially- and fully-modified Mth1491 trimers might be expected.

The full collection of Mth1491 mutants labeled, partially or to completion, with the full series of coumarin 343 maleimides was characterized by ultraviolet-visible (UV-Vis) spectroscopy and steady-state fluorometry (Figure 1.5). These spectra were then compared to free 1.3-Et after reaction of the maleimide with the sodium salt of 2-mercaptoethanesulfonate (MESNa). Reaction with MESNa was necessary for comparison of fluorescence data as maleimides are known to quench coumarin 343 emission through a photoinduced electron transfer (PET) pathway.<sup>38</sup> Formation of a thiosuccinimide from the maleimide and a thiol eliminates

## 1.8 Two-color Mth1491 conjugates

To further probe the effects of chromophore mobility on energy transfer, as controlled by pigment-protein linker composition, a dual-color donor-acceptor system was prepared. Through the preparation of pigment-protein conjugates with two distinct chromophore types as donor and acceptor, the effects of chromophore mobility can be extended from a homo-FRET system to a hetero-FRET system and, consequentially, be observed using a method other than fluorescence anisotropy. For this study, the Mth1491-43C mutant was examined as its conjugates had been observed to display the largest variation in chromophore mobility and energy transfer efficiency. The donor chromophores in this study were the linear series of coumarin 343 maleimides and Oregon Green 488 maleimide served as acceptor.

Dual-color conjugates were prepared by first partially labeling Mth1491-43C with Oregon Green 488 maleimide and then backfilling with a member of the coumarin 1.3 series (**Figure 1.6**).



**Figure 1.6.** Linker effects in two-color Mth1491-43C conjugates. a) Two-color systems were prepared by partial labeling with Oregon Green 488 maleimide (OG488) followed by complete labeling with coumarin maleimides 1.3. Deconvoluted mass spectra showing b) partially-modified OG488 maleimide-Mth1491-43C and c) a representative two-color conjugate (OG488-1.3-Et-Mth1491-43C). d) Excitation spectra of two-color Mth1491-43C conjugates and partially-labeled OG488-Mth1491-43C (black trace). Relative hetero-FRET efficiencies were determined ratiometrically by comparing OG488 emission due to donor excitation (438 nm) and due to direct OG488 excitation (498 nm). e) Effect of pigment-protein linker on hetero-FRET and homo-FRET between dyes on the Mth1491-43C trimer.

The result of this conjugation procedure is a batch of Mth149-43C conjugates consisting primarily of trimers with three donors, and trimers with two donors and one acceptor. Hetero-FRET efficiency was determined by comparing the relative contributions of donor and acceptor excitation to acceptor emission.

It was found that hetero-FRET efficiency generally increased with increasing linker length. Additionally, the five-atom ethereal linker demonstrated a marked increase in energy transfer efficiency over the pentyl linker. Much like in the all-coumarin 343 system, the construct with the propyl-linked donor significantly increased FRET efficiency compared to both the ethyl- and butyl-linked donors. The all-coumarin and dual-color systems exhibited very similar trends in energy transfer efficiencies for specific subsets of linkers. However, due to the differences in optimal alignment for energy transfer between a coumarin triad and a coumarin-coumarin-fluorescein triad, the global trend for each system differed. This finding again supports the hypothesis that pigment-protein linker composition can effect pigment-pigment interactions. The predictability of the effect of a specific linker on a specific donor-acceptor system, however, is perhaps quite low.

## **1.9 Towards homogeneous pigment-Mth1491 conjugates**

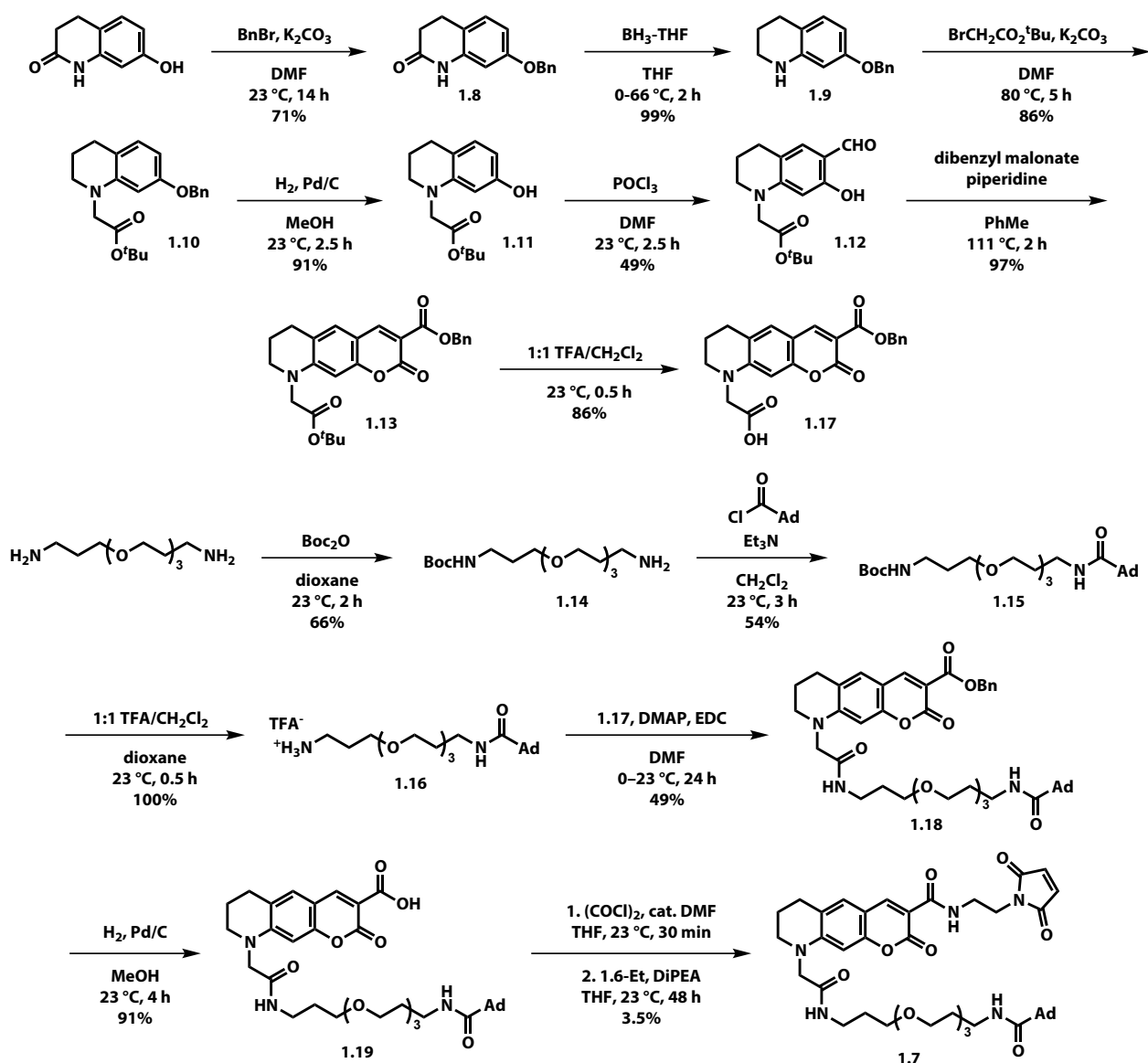
While the coumarin 343-Mth1491 conjugates discussed above clearly demonstrate the general importance of pigment-protein linker composition on pigment-protein and pigment-pigment interactions, the statistical labeling method used to create these materials introduces some uncertainty into the analysis of specific effects on the photophysical properties of these systems. To validate the findings of the anisotropy experiments in Sections 1.5 and 1.6 and to gain more insight specific pigment-pigment interactions leading to the observed spectral characteristics in Section 1.7, a method for the preparation of more well-defined pigment-protein conjugates was desired. More specifically, a method to prepare homogeneous samples of trimers with exactly one, two, or three pigments was needed. Due to the symmetry of the Mth1491 mutants, chemical strategies cannot be used to obtain conjugates with distinct levels of modification. Moreover, non-denaturing chromatographic techniques are generally unable to resolve modified and unmodified protein, let alone conjugates of varying degrees of modification, due to the small differences in size and polarity brought about by conjugation of a chromophore or other small molecule. Recently, we published a general method for preparing conjugates of Mth1491 mutants with controlled numbers of modifications through a new chromatography-mediated bioconjugation strategy, accomplished by affinity chromatography with  $\beta$ -cyclodextrin immobilized on a resin.<sup>34</sup> The adaptation of this general method to coumarin-Mth1491 conjugates is discussed in the following sections.

### **1.9.1 Synthesis of bifunctional coumarin reagent for chromatography-mediated bioconjugation**

In order to use the chromatography-mediated bioconjugation strategy for preparing homogeneous samples of pigment-protein conjugates, a cysteine-reactive chromophore with a pendant moiety capable of forming host-guest complexes with immobilized  $\beta$ -cyclodextrin was needed. To match the coumarin 1.3 series used above most closely, a 7-amino-3-carboxycoumarin derivative was envisioned. For the cysteine-reactive handle, an ethyl-linked maleimide would be installed through amide formation from the carboxylic acid. While the azo dye used in the published ac-

count of chromatography-mediated bioconjugation was found to be a highly effective affinity handle, this moiety would significantly alter the photophysical properties of the coumarin dye of interest. In fact, the majority of known  $\beta$ -cyclodextrin guests are aromatic molecules that could affect the properties of the desired light-harvesting model system.<sup>39</sup> Adamantane, known to bind strongly to  $\beta$ -cyclodextrin (1-adamantane carboxylic acid has an affinity constant of  $10^{4.5} \text{ M}^{-1}$ )<sup>39</sup> and being a relatively innocuous functionality, was chosen as the affinity handle for this bifunctional coumarin. To ensure that the adamantane group would be able to interact with the cyclodextrin resin from within the protein groove, a short discrete polyethylene glycol (dPEG) chain was placed between the adamantyl functionality and the coumarin.

Mertens *et al.* have reported the preparation of a bifunctionalized coumarin for use as an activity-based probe for cysteine cathepsins.<sup>40</sup> Inspired by this work, a synthetic route for the preparation of bifunctional coumarin (**1.7**) was envisioned (*Figure 1.7*). The phenol in commer-



*Figure 1.7.* Synthesis of bifunctional coumarin **1.7**.

cially available 7-hydroxy dihydroquinolinone **1.8** was protected with benzyl bromide in good yield followed by selective reduction of the lactam with borane-tetrahydrofuran complex to afford tetrahydroquinoline **1.9** in near quantitative yield. Alkylation of tetrahydroquinoline **1.9** with *tert*-butyl bromoacetate afforded *N*-alkylated tetrahydroquinoline **1.10** in very good yield. Selective deprotection of benzyl ether **1.10** was accomplished via catalytic hydrogenation with 10% palladium on carbon under 1 atm of H<sub>2</sub> in MeOH, affording 7-hydroxy tetrahydroquinoline **1.11** in excellent yield. Formylation of 7-hydroxy tetrahydroquinoline **1.11** at the 6-position was achieved under Vilsmeier-Haack conditions, affording salicylaldehyde derivative **1.12** in moderate yield. Knoevenagel condensation with dibenzyl malonate provided coumarin **1.13** in near quantitative yield, completing the synthesis of the core of the desired bifunctional coumarin.

The pendant adamantyl group of bifunctional coumarin **1.7** was prepared in three steps from commercially available bis(3-aminopropyl)diethylene glycol. Mono-Boc protection of the diethylene glycol derivative was achieved under standard conditions, affording mono-Boc diamine **1.14** in moderate yield. Interestingly, it was found that bis(3-aminopropyl)diethylene glycol and its derivatives displayed strongly concentration dependent retention factors during silica gel chromatography, significantly limiting the amount of pure material that could be isolated from a single round of purification. Mono-Boc diamine **1.14** was acylated with the acid chloride of 1-adamantanecarboxylic acid to afford adamantane derivative **1.15** in moderate yield. Deprotection of Boc-amine **1.15** with trifluoroacetic acid in methylene chloride afforded amino-dPEG<sub>3</sub>-adamantane as the trifluoroacetate salt **1.16** in near quantitative yield.

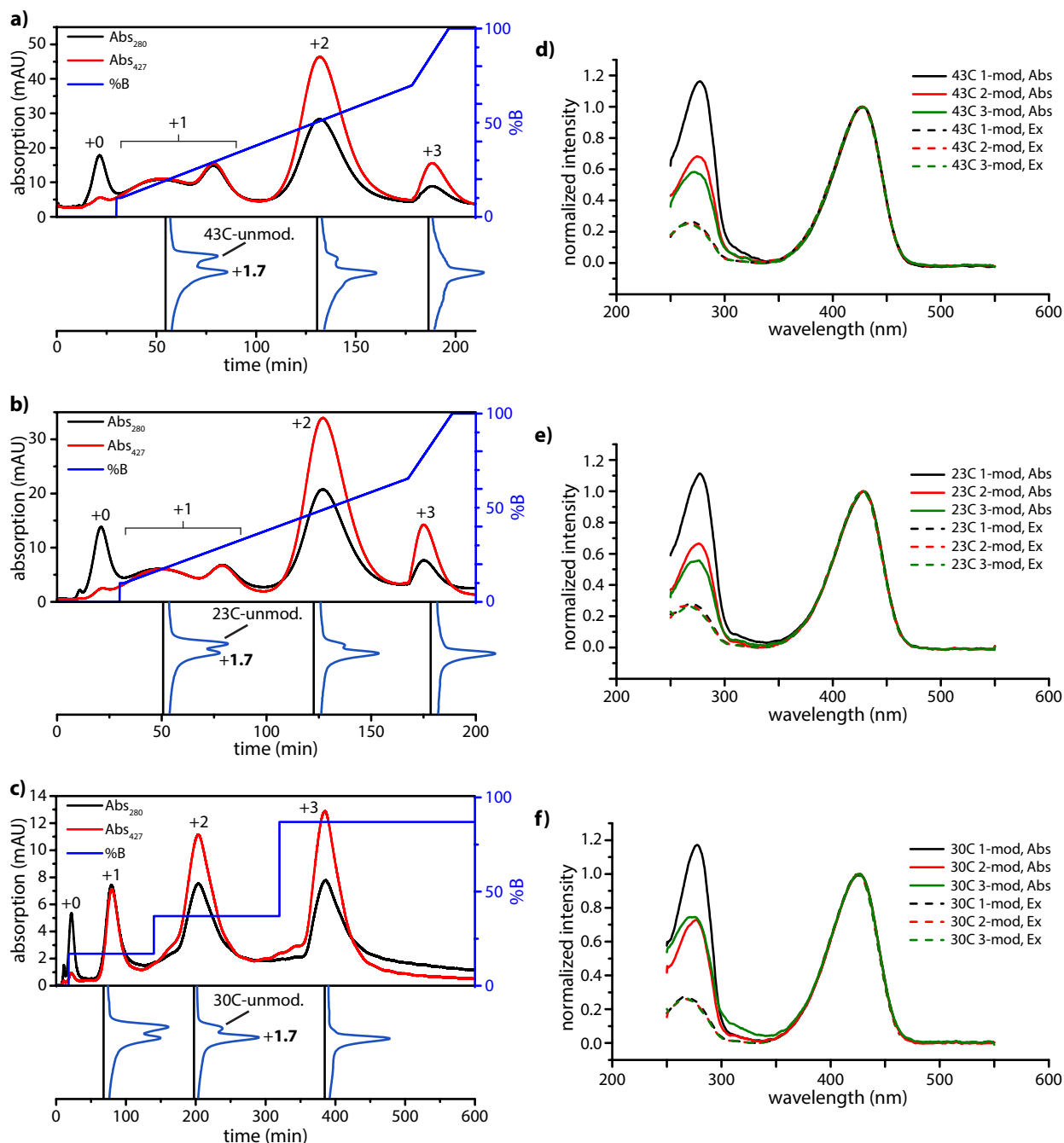
The *tert*-butyl ester in coumarin derivative **1.13** was selectively deprotected with 1:2 trifluoroacetic acid in methylene chloride, affording acid **1.17** in very good yield. Conjugation of amino-dPEG<sub>3</sub>-adamantane **1.16** to coumarin carboxylic acid **1.17** was achieved with EDCI and DMAP in good yield. Following attachment of the pendant adamantyl moiety, the benzyl ester in **1.18** was deprotected by catalytic hydrogenation with 10% palladium on carbon under 1 atm of H<sub>2</sub> in MeOH. The resultant carboxylic acid **1.19** was converted into the corresponding acid chloride with oxalyl chloride and treated with maleimide **1.6-Et** and triethylamine to afford the desired bifunctional coumarin derivative **1.7** in moderate yield. Overall, bifunctional coumarin **1.7** was prepared using a convergent strategy with a longest linear synthetic sequence of 10 steps and a total yield of 0.5% after HPLC purification.

## 1.9.2 Preparation and characterization of pigment-protein complexes using chromatography-mediated bioconjugation

Towards the study of homogenous pigment-protein conjugates, a chromatography-mediated bioconjugation strategy was utilized. The three mono-Cys Mth1491 mutants were modified to a level of roughly 50% using bifunctional coumarin **1.7**. The resulting mixtures of conjugates were then subjected to affinity purification on an immobilized  $\beta$ -cyclodextrin resin. Conjugates were eluted from the resin by competition with a gradient mobile phase containing up to 10 mM  $\beta$ -cyclodextrin (*Figure 1.8*).

Conjugates prepared from Mth1491-30C were chromatographically well-behaved and bound strongly to the resin. Singly-, doubly-, and triply-modified Mth149-30C could be isolated using an optimized step gradient elution. Perhaps as a result of the protein groove concealing the affinity handle from fully interacting with the resin, resolution of conjugates of the 23C and 43C mutants was found to be more difficult. Separation of singly-modified protein from unmodified pro-

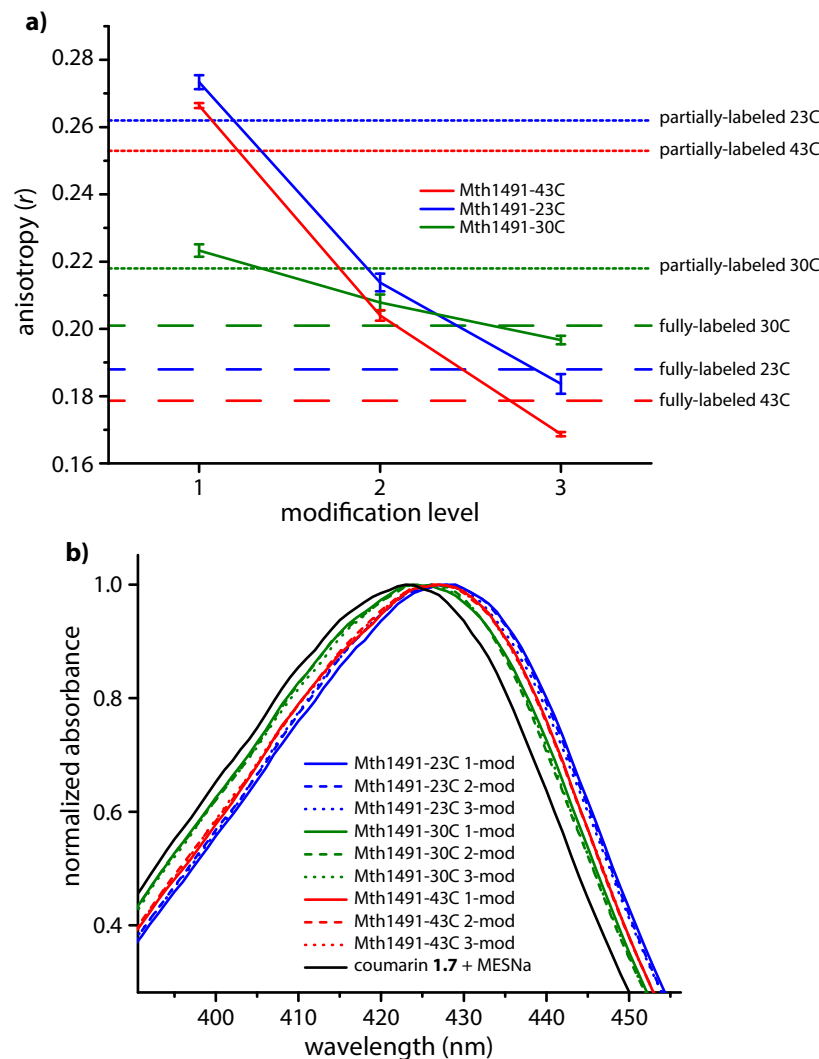
tein was challenging, as the singly-modified material eluted without the addition of  $\beta$ -cyclodextrin-containing buffer. However, conjugates with either one, two, or three bifunctional coumarins could be isolated from one another using a simple linear gradient elution.



**Figure 1.8.** Isolation and confirmation of homogeneous coumarin 1.7-Mth1491 conjugates. Affinity chromatography with Sepharose-immobilized  $\beta$ -cyclodextrin of Mth1491 mutants, which were partially labeled with coumarin 1.7, afforded homogeneous samples of singly-, doubly-, and triply-labeled a) Mth1491-43C, b) Mth1491-23C, and c) Mth1491-30C. Conjugates were eluted with a gradient of B (10 mM  $\beta$ -cyclodextrin in 20 mM Phos, pH 6.5). Below each chromatogram appear total ion count traces from ESI-TOF-MS analysis of pooled samples. d-f) Single, double, and triple modification was confirmed by comparing UV-Vis absorption and fluorescence excitation spectra, normalized to the primary coumarin peak at 424 nm. Excitation spectra were monitored by emission at 473 nm.

Isolated homogenous conjugates were first characterized by ESI-TOF-MS. Assuming the isolated samples were indeed singly-, doubly-, and triply-modified, Mth1491 conjugates ionized better than unmodified protein, which resulted in unrealistic ratios of modified to unmodified protein as measured by mass spectrometry. UV-Vis spectroscopy of these samples revealed the expected ratio of pigment ( $A_{425}$ ) to protein ( $A_{280}$ ) after adjusting for the contribution of pigment at  $A_{280}$ , confirming successful isolation of homogeneously-modified pigment-protein complexes.

Analysis by fluorescence anisotropy revealed similar trends to those found for Mth1491 statistically labeled with the coumarin 1.3 series (*Figure 1.9a*). Bifunctional coumarin 1.7 exhibited significantly higher anisotropy in singly-labeled Mth1491 in the 43C and 23C mutants than the 30C



**Figure 1.9.** Characterization of homogeneous coumarin 1.7-Mth1491 conjugates. a) Fluorescence anisotropy of partially- (dotted lines) and fully-labeled (dashed lines) 1.7-Mth1491 conjugates, prepared by statistical labeling, demonstrated increased pigment mobility and decreased homo-FRET efficiency, respectively, compared to resolved homogeneous 1.7-Mth1491 conjugates (solid lines). Error bars represent SD of  $\geq 9$  measurements. b) Partial UV-Vis absorption spectra of resolved 1.7-Mth1491 conjugates. Absorption maxima of the conjugates were redshifted relative to free coumarin 1.7-MESNa.

mutant. This was consistent with findings for the coumarin 343 maleimide with identical linker composition (1.3-Et). Anisotropy was found to decrease, as expected, with each added pigment to the trimer. Going from the singly- to doubly-labeled system, a large decrease in anisotropy was observed: 23.4%, 20.0%, and 6.9% for 43C, 23C, and 30C, respectively. Addition of the final pigment resulted in a smaller decrease in anisotropy: 17.3%, 18.4%, and 5.4% for 43C, 23C, and 30C, respectively. Overall, the energy transfer was found to be most efficient in conjugates of the 43C and 23C mutants, again in agreement with the statistically-labeled conjugates of coumarin 343 maleimide.

To verify the validity of using partially-labeled Mth1491 conjugates as a means to access singly-modified trimers, bifunctional coumarin 1.7-Mth1491 conjugates were prepared by statistical labeling, and were compared to the homogeneous conjugates. Partially-modified Mth1491-coumarin 1.7 conjugates were found to exhibit fluorescence anisotropy values roughly 0.01 units lower than that of the pure singly-labeled



trimer. This is consistent with the small percentage doubly- and triply-labeled trimer expected from statistical labeling. Fully-modified Mth1491-coumarin **1.7** conjugates gave anisotropy values roughly 0.005–0.01 units higher than that of pure triply-labeled trimer. Again, this is consistent with non-quantitative labeling of the trimer. Although statistically-labeled samples underestimated chromophore mobility and energy transfer efficiency, the observed trends are fully consistent with those of homogeneous samples. This finding validates the conclusions drawn from the coumarin 343-Mth1491 study in previous sections.

Unlike the series of coumarin 343-Mth1491 conjugates, the spectral properties of homogeneous coumarin **1.7**-Mth1491 conjugates were generally in line with those of the free dye. Emission spectra were consistent with free coumarin **1.7** conjugated to MESNa. Emission maxima varied by less than 2 nm and those for singly-, doubly-, and triply-labeled trimers of a given mutant clustered within 0.5 nm. A slight bathochromic shift in the absorption spectra of conjugates was observed (**Figure 1.9b**). Again, spectra were clustered by mutant, but shifts of roughly 2, 4, and 6 nm were observed for the 30C, 43C, and 23C mutants, respectively. Unfortunately, these well-defined conjugates of coumarin **1.7** did not provide much insight into the irregular absorption profiles of the series coumarin 343 conjugates.

## 1.10 Effects of linker composition in tobacco mosaic virus-templated pigment-protein complexes

Towards extending the study of pigment-protein linker composition to systems more relevant for applications in artificial light-harvesting, a selection of the coumarin 343-maleimide constructs were used to prepare pigment-protein conjugates from TMVcp. Double disks of RR-TMVcp were treated with coumarin 343-maleimides to prepare conjugates labeled at Cys123. Partially-modified and fully-modified double disks were prepared using coumarin **1.3-Et**, **1.3-EtOEt**, **1.3-R,R-Ch**, and **1.3-R,R-Ch**.

**Table 1.1.** Fluorescence anisotropy values of RR-TMVcp double disks modified partially or fully at Cys123 with coumarin **1.3**.

	$r_{\text{partial}}$	$r_{\text{full}}$	$\Delta r$
<b>1.3-Et</b>	0.30	0.15	0.15
<b>1.3-R,R-Ch</b>	0.29	0.21	0.08
<b>1.3-S,S-Ch</b>	0.32	0.18	0.14
<b>1.3-EtOEt</b>	0.29	–	–

linker. However, the ethereal linker also resulted in a comparable anisotropy value to that of the ethyl linker. Together, these findings would suggest that the local topography around Cys123 of RR-TMVcp is relatively featureless, much like Mth1491-A30C. The crystal structure of the circular permutant of TMVcp (PDB ID: 3KML) supports this hypothesis.<sup>17</sup>

While the observed mobility of coumarin 343 maleimides at Cys123 of RR-TMVcp double disks does not appear to be strongly influenced by linker composition, homo-FRET in these systems is effected. Upon complete modification of RR-TMVcp with coumarin **1.3-Et**, the anisotropy decreased by 0.15 units. This change in anisotropy is comparable to the most effective Mth1491-based system (Mth1491-43C with coumarin **1.3-Pr**). The *R,R*-cyclohexyl-linked RR-TMVcp conjugate

Fluorescence anisotropy measurements were performed for the collection of partially- and fully-modified TMV double disks (**Table 1.1**). Compared to the Mth1491 conjugates, the TMV conjugates demonstrated higher anisotropy values for singly-modified systems. This relative increase in anisotropy can be attributed to the greater-than-15-fold mass increase from Mth1491 to double disk RR-TMVcp, resulting in a decreased rate of tumbling. As in the Mth1491 system, the cyclohexyl linkers provided similar levels of rigidity as did the ethyl

was observed to transfer energy about as well as the ethyl-linked system. However, the diastereomeric system composed of *S,S*-cyclohexyl-linked coumarin 343 showed a markedly lower homo-FRET efficiency of 0.08 anisotropy units. While the *S,S*-cyclohexyl-linked system was demonstrated to be less effective at transferring energy within the ring of chromophores, most likely due to an unfavorable orientational arrangement of pigments, advantages of such a system could be envisioned for use in a larger array of synthetic pigment-protein complexes. By disfavoring energy transfer within a single complex, improved energy transfer between complexes may be achieved. If this is the case, artificial arrays much like those found in natural photosynthetic light-harvesting systems could be constructed using a combination of pigment-protein conjugates optimized for intra- and inter-complex energy transfer.

## 1.11 Conclusions

These studies demonstrate the importance of pigment-protein linker composition as a functional design element in artificial light-harvesting systems. The interplay of protein-bound chromophores with the protein surface is, in part, dictated by the length, rigidity, and hydrophobicity of the linking group. The resulting restrictions on chromophore mobility and orientation ultimately result in altered pigment-pigment interactions and energy transfer efficiencies. In ongoing studies, we are expanding the collection of linkers and exploring alternative bioconjugation strategies to create a modular toolkit for understanding natural antenna complexes and building next-generation synthetic light-harvesting systems.

## 1.12 Materials and Methods

### 1.12.1 General procedures and materials

Unless otherwise noted, all chemicals were purchased from commercial sources and used as received without further purification. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F<sub>254</sub> plates and visualized by ultraviolet (UV) irradiation at 254 nm and staining with potassium permanganate. Purifications by flash silica gel chromatography were performed using EM silica gel 60 (230–400 mesh). Chromatography solvents were used without distillation. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H<sub>2</sub>O) used in all non-synthetic procedures was deionized using a NANOpure™ purification system (Barnstead, USA). Centrifugations were performed with an Eppendorf 5424 R (Eppendorf, Hauppauge, NY). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA).

### 1.12.2 Instrumentation and sample analysis

**NMR.** <sup>1</sup>H and <sup>13</sup>C spectra were measured with a Bruker AVQ-400, AVB-400 (400 MHz), or AV-300 (300 MHz) spectrometer. <sup>1</sup>H NMR chemical shifts are reported as  $\delta$  in units of parts per million (ppm) relative to residual CHCl<sub>3</sub> ( $\delta$  7.26, singlet), CH<sub>3</sub>OH ( $\delta$  3.31, quintet), or DMSO ( $\delta$  2.50, quintet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (quintet), or br s (broad singlet). Coupling constants are reported as a *J* value in Hertz (Hz). The number of

protons (n) for a given resonance is indicated as nH and is based on spectral integration values.  $^{13}\text{C}$  NMR chemical shifts are reported as  $\delta$  in units of parts per million (ppm) relative to  $\text{CDCl}_3$  ( $\delta$  77.16, triplet),  $\text{CD}_3\text{OD}$  ( $\delta$  49.15, septet), or  $\text{DMSO-d}_6$  ( $\delta$  39.52, septet).

**Mass Spectrometry.** Low-resolution mass spectrometry (LRMS) of small molecules was performed on an AB Sciex 3200 Qtrap (AB Sciex, USA) equipped with an ESI Turbo V ion source connected in-line with an Agilent 1100 series high-performance liquid chromatograph (HPLC) (Agilent Technologies, USA). Chromatography was performed using a Gemini C18 column (Phenomenex, Torrance, CA) or a Chromolith RP-18 column (Merck, Darmstadt, Germany) with a  $\text{H}_2\text{O}/\text{MeCN}$  gradient mobile phase containing 0.1% formic acid. High-resolution mass spectrometry (HRMS) was performed at the QB3 Mass Spectrometry Facility (Berkeley, CA). Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) was performed on a Voyager-DE system (PerSeptive Biosystems, USA). Small molecule samples were co-crystallized with  $\alpha$ -cyano-4-hydroxycinnamic acid in 1:1  $\text{MeCN}/\text{H}_2\text{O}$  with 0.1% TFA. Proteins and protein conjugates were analyzed on an Agilent 6224 Time-of-Flight (TOF) mass spectrometer with a dual electrospray source connected in-line with an Agilent 1200 series HPLC (Agilent Technologies, USA). Chromatography was performed using a Poroshell 5 $\mu$  C18 300 Å (Agilent Technologies, USA) column with a  $\text{H}_2\text{O}/\text{MeCN}$  gradient mobile phase containing 0.1% formic acid. Mass spectra of proteins and protein conjugates were deconvoluted with the MassHunter Qualitative Analysis Suite B.05 (Agilent Technologies, USA).

**High Performance Liquid Chromatography (HPLC).** HPLC was performed on Agilent 1100 series HPLC systems (Agilent Technologies, USA) equipped with in-line diode array detector (DAD) and fluorescence detector (FLD). Analytical and preparative reverse-phase HPLC of small molecules was accomplished using Gemini C18 columns (Phenomenex, Torrance, CA) and a  $\text{H}_2\text{O}/\text{MeCN}$  gradient mobile phase containing 0.1% TFA. Size exclusion chromatography was accomplished on a Phenomenex PolySep-GFC-P 5000 or BioSep SEC-S-4000 column using an aqueous mobile phase (50 mM Tris-HCl, 100 mM NaCl, pH 7.2).

**Ultraviolet-Visible (UV-Vis) Spectrometry.** UV-Vis spectrometry was performed using 3x3 mm Suprasil quartz cuvettes (Hellma Analytics, Germany) with a Varian Cary 50 spectrophotometer (Varian Inc., Palo Alto, CA). All measurements were collected with a scan rate of 80 nm/s and were baseline corrected. Small-scale UV-Vis spectrometry was performed using a Nanodrop 1000 (Thermo Scientific, USA).

**Spectrofluorometry.** Fluorescence measurements were performed using 3x3 mm Suprasil quartz cuvettes (Hellma Analytics, Germany) with a Fluoromax-4 spectrofluorometer equipped with a Peltier temperature controller and automatic excitation and emission polarizers (HORIBA Jobin Yvon Inc., Edison, NJ). Samples were diluted to an  $\text{OD} \leq 0.03$  and filtered through a 0.22  $\mu\text{m}$  centrifugal filter prior to analysis. All measurements were collected with a 0.1 s integration time. Fluorescence anisotropy experiments were performed using 5.0 nm excitation and 1.0 nm emission slit widths at 25.0 °C. Anisotropy measurements were reported as the mean and standard deviation of a minimum of 9 measurements. The anisotropy of small molecule maleimide-containing fluorophores were measured after incubating with 2 equivalents of MESNa for 10 min.

**Dynamic Light Scattering (DLS).** DLS measurements were obtained using a Malvern Instruments

Zetasizer Nano ZS (Malvern Instruments, United Kingdom). Samples were filtered through 0.22  $\mu\text{m}$  centrifugal filter units (Millipore Corporation, Billerica, MA) prior to data collection.

### 1.12.3 Experimental procedures

**Expression and purification of wt-Mth1491 and Mth1491 mono-Cys mutants.** Mth1491 was expressed and purified according to a modified literature procedure.<sup>26</sup> BL21(Gold $\lambda$ DE3) cells were transformed with a pJexpress404 plasmid containing the ampicillin-resistant gene ampR and the Mth1491 gene with an N-terminal His<sub>6</sub>-tag followed by a thrombin cleavage site (DNA2.0 Inc., Menlo Park, CA). Colonies were selected for inoculation of lysogeny broth cultures with 100  $\mu\text{g}/\text{mL}$  ampicillin. When cultures reached mid-log phase as determined by OD<sub>600</sub>, expression was induced by addition of 10  $\mu\text{M}$  IPTG. Cultures were grown for 14–18 h at 37 °C and cells were isolated by centrifugation. Cells were resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0). Cells were lysed by sonication (Fisher Scientific, USA) and cleared by centrifugation. The His<sub>6</sub>-Mth1491 mutant was purified from the cleared lysate by affinity chromatography on Ni-NTA agarose, following the manufacturer's protocol (Qiagen, USA). The purified protein was buffer-exchanged into 50 mM Tris-HCl, pH 8.0 and then into cleavage buffer (50 mM Tris-HCl, pH 8.0 with 10 mM CaCl<sub>2</sub>). The His<sub>6</sub>-tag was cleaved using Thrombin CleanCleave (Sigma-Aldrich, USA), incubating for 24 h at 4 °C. Mth1491 was isolated from the cleaved His<sub>6</sub>-tag and residual His<sub>6</sub>-Mth1491 by buffer exchange into 50 mM Tris-HCl, pH 8.0 and then into lysis buffer and performing a pull-down with Ni-NTA agarose. Collection of the supernatant yielded 20–40 mg of protein per liter of culture. Isolated protein was exchanged into 25 mM potassium phosphate, pH 6.5 with 0.01% NaN<sub>3</sub> and stored at 4 °C.

**Generation of Mth1491 Mutants.** The QuickChange II Site-Directed Mutagenesis Kit (Agilent Technologies, USA) was used to generate wt-Mth1491, Mth1491-S16C, Mth1491-S23C, and Mth1491-A30C from the Mth1491-S43C plasmid.

Mth1491-43C was converted to the wild type 43S using the primers:

sense: 5' –CGTATCGAGGTGGTTGCGTACAGTATGGGCGTTAATGTGCTGC– 3'  
antisense: 5' –CTGACGACGCAGCACATTAACGCCATACTGTACGCAACCACCTCG– 3'

The serine of wt-Mth1491 at position 16 was converted to cysteine using the primers:

sense: 5' –CGATGAAGATGACGAGTGCCGCGTCCTGTTGC– 3'  
antisense: 5' –CTAATCAGCAACAGGACGCGGCACTCGTCATCTTC– 3'

The serine of wt-Mth1491 at position 23 was converted to cysteine using the primers:

sense: 5' –CCGCGTCCTGTTGCTGATTTGCAACGTCCGTAATC– 3'  
antisense: 5' –CGGACGTTGCAAATCAGCAACAGGACGCGGCTCTCG– 3'

The alanine of wt-Mth1491 at position 30 was converted to a cysteine using the primers:

Forward: 5' –GCAACGTCCGTAATCTGATGTGCGATCTGGAATCGGTACGTATC– 3'

Reverse: 5' –CCTCGATACGTACCGATTCCAGATCGCACATCAGATTACGGACG– 3'

Protocols for expression and purification of all mutants were identical.

**Expression and purification of TMV.** TMV plasmid construction, expression, and purification has been previously reported.<sup>17</sup>

**General procedure for Cys123 alkylation of TMV, full labeling.** An opaque Eppendorf tube was charged with a solution of RR-TMVcp (200  $\mu$ M monomer final concentration, 10 mM phosphate, pH 7.0). To this solution was added a maleimide-containing reagent (5–10 equivalents from a 100 mM DMF stock solution). The resulting mixture was stirred very gently with triangular magnetic stir bars at room temperature for 2 h. The reaction mixture was then purified with an illustra Nap-5 Sephadex G-25 column (GE Healthcare, USA), eluting with 10 mM phosphate, pH 7.0, followed by spin concentration using 100 kDa MWCO filters. Double disks were isolated by SEC purification prior to spectral analysis. Samples were analyzed by ESI-TOF-MS.

**General procedure for Cys123 alkylation of TMV, partial labeling.** An opaque Eppendorf tube was charged with a solution of RR-TMVcp (200  $\mu$ M monomer final concentration, 10 mM phosphate, pH 7.0). To this solution was added a maleimide-containing reagent (0.03 equivalents from a 100 mM DMF stock solution). The resulting mixture was stirred very gently with triangular magnetic stir bars at room temperature for 2 h. The reaction mixture was then purified with an illustra Nap-5 Sephadex G-25 column (GE Healthcare, USA), eluting with 10 mM phosphate, pH 7.0, followed by spin concentration using 100 kDa MWCO filters. Double disks were isolated by SEC purification prior to spectral analysis. Samples were analyzed by ESI-TOF-MS.

**General procedure for cysteine alkylation of Mth1491, full labeling.** An opaque Eppendorf tube was charged with a solution of Mth1491 (100  $\mu$ M monomer final concentration, 25 mM potassium phosphate, pH 6.5). To this solution was added a maleimide-containing reagent (300  $\mu$ M final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified with an illustra Nap-5 Sephadex G-25 column (GE Healthcare, USA), eluting with 25 mM potassium phosphate, pH 6.5.

**General procedure for cysteine alkylation of Mth1491, partial labeling.** An opaque Eppendorf tube was charged with a solution of Mth1491 (100  $\mu$ M monomer final concentration, 25 mM potassium phosphate, pH 6.5). To this solution was added a maleimide-containing reagent (10  $\mu$ M final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified with an illustra Nap-5 Sephadex G-25 column (GE Healthcare, USA), eluting with 25 mM potassium phosphate, pH 6.5.

**General procedure for preparation of dual-color Mth1491 conjugates.** An opaque Eppendorf tube was charged with a solution of Mth1491 (100  $\mu$ M monomer final concentration, 25 mM potassium phosphate, pH 6.5). To this solution was added Oregon Green 488 maleimide (10  $\mu$ M final

concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified with an illustra Nap-5 Sephadex G-25 column (GE Healthcare, USA), eluting with 25 mM potassium phosphate, pH 6.5.

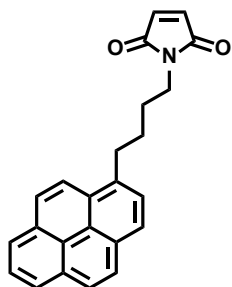
The partially-modified Oregon Green 488-Mth1491 conjugate (100  $\mu$ M monomer final concentration, 25 mM potassium phosphate, pH 6.5) was then treated with coumarin 343 maleimide (300  $\mu$ M final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified with an illustra Nap-5 Sephadex G-25 column (GE Healthcare, USA), eluting with 25 mM potassium phosphate, pH 6.5.

Relative hetero-FRET efficiency was determined using a ratiometric method. Excitation spectra were obtained, monitoring at 526 nm, for the two-dye systems and normalized to the excitation maximum of Oregon Green 488 (498 nm). Relative hetero-FRET efficiency was defined as the ratio of emission intensity with excitation at 438 nm (excitation of donor) to that with excitation at 498 nm (direct excitation of acceptor).

**General procedure for chromatography-mediated bioconjugation.** An opaque Eppendorf tube was charged with a solution of Mth1491 (100  $\mu$ M monomer final concentration, 25 mM potassium phosphate, pH 6.5). To this solution was added maleimide 1.7 (50  $\mu$ M final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then filtered through a 0.22  $\mu$ m centrifugal filter.

The filtered reaction mixture was then resolved into homogeneous fractions of distinct levels of chromophore-labeling using an Akta Pure M (GE Healthcare, USA) equipped with a 25 cm column packed with  $\beta$ -cyclodextrin-terminated Sepharose.<sup>34</sup> Conjugates were eluted using a gradient of 0–10 mM  $\beta$ -cyclodextrin in 20 mM phosphate buffer, pH 6.5, and a flow rate between 0.05 and 0.2 mL/min.

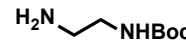
#### 1.12.4 Small molecule synthesis

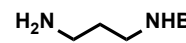


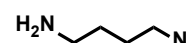
**Synthesis of 1-(4-(pyren-1-yl)butyl)-1H-pyrrole-2,5-dione (1.2).** To a solution of  $\text{PPh}_3$  (64 mg, 0.243 mmol, 1 equiv.) in dry THF (5 mL) was added a solution of diethyl azodicarboxylate in toluene (40 wt%, 0.11 mL, 0.243 mmol, 1 equiv.) at  $-78^\circ\text{C}$  under  $\text{N}_2$ . After 5 min, 1-pyrenebutanol (100 mg, 0.364 mmol, 1.5 equiv.) was added. After 5 min, maleimide (24 mg, 0.243 mmol, 1 equiv.) was added in one portion and the reaction was warmed slowly to room temperature and stirred for 12 h. Volatiles were removed under reduced pressure and the residue was purified via flash silica gel chromatography, eluting with 20% EtOAc/hexanes to afford a yellow solid (36 mg, 42%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.24 (d,  $J = 9.2$  Hz, 1H), 8.20–7.96 (m, 7H), 7.84 (d,  $J = 7.7$  Hz, 1H), 6.59 (s, 2H), 3.58 (t,  $J = 7.0$  Hz, 2H), 3.35 (t,  $J = 7.4$  Hz, 2H), 1.91–1.72 (m, 4H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.90, 136.27, 134.01, 131.48, 130.96, 129.93, 128.65, 127.58, 127.36, 126.70, 125.89, 125.13, 125.07, 124.95, 124.90, 124.79, 123.39, 37.71, 32.97, 28.87, 28.62.


**Typical procedure for the synthesis of mono-N-Boc alkyl diamines (Example of tert-butyl (2-aminoethyl)carbamate (1.4-Et)).** To a solution of ethylenediamine (2.55 mL, 38.0 mmol, 5.2 equiv.) in  $\text{CHCl}_3$  (35 mL) was added a solution of  $\text{Boc}_2\text{O}$  (1.59 g, 7.3 mmol, 1 equiv.) in  $\text{CHCl}_3$

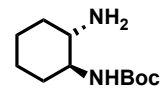
(400 mL) over 10 min at room temperature under N<sub>2</sub>. After 12 h, the reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub>, washed twice with 1 M aqueous NaHCO<sub>3</sub>, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude product was purified via flash silica gel chromatography, eluting with 10% MeOH/CHCl<sub>3</sub> with 0.1% Et<sub>3</sub>N.

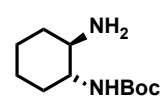
 **tert-butyl (2-aminoethyl)carbamate (1.4-Et)**. Pale yellow oil (98% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.97 (br s, 1H), 3.09 (t, *J* = 6.1 Hz, 2H), 2.70 (t, *J* = 6.1 Hz, 2H), 1.42 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 156.43, 79.28, 43.61, 41.99, 28.58.

 **tert-butyl (3-aminopropyl)carbamate (1.4-Pr)**. Pale yellow oil (97% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.96 (s, 1H), 3.18 (dd, *J* = 12.8, 6.4 Hz, 2H), 2.73 (t, *J* = 6.6 Hz, 2H), 1.58 (p, *J* = 6.6 Hz, 2H), 1.41 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 155.93, 78.17, 39.03, 37.66, 32.94, 28.01.

 **tert-butyl (4-aminobutyl)carbamate (1.4-Bu)**. Pale yellow oil (99% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.69 (s, 1H), 3.12 (m, 2H), 2.73 (t, *J* = 6.4 Hz, 2H), 1.52 (m, 4H), 1.43 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 155.78, 78.23, 41.19, 39.82, 30.29, 27.86, 26.97.

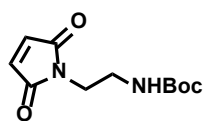
 **tert-butyl (5-aminopentyl)carbamate (1.4-Pe)**. Pale yellow oil (95% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.57 (br s, 1H), 3.10 (d, *J* = 6.2 Hz, 2H), 2.68 (t, *J* = 6.9 Hz, 2H), 1.78 (br s, 2H), 1.51–1.44 (m, 4H), 1.42 (s, 9H), 1.37–1.31 (m, 2H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 156.90, 79.89, 42.81, 41.33, 34.02, 30.68, 29.17, 24.90.

 **tert-butyl ((1*S*,2*S*)-2-aminocyclohexyl)carbamate (1.4-*S,S*-Ch)**. Tan solid (86% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.51 (d, *J* = 8.3 Hz, 1H), 3.12 (t, *J* = 10.9 Hz, 1H), 2.34 (dt, *J* = 11.7, 5.9 Hz, 1H), 2.02–1.91 (m, 2H), 1.81–1.73 (m, 2H), 1.69 (dt, *J* = 12.0, 2.7 Hz, 2H), 1.44 (s, 7H), 1.29–1.07 (m, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 156.12, 57.46, 55.64, 35.02, 32.85, 28.37, 25.13, 25.00.

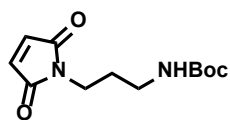
 **tert-butyl ((1*R*,2*R*)-2-aminocyclohexyl)carbamate (1.4-*R,R*-Ch)**. Tan solid (90% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.51 (d, *J* = 8.3 Hz, 1H), 3.12 (t, *J* = 10.9 Hz, 1H), 2.34 (dt, *J* = 11.7, 5.9 Hz, 2H), 2.02–1.91 (m, 2H), 1.81–1.73 (m, 2H), 1.69 (dt, *J* = 12.0, 2.7 Hz, 2H), 1.44 (s, 9H), 1.29–1.07 (m, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 156.13, 57.46, 55.64, 35.06, 32.85, 28.37, 25.13, 25.01.

**Typical procedure for the synthesis of *N*-Boc-aminoalkylmaleimides (Example of *tert*-butyl (2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl)carbamate (1.5-Et)).** *tert*-Butyl (2-aminoethyl) carbamate (3.45 g, 21.5 mmol, 1 equiv.) was added to a solution of maleic anhydride (2.47 g, 25.2 mmol, 1.17 equiv.) in acetone (15 mL) and the reaction mixture was stirred at 5 °C for 1 h. Solvent was removed under reduced pressure and the residue was taken up in warm EtOAc, washed twice with water, twice with 1 M aqueous HCl, once with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure to afford the crude maleamic acid as an off-white solid. The crude maleamic acid was then heated with anhydrous NaOAc (1.64 g, 20.0 mmol, 0.93 equiv.) in Ac<sub>2</sub>O (17 mL) at 100 °C for 1 h under N<sub>2</sub>. After cooling over ice, the reaction mixture was added to

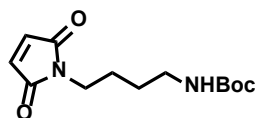
a biphasic mixture of EtOAc and half saturated aqueous NaHCO<sub>3</sub>. Additional saturated aqueous NaHCO<sub>3</sub> was added slowly until effervescence subsided. The organic layer was isolated and washed twice with saturated aqueous NaHCO<sub>3</sub>, once with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude product was purified via flash silica gel chromatography, eluting with 10% CH<sub>2</sub>Cl<sub>2</sub>/EtOAc and then recrystallized from EtOAc/hexanes to afford the desired maleimide.



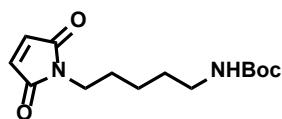
**tert-butyl (2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)carbamate (1.5-Et).** White solid (65%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.67 (s, 2H), 4.85 (br s, 1H), 3.59 (dd, *J* = 6.0, 4.2 Hz, 2H), 3.26 (q, *J* = 6.0 Hz, 2H), 1.44 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.82, 155.93, 134.15, 79.46, 28.28.



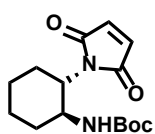
**tert-butyl (3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)propyl)carbamate (1.5-Pr).** White solid (68%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.70 (s, 2H), 4.92 (br s, 1H), 3.58 (t, *J* = 6.6 Hz, 2H), 3.07 (q, *J* = 6.6 Hz, 2H), 1.74 (p, *J* = 6.6 Hz, 2H), 1.43 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.82, 155.93, 134.15, 79.46, 28.28, 27.10.



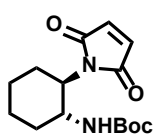
**tert-butyl (4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)butyl)carbamate (1.5-Bu).** White solid (55%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.69 (s, 2H), 4.56 (br s, 1H), 3.50 (t, *J* = 7.1 Hz, 2H), 3.15–3.09 (m, 2H), 1.63–1.50 (m, 2H), 1.43 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.87, 155.82, 134.33, 79.30, 39.63, 37.22, 28.32, 26.99, 25.79.



**tert-butyl (5-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)pentyl)carbamate (1.5-Pe).** White solid (72%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.70 (s, 2H), 4.55 (br s, 1H), 3.48 (t, *J* = 6.1 Hz, 2H), 3.15–3.08 (m, 2H), 1.43 (s, 9H), 1.67–1.18 (m, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.79, 155.86, 133.89, 79.82, 39.54, 37.22, 28.30, 27.01, 25.73, 22.77.



**tert-butyl ((1S,2S)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)cyclohexyl)carbamate (1.5-S,S-Ch).** White solid (65%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.62 (s, 2H), 4.30 (d, 1H, *J* = 9.2 Hz), 3.97 (br m, 1H), 3.64 (m, 1H, *J* = 4 Hz), 2.36 (q, 1H, *J* = 10 Hz), 2.04 (d, 1H, *J* = 10.5 Hz), 1.82–1.71 (m, 4H), 1.44–1.22 (m, 11H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 171.04, 155.17, 133.80, 79.13, 55.66, 50.21, 33.07, 28.56, 28.22, 25.30, 24.67.

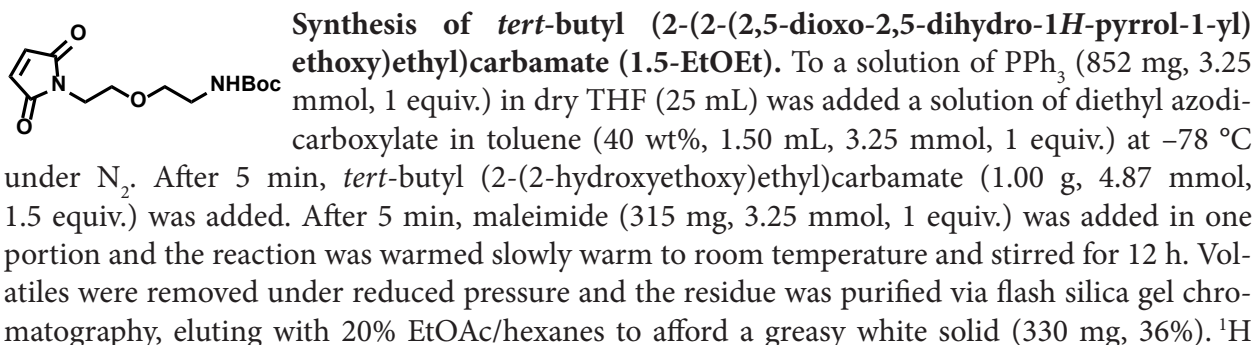
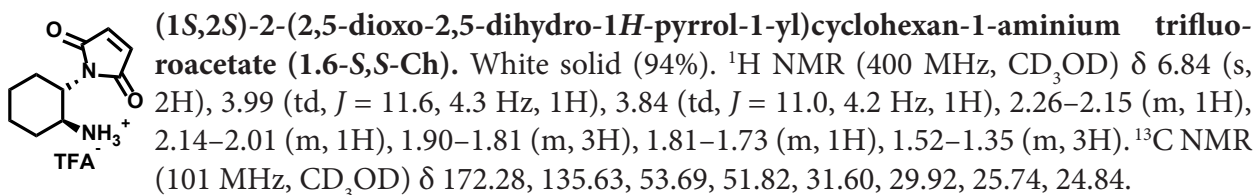
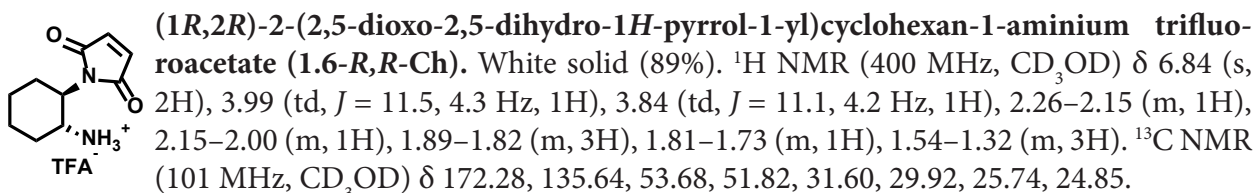
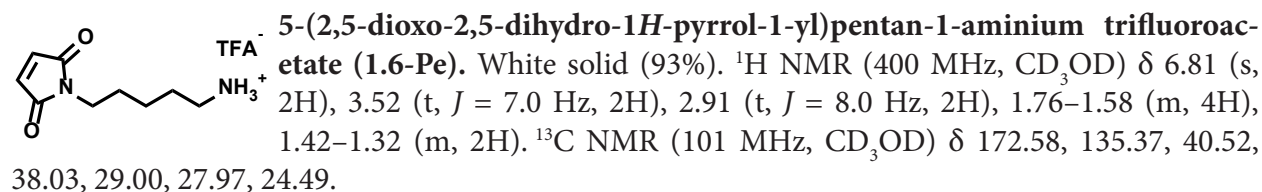
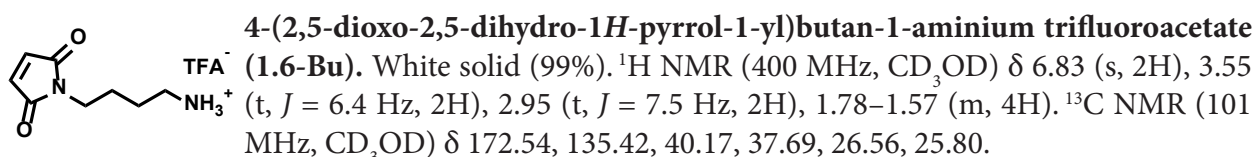
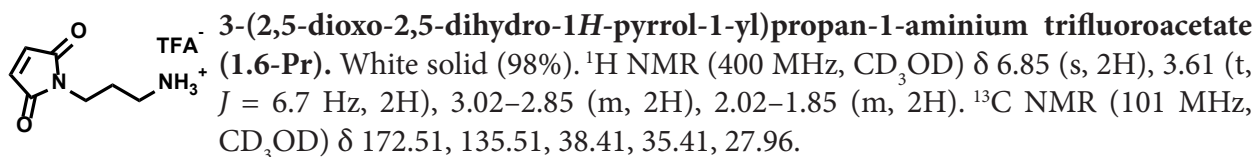
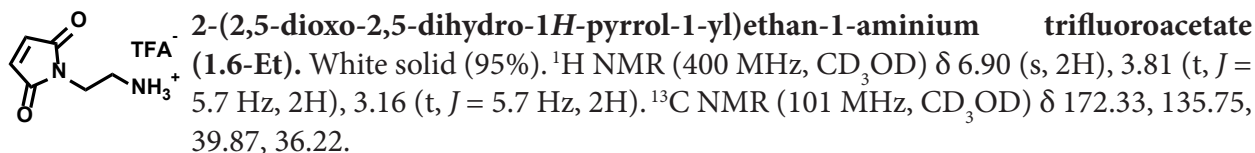


**tert-butyl ((1R,2R)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)cyclohexyl)carbamate (1.5-R,R-Ch).** White solid (67%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.62 (s, 2H), 4.30 (d, 1H, *J* = 9.2 Hz), 3.97 (m, 1H), 3.64 (m, 1H, *J* = 4 Hz), 2.36 (q, 1H, *J* = 10 Hz), 2.04 (d, 1H, *J* = 10.5 Hz), 1.82–1.71 (m, 4H), 1.44–1.22 (m, 11H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 171.04, 155.17, 133.80, 79.13, 55.66, 50.21, 33.07, 28.56, 28.22, 25.30, 24.67.

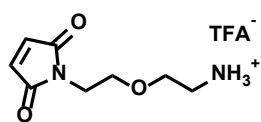
**Typical procedure for the synthesis of aminoalkylmaleimides (Example of 2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethan-1-aminium trifluoroacetate (1.6-Et)).** To a solution of *tert*-butyl (2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)carbamate (500 mg, 2.08 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (3 mL) was added neat trifluoroacetic acid (3 mL) dropwise at room temperature and the reaction was stirred under N<sub>2</sub>. After 30 min, volatiles were removed under reduced pressure. The residue was



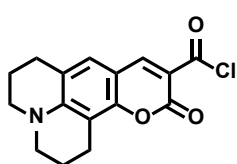
dissolved in a minimal amount of warm MeOH and precipitated with Et<sub>2</sub>O to afford the desired ammonium salt.



NMR (400 MHz CDCl<sub>3</sub>) δ 6.70 (d, *J* = 1.0 Hz, 2H), 4.88 (br-s, 1H), 3.70 (t, *J* = 5.6 Hz, 2H), 3.58 (t, *J* = 5.6 Hz, 2H), 3.47 (t, *J* = 5.2 Hz, 2H), 3.24 (q, *J* = 5.5 Hz, 2H), 1.42 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.79, 156.02, 134.28, 79.32, 69.93, 67.91, 40.41, 37.35, 28.52.

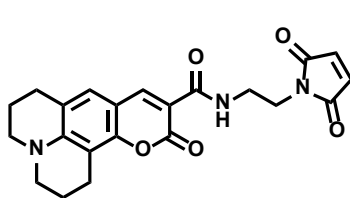


**Synthesis of 2-(2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethoxy)ethan-1-aminium trifluoroacetate (1.6-EtOEt).** To a solution of *tert*-butyl (2-(2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethoxy)ethyl)carbamate (300 mg, 1.06 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (6 mL) was added neat trifluoroacetic acid (2 mL) at room temperature and the reaction mixture was stirred for 30 min. The solution was concentrated under reduced pressure to afford a hygroscopic white solid (298 mg, 100%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 6.84 (s, 2H), 3.77–3.62 (m, 6H), 3.09 (t, *J* = 5.0 Hz, 2H). <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ 172.60, 161.88, 135.49, 119.15, 69.34, 67.18, 40.57, 37.77.

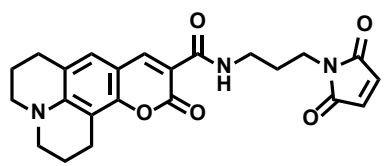


**Synthesis of 11-oxo-2,3,6,7-tetrahydro-1*H*,5*H*,11*H*-pyrano[2,3-*f*]pyrido[3,2,1-*ij*]quinoline-10-carbonyl chloride.** To a suspension of coumarin 343 (419 mg, 1.47 mmol, 1 equiv.) in CHCl<sub>3</sub> (20 mL) was added oxalyl chloride (0.63 mL, 7.34 mmol, 5 equiv.) at room temperature under N<sub>2</sub>. One drop of DMF was added and the reaction was stirred for 20 min until no more effervescence was observed and then heated to reflux until all solids had dissolved. To the hot reaction mixture was added cyclohexane (30 mL) and a red precipitate formed. After cooling to room temperature, the precipitate was collected by filtration and washed with cyclohexane to afford the acid chloride as a red solid (343 mg, 1.13 mmol, 77%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.44 (s, 1H), 6.98 (s, 1H), 3.39 (dt, *J* = 6.0, 4.1 Hz, 3H), 2.87 (t, *J* = 6.4 Hz, 2H), 2.78 (t, *J* = 6.3 Hz, 2H), 1.99 (dtd, *J* = 11.1, 6.4, 2.9 Hz, 4H).

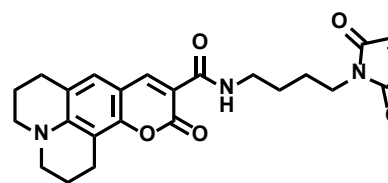
**Typical procedure for the synthesis of coumarin 343-maleimides (Example of *N*-(2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl)-11-oxo-2,3,6,7-tetrahydro-1*H*,5*H*,11*H*-pyrano[2,3-*f*]pyrido[3,2,1-*ij*]quinoline-10-carboxamide (1.3-Et)).** To a solution of coumarin 343 acid chloride (50 mg, 0.165 mmol, 1 equiv.) in dry CH<sub>2</sub>Cl<sub>2</sub> (3 mL) was added 2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethan-1-aminium trifluoroacetate (41 mg, 0.173 mmol, 1.05 equiv.) followed by dropwise addition of Et<sub>3</sub>N (0.10 mL, 0.740 mmol, 4.5 equiv.) at room temperature under N<sub>2</sub>. The reaction mixture was stirred for 12 h and then washed twice with 1 M aqueous HCl, three times with 1 M aqueous NaOH, once with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The residue was purified via flash silica gel chromatography, eluting with 50–66% EtOAc/hexanes.



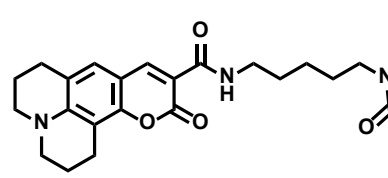
***N*-(2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl)-11-oxo-2,3,6,7-tetrahydro-1*H*,5*H*,11*H*-pyrano[2,3-*f*]pyrido[3,2,1-*ij*]quinoline-10-carboxamide (1.3-Et).** Orange-yellow solid (45%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.92 (t, *J* = 5.9 Hz, 1H), 8.55 (s, 1H), 6.97 (s, 1H), 6.69 (s, 2H), 3.78 (t, *J* = 5.8 Hz, 2H), 3.69–3.59 (m, 2H), 3.31 (q, *J* = 5.6 Hz, 4H), 2.85 (t, *J* = 6.4 Hz, 2H), 2.75 (t, *J* = 6.3 Hz, 2H), 1.95 (p, *J* = 6.3 Hz, 4H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.78, 164.23, 163.07, 152.79, 148.28, 148.25, 134.24, 127.14, 119.72, 108.72, 108.33, 105.74, 50.33, 49.92, 38.23, 37.57, 27.56, 21.24, 20.29, 20.19. HRMS (ESI) calculated for C<sub>22</sub>H<sub>21</sub>N<sub>3</sub>O<sub>5</sub> [M + H]<sup>+</sup> 408.1554, found 408.1551.



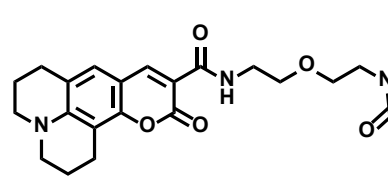
**N-(3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)propyl)-11-oxo-2,3,6,7-tetrahydro-1H,5H,11H-pyrano[2,3-f]pyrido[3,2,1-ij]quinoline-10-carboxamide (1.3-Pr).** Orange-yellow solid (52%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.01 (t, *J* = 5.9 Hz, 1H), 8.57 (s, 1H), 6.99 (s, 1H), 6.70 (s, 2H), 3.63 (t, *J* = 6.8 Hz, 2H), 3.42 (q, *J* = 6.6 Hz, 2H), 3.32 (q, *J* = 5.1 Hz, 4H), 2.88 (t, *J* = 6.4 Hz, 2H), 2.76 (t, *J* = 6.3 Hz, 2H), 2.02–1.85 (m, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.92, 163.83, 163.10, 152.78, 148.21, 148.15, 134.24, 127.12, 119.68, 109.12, 108.35, 105.79, 50.35, 49.93, 36.94, 35.69, 28.82, 27.59, 21.27, 20.34, 20.23. HRMS (ESI) calculated for C<sub>23</sub>H<sub>23</sub>N<sub>3</sub>O<sub>5</sub> [M + H]<sup>+</sup> 422.1710, found 422.1710.



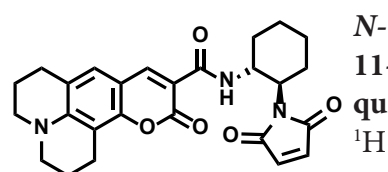
**N-(4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)butyl)-11-oxo-2,3,6,7-tetrahydro-1H,5H,11H-pyrano[2,3-f]pyrido[3,2,1-ij]quinoline-10-carboxamide (1.3-Bu).** Orange-yellow solid (56%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.92–8.77 (m, 1H), 8.55 (s, 1H), 6.96 (s, 1H), 6.66 (s, 2H), 3.54 (t, *J* = 6.9 Hz, 2H), 3.42 (q, *J* = 6.5 Hz, 2H), 3.35–3.25 (m, 4H), 2.85 (t, *J* = 6.4 Hz, 2H), 2.74 (t, *J* = 6.3 Hz, 2H), 2.01–1.89 (m, 4H), 1.73–1.51 (m, 4H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.85, 163.65, 163.13, 152.67, 148.15, 148.07, 134.13, 127.04, 119.66, 109.00, 108.28, 105.67, 50.28, 49.86, 39.00, 37.59, 27.52, 27.10, 26.22, 21.20, 20.26, 20.17. HRMS (ESI) calculated for C<sub>24</sub>H<sub>25</sub>N<sub>3</sub>O<sub>5</sub> [M + H]<sup>+</sup> 436.1867, found 436.1871.



**N-(5-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)pentyl)-11-oxo-2,3,6,7-tetrahydro-1H,5H,11H-pyrano[2,3-f]pyrido[3,2,1-ij]quinoline-10-carboxamide (1.3-Pe).** Orange-yellow solid (40%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.84 (t, *J* = 5.7 Hz, 1H), 8.57 (s, 1H), 6.98 (s, 1H), 6.67 (s, 2H), 3.51 (t, *J* = 7.2 Hz, 2H), 3.40 (q, *J* = 6.8 Hz, 2H), 3.31 (q, *J* = 5.2 Hz, 4H), 2.87 (t, *J* = 6.4 Hz, 2H), 2.75 (t, *J* = 6.2 Hz, 2H), 2.01–1.90 (m, 4H), 1.62 (m, 4H), 1.36 (m, 2H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.93, 163.61, 163.17, 152.70, 148.09, 134.14, 127.08, 119.67, 109.19, 108.34, 105.73, 50.32, 49.90, 39.46, 37.81, 29.26, 28.38, 27.57, 24.32, 21.25, 20.32, 20.23. HRMS (ESI) calculated for C<sub>25</sub>H<sub>27</sub>N<sub>3</sub>O<sub>5</sub> [M + H]<sup>+</sup> 450.2023, found 450.2025.

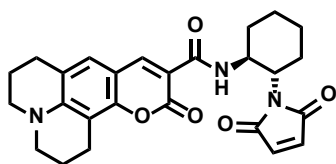


**N-(2-(2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethoxy)ethyl)-11-oxo-2,3,6,7-tetrahydro-1H,5H,11H-pyrano[2,3-f]pyrido[3,2,1-ij]quinoline-10-carboxamide (1.3-EtOEt).** Orange-yellow solid (34%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.97 (t, *J* = 5.1 Hz, 1H), 8.56 (s, 1H), 6.98 (s, 1H), 6.69 (s, 2H), 3.74 (t, *J* = 5.6 Hz, 2H), 3.67–3.54 (m, 6H), 3.35–3.27 (m, 4H), 2.87 (t, *J* = 6.4 Hz, 2H), 2.76 (t, *J* = 6.3 Hz, 2H), 2.01–1.91 (m, 4H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.64, 163.57, 162.63, 152.50, 148.00, 147.76, 134.05, 126.87, 119.49, 108.66, 108.00, 105.42, 69.24, 67.61, 50.11, 49.68, 39.25, 37.03, 27.33, 21.00, 20.06, 20.00. HRMS (ESI) calculated for C<sub>24</sub>H<sub>25</sub>N<sub>3</sub>O<sub>6</sub> [M + H]<sup>+</sup> 452.1816, found 452.1818.

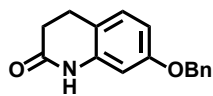


**N-((1R,2R)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)cyclohexyl)-11-oxo-2,3,6,7-tetrahydro-1H,5H,11H-pyrano[2,3-f]pyrido[3,2,1-ij]quinoline-10-carboxamide (1.3-R,R-Ch).** Orange-yellow solid (41%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.67 (d, *J* = 8.7 Hz, 1H), 8.46 (s, 1H), 6.93

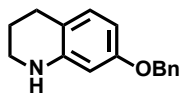
(s, 1H), 6.56 (s, 2H), 4.57 (qd,  $J = 11.0, 8.9, 4.0$  Hz, 1H), 3.92 (td,  $J = 12.2, 4.0$  Hz, 1H), 3.31 (q,  $J = 5.6$  Hz, 4H), 2.85 (t,  $J = 6.4$  Hz, 2H), 2.74 (t,  $J = 6.3$  Hz, 2H), 2.37 (qd,  $J = 12.7, 3.5$  Hz, 1H), 2.17 (d,  $J = 12.5$  Hz, 1H), 1.95 (p,  $J = 6.2$  Hz, 4H), 1.90–1.73 (m, 4H), 1.56–1.29 (m, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.00, 163.25, 162.89, 152.73, 148.18, 148.15, 133.87, 127.04, 119.61, 108.97, 108.31, 105.80, 54.87, 50.34, 49.94, 49.61, 33.13, 29.09, 27.59, 25.56, 24.75, 21.28, 20.35, 20.22. HRMS (ESI) calculated for  $\text{C}_{26}\text{H}_{27}\text{N}_3\text{O}_5$   $[\text{M} + \text{H}]^+$  462.2023, found 462.2022.



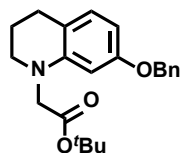
***N*-((1*S*,2*S*)-2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)cyclohexyl)-11-oxo-2,3,6,7-tetrahydro-1*H*,5*H*,11*H*-pyrano[2,3-*f*]pyrido[3,2,1-*ij*]quinoline-10-carboxamide (1.3-*S,S*-Ch).** Orange-yellow solid (38  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.67 (d,  $J = 8.7$  Hz, 1H), 8.45 (s, 1H), 6.92 (s, 1H), 6.55 (s, 2H), 4.56 (qd,  $J = 11.1, 4.2$  Hz, 1H), 3.91 (td,  $J = 12.3, 10.8, 4.0$  Hz, 1H), 3.31 (q,  $J = 5.6$  Hz, 4H), 2.84 (t,  $J = 6.4$  Hz, 2H), 2.74 (t,  $J = 6.3$  Hz, 2H), 2.36 (qd,  $J = 12.8, 3.5$  Hz, 1H), 2.16 (d,  $J = 11.6$  Hz, 1H), 1.95 (p,  $J = 6.3$  Hz, 4H), 1.89–1.70 (m, 3H), 1.55–1.28 (m, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.98, 163.23, 162.87, 152.70, 148.17, 148.12, 133.86, 127.02, 119.60, 108.91, 108.27, 105.76, 54.85, 50.32, 49.91, 49.58, 33.11, 29.07, 27.56, 25.53, 24.72, 21.25, 20.32, 20.19. HRMS (ESI) calculated for  $\text{C}_{26}\text{H}_{27}\text{N}_3\text{O}_5^+$  ( $[\text{M} + \text{H}]^+$ ) 462.2023, found 462.2023.



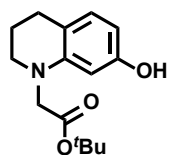
**Synthesis of 7-(benzyloxy)-3,4-dihydroquinolin-2(1*H*)-one.** To a mixture of 7-hydroxy-3,4-dihydro-2(1*H*)-quinolinone (8.16 g, 50.0 mmol, 1 equiv.) and  $\text{K}_2\text{CO}_3$  (8.29 g, 60.0 mmol, 1.2 equiv.) in dry DMF (50 mL) was added benzyl bromide (8.90 mL, 75.0 mmol, 1.5 equiv.) at room temperature under  $\text{N}_2$  and the reaction stirred for 12 h. The reaction was poured into ice cold  $\text{H}_2\text{O}$  (375 mL) and the resulting precipitate was collected by filtration and washed with additional cold  $\text{H}_2\text{O}$ . The crude off-white solid was recrystallized from MeOH to afford the desired product as a white crystalline solid (9.0 g, 71%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.72 (br s, 1H), 7.44–7.30 (m, 5H), 7.06 (d,  $J = 8.3$  Hz, 1H), 6.60 (dd,  $J = 8.3, 2.5$  Hz, 1H), 6.38 (d,  $J = 2.5$  Hz, 1H), 5.04 (s, 2H), 2.90 (t,  $J = 7.5$  Hz, 2H), 2.62 (t,  $J = 7.5$  Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.11, 158.34, 138.22, 136.84, 128.66, 128.60, 128.07, 127.41, 116.08, 109.07, 102.63, 70.26, 31.05, 24.60.



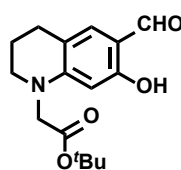
**Synthesis of 7-(benzyloxy)-1,2,3,4-tetrahydroquinoline (1.9).** To a suspension of 7-(benzyloxy)-3,4-dihydroquinolin-2(1*H*)-one (8.72 g, 34.4 mmol, 1 equiv.) in dry THF (60 mL) was added slowly a solution of  $\text{BH}_3$ -THF in THF (0.94 M, 100 mL, 94.0 mmol, 2.73 equiv.) at 0  $^\circ\text{C}$  under  $\text{N}_2$ . The reaction was warmed to room temperature and stirred for 1 h and then heated to reflux for 2 h. The reaction mixture was cooled to 0  $^\circ\text{C}$  and cautiously quenched with MeOH (70 mL). Volatiles were removed under reduced pressure and the crude material was purified via flash silica column chromatography, eluting with 10–20% EtOAc/hexanes to afford a viscous oil, which solidified under high vacuum as a hard white solid (8.24 g, 99%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.50–7.27 (m, 5H), 6.86 (d,  $J = 8.3$  Hz, 1H), 6.29 (dd,  $J = 8.2, 2.6$  Hz, 1H), 6.13 (d,  $J = 2.5$  Hz, 1H), 5.01 (s, 2H), 3.82 (br s, 1H), 3.35–3.22 (m, 2H), 2.71 (t,  $J = 6.4$  Hz, 2H), 2.00–1.87 (m, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  158.21, 145.67, 137.68, 130.20, 128.61, 127.85, 127.50, 114.49, 103.80, 100.58, 70.04, 42.03, 26.42, 22.52.



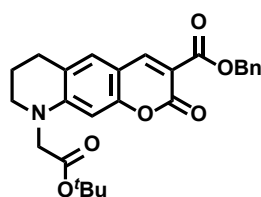
**Synthesis of *tert*-butyl 2-(7-(benzyloxy)-3,4-dihydroquinolin-1(2*H*)-yl)acetate (1.10).** To a mixture of 7-(benzyloxy)-1,2,3,4-tetrahydroquinoline (8.24 g, 34.4 mmol, 1 equiv.) and  $K_2CO_3$  (5.23 g, 37.9 mmol, 1.1 equiv.) in dry DMF (170 mL) was added *tert*-butyl bromoacetate (5.59 mL, 37.9 mmol, 1.1 equiv.) at room temperature under  $N_2$ . The reaction was heated to 80 °C for 5 h and then concentrated under reduced pressure. The crude material was dissolved in warm EtOAc and filtered. The filtrate was concentrated under reduced pressure and purified via flash silica column chromatography, eluting with 5–10% EtOAc/hexanes to afford a bright white solid (10.51 g, 86%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.45–7.27 (m, 5H), 6.84 (dt,  $J = 8.3, 1.0$  Hz, 1H), 6.24 (dd,  $J = 8.1, 2.4$  Hz, 1H), 6.10 (d,  $J = 2.4$  Hz, 1H), 4.99 (s, 2H), 3.85 (s, 2H), 3.37 (t,  $J = 6.3$  Hz, 2H), 2.71 (t,  $J = 6.3$  Hz, 2H), 2.01–1.89 (m, 2H), 1.44 (s, 9H).  $^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta$  170.02, 158.48, 145.79, 137.83, 129.61, 128.63, 127.87, 127.61, 115.78, 102.28, 98.36, 81.56, 70.12, 54.19, 50.70, 28.28, 27.38, 22.69.



**Synthesis of *tert*-butyl 2-(7-hydroxy-3,4-dihydroquinolin-1(2*H*)-yl)acetate (1.11).** To a solution of *tert*-butyl 2-(7-(benzyloxy)-3,4-dihydroquinolin-1(2*H*)-yl)acetate (10.51 g, 29.7 mmol) in MeOH (200 mL) was added 10% Pd/C (1 g) and the reaction mixture was stirred under an atmosphere of  $H_2$  (1 atm) for 2.5 h. The suspension was filtered through Celite and concentrated under reduced pressure. The crude mixture was purified via flash silica column chromatography, eluting with 20–50% EtOAc/hexanes to afford a waxy white solid (7.13 g, 91%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  6.79 (dd,  $J = 8.0, 1.1$  Hz, 1H), 6.07 (dd,  $J = 8.0, 2.4$  Hz, 1H), 5.93 (d,  $J = 2.4$  Hz, 1H), 4.54 (s, 1H), 3.84 (s, 2H), 3.36 (t,  $J = 5.6$  Hz, 2H), 2.69 (t,  $J = 6.3$  Hz, 2H), 2.00–1.89 (m, 2H), 1.44 (s, 9H).  $^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta$  170.36, 154.96, 144.72, 129.86, 115.44, 103.29, 97.83, 81.72, 54.22, 50.69, 28.25, 27.32, 22.64.

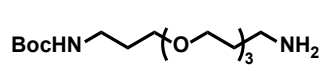


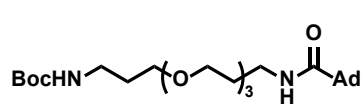
**Synthesis of *tert*-butyl 2-(6-formyl-7-hydroxy-3,4-dihydroquinolin-1(2*H*)-yl)acetate (1.12).**  $POCl_3$  (1.34 mL, 14.4 mmol, 1 equiv.) was added dropwise to dry DMF (2.80 mL, 36.3 mmol, 2.52 equiv.) at room temperature and stirred for 2 h under  $N_2$ . To this solution was added *tert*-butyl 2-(7-hydroxy-3,4-dihydroquinolin-1(2*H*)-yl)acetate (3.79 g, 14.4 mmol, 1 equiv.) in dry DMF (80 mL) at room temperature and the reaction was stirred for 12 h. The reaction mixture was poured over crushed ice (100 mL) and extracted into  $CH_2Cl_2$ . The aqueous layer was adjusted to pH 8 with 2 M NaOH and extracted into  $CH_2Cl_2$  three times. The combined organic layers were washed with brine, dried over  $Na_2SO_4$ , and concentrated under reduced pressure. The crude mixture was purified via flash silica column chromatography, eluting with 20–33–50% EtOAc/hexanes to afford a white solid (2.06 g, 49%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  11.51 (s, 1H), 9.47 (s, 1H), 7.00 (d,  $J = 1.1$  Hz, 1H), 5.86 (s, 1H), 3.92 (s, 2H), 3.44 (t,  $J = 5.8$  Hz, 2H), 2.71 (t,  $J = 6.3$  Hz, 2H), 1.98 (p,  $J = 6.0$  Hz, 2H), 1.47 (s, 9H).  $^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta$  192.63, 168.54, 163.15, 152.24, 133.40, 115.30, 111.99, 96.33, 82.60, 53.94, 51.06, 28.22, 27.12, 21.96.

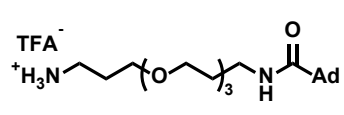


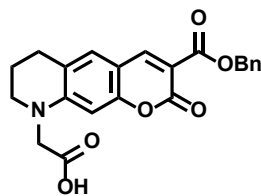
**Synthesis of benzyl 9-(2-(*tert*-butoxy)-2-oxoethyl)-2-oxo-6,7,8,9-tetrahydro-2*H*-pyrano[3,2-*g*]quinoline-3-carboxylate (1.13).** To a solution of dibenzyl malonate (1.94 mL, 7.78 mmol, 1.1 equiv.) and *tert*-butyl 2-(6-formyl-7-hydroxy-3,4-dihydroquinolin-1(2*H*)-yl)acetate (2.06 g, 7.07 mmol, 1 equiv.) in dry toluene (40 mL) was added piperidine (0.70 mL, 7.07 mmol, 1 equiv.) and the reaction was heated to reflux under  $N_2$  for 2 h.

Volatiles were removed under reduced pressure and the crude mixture was purified via flash silica column chromatography, eluting with 50% EtOAc/hexanes to afford a bright yellow solid (3.08 g, 97%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.39 (s, 1H), 7.51–7.44 (m, 2H), 7.43–7.27 (m, 3H), 7.08 (s, 1H), 6.22 (s, 1H), 5.35 (s, 2H), 3.94 (s, 2H), 3.48 (t, *J* = 5.8 Hz, 2H), 2.79 (t, *J* = 6.2 Hz, 2H), 2.01 (p, *J* = 6.1 Hz, 2H), 1.47 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 168.20, 164.01, 158.35, 157.40, 151.23, 149.38, 129.17, 128.68, 128.31, 128.24, 121.06, 109.38, 108.55, 96.06, 82.99, 66.82, 53.95, 51.05, 28.23, 27.39, 21.52.

 **Synthesis of *tert*-butyl (3-(2-(2-(3-aminopropoxy)ethoxy)ethoxy)propyl)carbamate (1.14).** To a solution of 4,7,10-trioxa-1,13-tridecane-diamine (16.1 mL, 73.5 mmol, 8 equiv.) in dioxane (120 mL) was added a solution of di-*tert*-butyl dicarbonate (2.11 mL, 9.16 mmol, 1 equiv.) in dioxane (50 mL) over 2 h at room temperature under N<sub>2</sub> and the reaction mixture was stirred for 6 h. Solvent was removed under reduced pressure and the residue was taken up in H<sub>2</sub>O (100 mL) and extracted three times into EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure to afford the desired product as a clear colorless oil (1.95 g, 66%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.13 (br s, 1H), 3.63 (dt, *J* = 5.8, 3.1 Hz, 4H), 3.60–3.49 (m, 8H), 3.21 (q, *J* = 6.2 Hz, 2H), 2.77 (t, *J* = 6.7 Hz, 2H), 1.74 (p, *J* = 6.1 Hz, 4H), 1.42 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 157.03, 78.93, 70.50, 70.47, 70.18, 70.11, 69.55, 69.43, 39.64, 38.42, 32.55, 29.61.

 **Synthesis of *tert*-butyl (1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1-oxo-6,9,12-trioxa-2-azapentadecan-15-yl)carbamate (1.15).** To a solution of *tert*-butyl (3-(2-(2-(3-aminopropoxy)ethoxy)ethoxy)propyl)carbamate (1.95 g, 6.09 mmol, 1 equiv.) and Et<sub>3</sub>N (2.10 mL, 15.03 mmol, 2.45 equiv.) in CH<sub>2</sub>Cl<sub>2</sub> (75 mL) was added 1-adamantanecarbonyl chloride (1.29 g, 6.51 mmol, 1.07 equiv.) at room temperature and the reaction was stirred under N<sub>2</sub> for 3 h. The reaction mixture was washed with two portions of H<sub>2</sub>O, one portion of 1 M aqueous HCl, one portion of saturated aqueous NaHCO<sub>3</sub>, one portion of brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified by flash silica gel chromatography, eluting with 33% acetone/cyclohexane to afford a clear colorless oil (1.57 g, 53%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.29 (br s, 1H), 4.96 (br s, 1H), 3.70–3.50 (m, 12H), 3.35 (q, *J* = 5.8 Hz, 2H), 3.22 (q, *J* = 6.3 Hz, 2H), 2.06–1.98 (m, 3H), 1.85–1.81 (m, 6H), 1.81–1.63 (m, 4H), 1.43 (s, 6H), 1.42 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 177.11, 156.14, 78.50, 70.87, 70.73, 70.68, 70.54, 70.35, 69.79, 40.59, 39.35, 38.25, 37.94, 36.72, 29.15, 28.59, 28.32, 27.04.

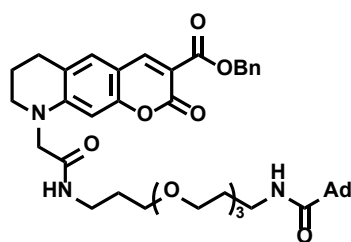
 **Synthesis of 1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1-oxo-6,9,12-trioxa-2-azapentadecan-15-aminium trifluoroacetate (1.16).** To a solution of *tert*-butyl (1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1-oxo-6,9,12-trioxa-2-azapentadecan-15-yl)carbamate (1.57 g, 3.25 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) was added neat trifluoroacetic acid (5 mL) dropwise at room temperature and the reaction was stirred under N<sub>2</sub>. After 2 h, volatiles were removed under reduced pressure. The residue was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) and once again concentrated under reduced pressure to remove residual trifluoroacetic acid, affording the product as a highly viscous oil (1.6 g, 99%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 7.72 (br s, 3H), 7.34 (t, *J* = 5.6 Hz, 1H), 3.56–3.43 (m, 10H), 3.36 (t, *J* = 6.3 Hz, 2H), 3.07 (q, *J* = 6.5 Hz, 2H), 2.85 (h, *J* = 6.1 Hz, 2H), 1.99–1.91 (m, 3H), 1.82–1.56 (m, 16H). LRMS (ESI) calculated for C<sub>21</sub>H<sub>39</sub>N<sub>2</sub>O<sub>4</sub><sup>+</sup> ([M+H]<sup>+</sup>) 383.3, found 383.4.



### Synthesis of 2-(3-((benzyloxy)carbonyl)-2-oxo-7,8-dihydro-2H-pyrano

[3,2-g]quinolin-9(6H)-yl)acetic acid (1.17). To a solution of benzyl 9-(2-(*tert*-butoxy)-2-oxoethyl)-2-oxo-6,7,8,9-tetrahydro-2H-pyrano[3,2-g]quinoline-3-carboxylate (1.0 g, 2.23 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) was added neat trifluoroacetic acid (10 mL) dropwise at room temperature under N<sub>2</sub>. After 2 h, volatiles were removed under reduced pressure. The residue was dissolved

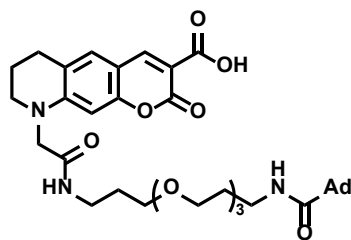
in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and once again concentrated under reduced pressure to remove residual trifluoroacetic acid. The crude solid was triturated with hot EtOAc, diluted with hexanes, and filtered to afford a bright yellow solid (756 mg, 86%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.55 (s, 1H), 7.47 (d, *J* = 7.5 Hz, 2H), 7.43–7.30 (m, 3H), 6.39 (s, 1H), 5.27 (s, 2H), 4.24 (s, 2H), 3.42 (t, *J* = 5.4 Hz, 2H), 2.73 (t, *J* = 4.8 Hz, 3H), 1.95–1.81 (m, 2H). LRMS (ESI) calculated for C<sub>22</sub>H<sub>18</sub>NO<sub>6</sub><sup>-</sup> ([M+H]<sup>-</sup>) 392.1, found 392.3.



### Synthesis of benzyl 9-(1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1,17-dioxo-6,9,12-trioxa-2,16-diazaoctadecan-18-yl)-2-oxo-6,7,8,9-tetra

hydro-2H-pyrano[3,2-g]quinoline-3-carboxylate (1.18). To a solution of 2-(3-((benzyloxy)carbonyl)-2-oxo-7,8-dihydro-2H-pyrano [3,2-g]quinolin-9(6H)-yl)acetic acid (500 mg, 1.27 mmol, 1 equiv.), DMAP (776 mg, 6.36 mmol, 5 equiv.), and 1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1-oxo-6,9,12-trioxa-2-azapentadecan-15-aminium trifluoroacetate (947 mg, 1.91 mmol, 1.5 equiv.) in dry DMF (25 mL) was added EDCI (268 mg, 1.40 mmol, 1.1 equiv.) at 0 °C under N<sub>2</sub>. The reaction was warmed to room temperature and stirred for 24 h.

The reaction mixture was concentrated under reduced pressure and purified via flash silica gel chromatography, eluting with 2.5–5% MeOH/CH<sub>2</sub>Cl<sub>2</sub> to afford a yellow solid (680 mg, 71%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.30 (s, 1H), 7.45 (d, *J* = 7.2 Hz, 2H), 7.35 (dt, *J* = 14.1, 7.2 Hz, 3H), 6.99 (s, 1H), 6.24 (s, 1H), 5.32 (s, 2H), 3.95 (s, 2H), 3.63–3.45 (m, 12H), 3.44–3.35 (m, 4H), 3.30 (t, *J* = 6.1 Hz, 2H), 2.73 (t, *J* = 5.7 Hz, 2H), 2.13–1.60 (m, 21H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 178.06, 167.80, 163.78, 157.20, 151.37, 149.10, 136.10, 129.09, 128.64, 128.32, 128.28, 121.18, 112.23, 109.08, 108.45, 96.37, 70.50, 70.48, 70.31, 70.03, 69.95, 66.77, 55.48, 51.29, 40.52, 39.28, 37.99, 36.62, 34.74, 31.66, 29.20, 29.13, 28.23, 27.25, 25.35, 22.73, 21.40, 14.21.

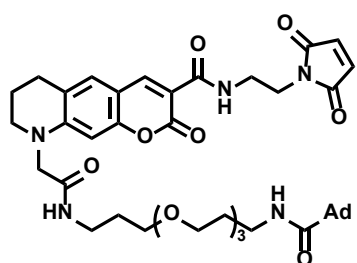


### Synthesis of 9-(1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1,17-dioxo-6,9,12-trioxa-2,16-diazaoctadecan-18-yl)-2-oxo-6,7,8,9-tetrahydro-2H-pyrano[3,2-g]quinoline-3-carboxylic acid (1.19).

To a solution of benzyl 9-(1-((3*r*,5*r*,7*r*)-adamantan-1-yl)-1,17-dioxo-6,9,12-trioxa-2,16-diazaoctadecan-18-yl)-2-oxo-6,7,8,9-tetrahydro-2H-pyrano[3,2-g]quinoline-3-carboxylate (680 mg, 0.897 mmol) in MeOH (100 mL) was added 10% Pd/C (250 mg) and the suspension

was placed under an atmosphere of H<sub>2</sub> (1 atm) and stirred for 4 h. The reaction mixture was filtered through Celite, eluting with MeOH, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 5% MeOH/CH<sub>2</sub>Cl<sub>2</sub> with 0.1% trifluoroacetic acid to afford a bright yellow solid (540 mg, 90%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 10.76 (br s, 1H), 8.45 (s, 1H), 7.23 (t, *J* = 5.0 Hz, 1H), 7.14 (s, 1H), 6.64 (t, *J* = 5.0 Hz, 1H), 6.34 (s, 1H), 4.06 (s, 2H), 3.65–3.49 (m, 14H), 3.46–3.38 (m, 2H), 3.37–3.29 (m, 2H), 2.83 (t, *J* = 6.3 Hz, 2H), 2.11–1.98 (m, 5H), 1.86–1.59 (m, 16H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 176.86, 169.99, 165.27, 156.75,

152.46, 149.95, 129.78, 122.89, 118.99, 118.78, 109.39, 106.02, 96.63, 70.42, 70.39, 70.35, 70.31, 69.99, 69.92, 55.42, 51.41, 39.79, 38.99, 38.42, 38.27, 36.47, 28.89, 28.09, 27.18, 21.05.



**Synthesis of 9-(1-((3s)-adamantan-1-yl)-1,17-dioxo-6,9,12-trioxo-2,16-diazaoctadecan-18-yl)-N-(2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)-2-oxo-6,7,8,9-tetrahydro-2H-pyrano[3,2-g]quinoline-3-carboxamide (1.7).** To a solution of 9-(1-((3r,5r,7r)-adamantan-1-yl)-1,17-dioxo-6,9,12-trioxo-2,16-diazaoctadecan-18-yl)-2-oxo-6,7,8,9-tetrahydro-2H-pyrano[3,2-g]quinoline-3-carboxylic acid (50 mg, 0.75 mmol, 1 equiv.) in dry THF (2.5 mL) was added oxalyl chloride (0.03 mL, 0.375 mmol, 5 equiv.) followed by one

drop of DMF at room temperature under  $N_2$ . After 30 min, volatiles were removed under reduced pressure. Excess oxalyl chloride was removed by azeotropic distillation with  $CH_2Cl_2$ . The resulting acid chloride was dissolved in dry THF (5 mL) and 2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethan-1-aminium trifluoroacetate (21 mg, 0.083 mmol, 1.1 equiv.) was added followed by *N,N*-diisopropyldeethylamine (0.065 mL, 0.375 mmol, 5 equiv.) at room temperature under  $N_2$ . After 24 h, volatiles were removed and the crude mixture was purified via reverse phase HPLC (C18, 5–95% MeCN/ $H_2O$  with 0.1% TFA) and isolated as a yellow solid (2.1 mg, 3.5%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  8.88 (t,  $J = 6.0$  Hz, 1H), 8.62 (s, 1H), 7.15 (s, 1H), 6.86 (t,  $J = 5.7$  Hz, 1H), 6.69 (s, 2H), 6.31 (s, 1H), 6.22 (t,  $J = 4.5$  Hz, 1H), 3.94 (d,  $J = 5.4$  Hz, 2H), 3.82–3.74 (m, 2H), 3.67–3.59 (m, 2H), 3.59–3.47 (m, 14H), 3.39 (q,  $J = 6.1$  Hz, 2H), 3.31 (q,  $J = 6.1$  Hz, 2H), 2.83 (t,  $J = 6.2$  Hz, 2H), 2.09–1.96 (m, 6H), 1.85–1.61 (m, 15H). HRMS (ESI) calculated for  $C_{42}H_{56}N_5O_{10}^+$  ( $[M+H]^+$ ) 790.4022, found 790.4028.

### 1.13 References

- (1) Nelson, N.; Ben-Shem, A. The Complex Architecture of Oxygenic Photosynthesis. *Nature* **2004**, *5* (12), 971–982.
- (2) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*; Wiley-Blackwell: Oxford, 2013.
- (3) Green, B. R.; Parson, W. W. *Light-Harvesting Antennas in Photosynthesis*; Kluwer Academic: Dordrecht, 2003.
- (4) Schwartz, E.; Le Gac, S.; Cornelissen, J. J. L. M.; Nolte, R. J. M.; Rowan, A. E. Macromolecular Multi-Chromophoric Scaffolding. *Chem. Soc. Rev.* **2010**, *39* (5), 1576.
- (5) Maggini, L.; Bonifazi, D. Hierarchised Luminescent Organic Architectures: Design, Synthesis, Self-Assembly, Self-Organisation and Functions. *Chem. Soc. Rev.* **2011**, *41* (1), 211.
- (6) Teo, Y. N.; Kool, E. T. DNA-Multichromophore Systems. *Chem. Rev.* **2012**, *112* (7), 4221–4245.
- (7) Gilat, S. L.; Adronov, A.; Fréchet, J. M. J. Light Harvesting and Energy Transfer in Novel Convergently Constructed Dendrimers. *Angew. Chem. Int. Ed. Engl.* **1999**, *38* (10), 1422–1427.



- (8) Adronov, A.; Gilat, S. L.; Fréchet, J. M. J.; Ohta, K.; Neuwahl, F. V. R.; Fleming, G. R. Light Harvesting and Energy Transfer in Laser–Dye-Labeled Poly(Aryl Ether) Dendrimers. *J. Am. Chem. Soc.* **2000**, *122* (6), 1175–1185.
- (9) Choi, M.-S.; Yamazaki, T.; Yamazaki, I.; Aida, T. Bioinspired Molecular Design of Light-Harvesting Multiporphyrin Arrays. *Angew. Chem. Int. Ed. Engl.* **2004**, *43* (2), 150–158.
- (10) List, E.; Creely, C.; Leising, G.; Schulte, N. Excitation Energy Migration in Highly Emissive Semiconducting Polymers. *Chem. Phys.* **2000**, *325* (1–3), 132–138.
- (11) Kim, J.; McQuade, D. T.; Rose, A.; Zhu, Z. Directing Energy Transfer Within Conjugated Polymer Thin Films. *J. Am. Chem. Soc.* **2001**, *123* (46), 11488–11489.
- (12) Klare, J. E.; Murray, I. P.; Goldberger, J.; Stupp, S. I. Assembling P-Type Molecules on Single Wall Carbon Nanotubes for Photovoltaic Devices. *Chem. Commun.* **2009**, *25*, 3705–3707.
- (13) Grätzel, M. Dye-Sensitized Solar Cells. *J. Photochem. Photobiol., C* **2003**, *4*, 145–153.
- (14) Chen, P.-Y.; Dang, X.; Klug, M. T.; Qi, J.; Dorval Courchesne, N.-M.; Burpo, F. J.; Fang, N.; Hammond, P. T.; Belcher, A. M. Versatile Three-Dimensional Virus-Based Template for Dye-Sensitized Solar Cells with Improved Electron Transport and Light Harvesting. *ACS Nano* **2013**, *7* (8), 6563–6574.
- (15) 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.
- (16) Ma, Y.-Z.; Miller, R. A.; Fleming, G. R.; Francis, M. B. Energy Transfer Dynamics in Light-Harvesting Assemblies Templated by the Tobacco Mosaic Virus Coat Protein. *J. Phys. Chem. B* **2008**, *112* (22), 6887–6892.
- (17) 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.
- (18) Miller, R. A.; Stephanopoulos, N.; McFarland, J. M.; Rosko, A. S.; Geissler, P. L.; Francis, M. B. Impact of Assembly State on the Defect Tolerance of TMV-Based Light Harvesting Arrays. *J. Am. Chem. Soc.* **2010**, *132* (17), 6068–6074.
- (19) Noriega, R.; Finley, D. T.; Haberstroh, J.; Geissler, P. L.; Francis, M. B.; Ginsberg, N. S. Manipulating Excited-State Dynamics of Individual Light-Harvesting Chromophores Through Restricted Motions in a Hydrated Nanoscale Protein Cavity. *J. Phys. Chem. B* **2015**, *119* (23), 6963–6973.
- (20) Endo, M.; Fujitsuka, M.; Majima, T. Porphyrin Light-Harvesting Arrays Constructed in the Recombinant Tobacco Mosaic Virus Scaffold. *Chem. Eur. J.* **2007**, *13* (31), 8660–8666.
- (21) Nam, Y. S.; Shin, T.; Park, H.; Magyar, A. P.; Choi, K.; Fantner, G.; Nelson, K. A.; Belcher, A. M. Virus-Templated Assembly of Porphyrins Into Light-Harvesting Nanoantennae. *J. Am. Chem. Soc.* **2010**, *132* (5), 1462–1463.
- (22) Yoneda, Y.; Noji, T.; Katayama, T.; Mizutani, N.; Komori, D.; Nango, M.; Miyasaka, H.;

- Itoh, S.; Nagasawa, Y.; Dewa, T. Extension of Light-Harvesting Ability of Photosynthetic Light-Harvesting Complex 2 (LH2) Through Ultrafast Energy Transfer From Covalently Attached Artificial Chromophores. *J. Am. Chem. Soc.* **2015**, *137* (40), 13121–13129.
- (23) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucl. Acids Res.* **2000**, *28* (1), 235–242.
- (24) Levy, E. D.; Pereira-Leal, J. B.; Chothia, C.; Teichmann, S. A. 3D Complex: a Structural Classification of Protein Complexes. *PLoS Comput. Biol.* **2006**, *2* (11), e155.
- (25) Christendat, D.; Saridakis, V.; Kim, Y.; Kumar, P. A.; Xu, X.; Semesi, A.; Joachimiak, A.; Arrowsmith, C. H.; Edwards, A. M. The Crystal Structure of Hypothetical Protein MTH1491 From *Methanobacterium Thermoautotrophicum*. *Protein Sci.* **2002**, *11* (6), 1409–1414.
- (26) Galvagnion, C.; Smith, M. T. J.; Broom, A.; Vassall, K. A.; Meglei, G.; Gaspar, J. A.; Stathopoulos, P. B.; Cheyne, B.; Meiering, E. M. Folding and Association of Thermophilic Dimeric and Trimeric DsrEFH Proteins: Tm0979 and Mth1491. *Biochemistry* **2009**, *48* (13), 2891–2906.
- (27) Betcher-Lange, S. L.; Lehrer, S. S. Pyrene Excimer Fluorescence in Rabbit Skeletal  $\alpha$ Tropomyosin Labeled with *N*-(1-Pyrene)Maleimide. a Probe of Sulfhydryl Proximity and Local Chain Separation. *J. Biol. Chem.* **1978**, *253* (11), 3757–3760.
- (28) Niwayama, S.; Kassir, A. S.; Zhao, T.; Sutton, R. B.; Altenberg, G. A. A Pyrene Maleimide with a Flexible Linker for Sampling of Longer Inter-Thiol Distances by Excimer Formation. *PLoS ONE* **2011**, *6* (10), e26691.
- (29) Lee, D.; Greenman, L.; Sarovar, M.; Whaley, K. B. Ab Initio Calculation of Molecular Aggregation Effects: a Coumarin-343 Case Study. *J. Phys. Chem. A* **2013**, *117* (43), 11072–11085.
- (30) Liu, X.; Cole, J. M.; Low, K. S. Molecular Origins of Dye Aggregation and Complex Formation Effects in Coumarin 343. *J. Phys. Chem. C* **2013**, *117* (28), 14723–14730.
- (31) Garimella, P. D.; Datta, A.; Romanini, D. W.; Raymond, K. N.; Francis, M. B. Multivalent, High-Relaxivity MRI Contrast Agents Using Rigid Cysteine-Reactive Gadolinium Complexes. *J. Am. Chem. Soc.* **2011**, *133* (37), 14704–14709.
- (32) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Springer Verlag: New York, 2006.
- (33) Runnels, L. W.; Scarlata, S. F. Theory and Application of Fluorescence Homotransfer to Melittin Oligomerization. *Biophys. J.* **1995**, *69* (4), 1569–1583.
- (34) 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.
- (35) Warren, S.; Margineanu, A.; Katan, M.; Dunsby, C.; French, P. Homo-FRET Based Biosensors and Their Application to Multiplexed Imaging of Signalling Events in Live Cells. *Int. J. Mol. Sci.* **2015**, *16* (7), 14695–14716.
- (36) Papiz, M. Z.; Prince, S. M.; Howard, T.; Cogdell, R. J.; Isaacs, N. W. The Structure and Ther-

- mal Motion of the B800–850 LH2 Complex From *Rps. Acidophila* At 2.0 Å Resolution and 100 K: New Structural Features and Functionally Relevant Motions. *J. Mol. Biol.* **2003**, 326 (5), 1523–1538.
- (37) Sturgis, J. N.; Robert, B. The Role of Chromophore Coupling in Tuning the Spectral Properties of Peripheral Light-Harvesting Protein of Purple Bacteria. *Photosynth. Res.* **1996**, 50 (1), 5–10.
- (38) Lin, Q.; Bao, C.; Cheng, S.; Yang, Y.; Ji, W.; Zhu, L. Target-Activated Coumarin Phototriggers Specifically Switch on Fluorescence and Photocleavage Upon Bonding to Thiol-Bearing Protein. *J. Am. Chem. Soc.* **2012**, 134 (11), 5052–5055.
- (39) Connors, K. A. The Stability of Cyclodextrin Complexes in Solution. *Chem. Rev.* **1997**, 97 (5), 1325–1358.
- (40) Mertens, M. D.; Schmitz, J.; Horn, M.; Furtmann, N.; Bajorath, J.; Mareš, M.; Gütschow, M. A Coumarin-Labeled Vinyl Sulfone as Tripeptidomimetic Activity-Based Probe for Cysteine Cathepsins. *ChemBioChem* **2014**, 15 (7), 955–959.

## Chapter 2

### Development of general protein-based formulations for hydrophobic actives using hydrophobicity modulators

#### Abstract

Virus-like particles (VLPs) have gained widespread attention as potential nanocarriers for drug delivery and diagnostic imaging applications. The techniques used to formulate these nanocarriers generally rely on the cargo of interest containing specific chemical functionalities, or they require modification of the cargo prior to loading. In this chapter, efforts towards the development of a new VLP-based formulation for general hydrophobic actives will be discussed. Two new biocompatible bond cleavage reactions were investigated, one metal-induced and one photo-induced, and utilized to create a hydrophobic surface on the interior of the MS2 capsid. Efforts towards the preparation of hydrophobicity modulating polymers and polymer-protein hybrids, for the creation of a hydrophobic droplet within the MS2 capsid, will be discussed.

## 2.1 Introduction

### 2.1.1 Virus-like particles as drug delivery vehicles

Due to their biocompatibility, uniformity, and processability, virus-like particles (VLPs) have been studied extensively as protein-based nanocarrier formulations for targeted drug delivery.<sup>1-3</sup> Such nanocarriers have been prepared by loading cargo into the interior of the particle and by selective attachment of targeting groups to the exterior surface. The active payload has been incorporated into VLPs through conjugation, infusion, and encapsidation. Conjugation is by far the most common method of loading and involves the covalent attachment of cargo molecules to specific amino acids on the interior surface of the coat protein. Examples include conjugation of taxol via a maleimide-functionalized cleavable linker to the MS2 capsid,<sup>4</sup> conjugation of doxorubicin to VP6 through carbodiimide coupling,<sup>5</sup> and conjugation of doxorubicin via a cleavable maleimide linker to small heat shock protein (Hsp).<sup>6</sup> The conjugation method of loading suffers from two major limitations. First, the number of cargo molecules is limited by the number of targeted reactive residues per capsid (generally one per capsid protein monomer). Second, the cargo molecule must have a functional handle for the attachment of a bioconjugation reagent. Although not yet demonstrated for the loading of active drug molecules, inwardly-directed polymer-VLP conjugates with reactive handles have addressed the first of these limitations. Very high densities of fluorophores and gadolinium chelating agents (22 per monomer) for MRI contrast have been incorporated into P22-polymer hybrids.<sup>7</sup> Although these hybrid materials may resolve the cargo density limitations of standard covalent modification, the requirement for the modification of the cargo molecule for attachment persists.

Non-covalent loading strategies avoid the issue of cargo modification. Due to the presence of RNA and/or DNA in untreated VLPs, cationic cargo can be loaded by infusion and is retained within the VLP by electrostatic interaction with the anionic phosphate ester backbone of the RNA/DNA. This loading strategy has been used to load 1500 copies (8.3 per monomer) of the potent cytotoxic agent doxorubicin into cucumber mosaic virus (CMV) capsids,<sup>8</sup> up to 1270 copies (7 per monomer) of doxorubicin to red clover necrotic mosaic virus (RCNMV) capsids,<sup>9</sup> and 140 copies (1.2 per monomer) of the potential cytotoxic agent proflavine to cowpea mosaic virus (CPMV)<sup>10</sup>. Taking advantage of the natural self-assembly behavior of capsid proteins, VLPs can be disassembled and reassembled around suitable anionic macromolecules.<sup>11</sup> Assembly of hibiscus chlorotic ringspot virus (HCRSV) capsid proteins around poly(styrene sulfonate) pre-incubated with doxorubicin afforded capsids with 900 copies (5 copies per monomer) of the active drug.<sup>12</sup> Through infusion and encapsidation methods, significantly higher cargo densities can be achieved than through traditional conjugation strategies. However, these methods are severely limited by the requirement that the small molecule cargo must be charged.

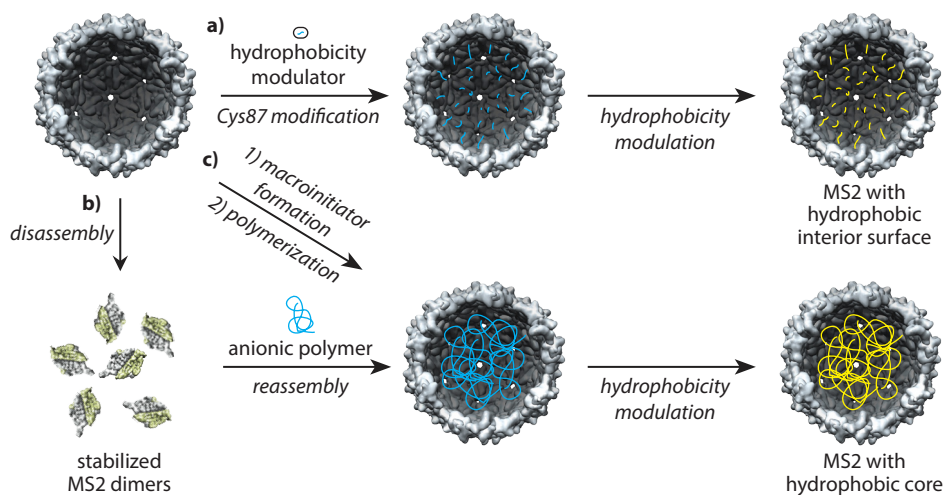
As many active drug molecules cannot be modified for bioconjugation without significant synthetic efforts or effecting their efficacy, and are not cationic, but are often relatively hydrophobic, the development of a general VLP-based formulation for the delivery of hydrophobic actives would be warranted. Kwak *et al.* have developed a proof of concept system that achieves hydrophobic cargo loading.<sup>13</sup> Using an encapsidation strategy, DNA amphiphile micelles pre-loaded with pyrene were used for the templation of cowpea chlorotic mottle virus (CCMV) capsids. Inspired by this step towards a more general VLP-based formulation, we sought to develop new protein nanocarriers for the delivery of hydrophobic actives (e.g., small molecule drugs, lipid-soluble vitamins, crop

protection agents, etc.) using the bacteriophage MS2 capsid. Progress towards the construction of these new materials will be discussed in this chapter.

## 2.1.2 Hydrophobicity modulation as a strategy to prepare general protein nanocarriers

As discussed above, the Francis group has utilized the MS2 capsid extensively for the preparation of nanocarriers using conjugation and encapsidation strategies. We envisioned that by modulating the hydrophobicity of the interior of the MS2 capsid, the MS2-based nanocarrier could be loaded with most any hydrophobic cargo through an infusion strategy. Three strategies were proposed for the preparation of MS2 particles with hydrophobic cores through the use of a switchable hydrophobicity modulating compounds (**Figure 2.1**): 1) increasing the interior surface hydrophobicity using a bioconjugation reagent capable of modulating hydrophobicity (Section 2.2), 2) generation of a hydrophobic droplet within the particle through the encapsidation of hydrophobicity modulating polymers (Section 2.3), and 3) incorporation of a hydrophobic corona or droplet by inwardly grafting a hydrophobicity modulating polymer from the MS2 surface (Section 2.4). The success of each of these strategies depended on the development of a biocompatible bond cleavage reaction that results in a dramatic increase in local hydrophobicity.

While the toolbox of protein bioconjugation reactions contains numerous robust methods for the selective modification of specific amino acid residues and is being continuously expanded, the development of a toolbox of biocompatible bond cleavage reactions is still in early stages.<sup>14</sup> Even with this limited toolbox, biocompatible bond cleavage reactions have been used primarily in the field of chemical biology with applications ranging from chemically controlled spatial and temporal activation of proteins, to the release of small-molecule drugs, to direct manipulation of intact cells.<sup>15</sup> In tandem, the biocompatible bond formation and bond cleavage toolboxes allow for the



**Figure 2.1.** Creating a general MS2-based nanocarrier for hydrophobic actives. Three strategies were proposed for the creation of a hydrophobically-modified MS2 capsid. a) Modification of cysteines on the interior surface of MS2 with a small molecule hydrophobicity modulating reagent. b) Encapsidation of hydrophobicity modulating polymer using a disassembly-reassembly procedure. c) Grafting of a hydrophobicity modulating polymer from the interior surface of the MS2 capsid. In each strategy a biocompatible bond cleavage reaction capable of increasing local hydrophobicity would be utilized to modulate hydrophobicity. Capsids with hydrophobic interiors could then be loaded with hydrophobic actives by diffusion.

preparation of complex and well-defined protein-based materials, which are not accessible through bond forming reactions alone.

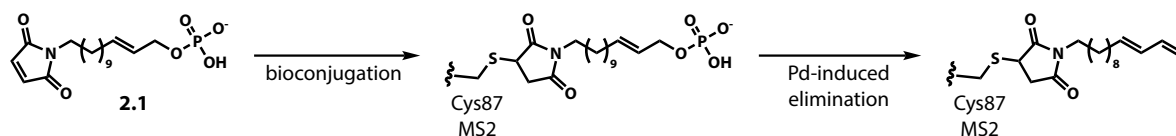
Currently, biocompatible bond cleavage reactions fall primarily under the class of deprotection or uncaging chemistry and result in the exposure of a heteroatom. Established methods can be divided into five classes: reduction-induced, pH-induced, photo-induced, metal-induced, and small molecule-induced bond cleavage. Disulfide-containing molecules introduced to proteins via bioconjugation can be reduced with dithiothreitol (DTT) or tris(2-carboxyphenyl)phosphine (TCEP) to uncage a free thiol.<sup>14</sup> The Staudinger reduction of azides affords free amines.<sup>16</sup> Azo compounds can be reductively cleaved under mild conditions to expose reactive anilines.<sup>17</sup> Through pH-modulation, protecting group removal and release of small molecules can be achieved. Hydrolytically labile groups, most notably the hydrazone linkage, allow for cargo release at low pH.<sup>18</sup> Cargo release can also be achieved at low pH using a thiomaleamic acid linker, which is cleaved through cyclization to the thiomaleic anhydride.<sup>19</sup> Under alkaline treatment, thioacetates can be deprotected to expose free thiols.<sup>14</sup> Of the biocompatible photo-induced bond cleavage reactions, UV irradiation of an *o*-nitrobenzyl carbamate-protected amines is by far the most commonly used.<sup>20</sup> Allyloxycarbonyl (Alloc)-protected,<sup>21</sup> propargyloxycarbonyl (Poc)-protected,<sup>22</sup> and tertiary propargyl carbamate (Ynoc)-protected amines<sup>23</sup> have been uncaged under biological conditions with Ru, Pd, and Cu catalysts, respectively. Under the class of small molecule-induced cleavage reactions, *trans*-cyclooctene caged lysines have been uncaged by the addition of tetrazines.<sup>24</sup> Aryl azides have been converted to anilines by reaction with *trans*-cyclooctene via hydrolysis of the rearranged cycloaddition product.<sup>25</sup>

While the currently available biocompatible bond cleavage reactions could be used to decrease local hydrophilicity through the removal of a strong water-solubilizing group, a polar heteroatom functionality would remain. In order to achieve a truly significant switch from hydrophilic to hydrophobic character, new biocompatible bond cleavage reactions would need to be utilized. In this work, a metal-induced and a photo-induced bond cleavage reaction, new to the bioconjugation toolbox, were examined for their ability to unmask heteroatom-free functionalities for hydrophobicity modulation.

## 2.2 Interior surface hydrophobicity modulation of MS2

### 2.2.1 Synthesis of thiol-reactive $\pi$ -allylpalladium substrate

The first new biocompatible bond cleavage reaction examined for use in hydrophobicity modulation was inspired by the Francis group's previous work towards selective tyrosine modification. Tilley and Francis found that tyrosine residues could be selectively alkylated with  $\pi$ -allylpalladium complexes.<sup>26</sup> By modifying a farnesyl group with a highly water-soluble allyl taurine carbamate, synthetic lipoproteins could be readily prepared through solubility switching of the reagent upon conjugation. Tilley also reported a previously unobserved side reaction for aqueous  $\pi$ -allylpalladium reactions, which led to the formation of a terminal diene via  $\beta$ -elimination of the  $\pi$ -allylpalladium complex. Although this eliminative pathway was not previously observed in aqueous systems, it is a known pathway in organic solvent.<sup>27</sup> By preparing a bifunctional variant of Tilley's lipoprotein forming reagent, hydrophobicity modulation could be achieved in a two-step conjugation-cleavage procedure. Although Tilley's reagent allows for the direct incorporation of a hydrophobic moiety, it is limited to modification of tyrosine. The proposed bifunctional hydro-



**Figure 2.2.** Envisioned surface hydrophobicity modulation using  $\pi$ -allylpalladium chemistry. Cys87 on the interior surface of MS2 would first be modified with hydrophobicity modulator **2.1**. Treatment with a water-soluble palladium source should induce eliminative bond cleavage to afford a hydrophobic terminal diene.

phobicity modulator could be modified for selective labeling of any of the reactive amino acids. Although not the focus of the current work, a two-step process would allow for increased temporal control over incorporation of synthetic lipoproteins into lipid bilayers.

Towards the development of a thiol-reactive palladium-switchable hydrophobicity modulator, maleimide allyl phosphate **2.1** was chosen as a target (**Figure 2.2**). A maleimide was chosen as the thiol-reactive group, as maleimides are relatively specific agents for labeling cysteines. This is especially true in the case of Cys87 labeling in MS2, wherein side-reactivity with maleimide-based agents is known to be quite limited.<sup>28</sup> For the cleavable end of the reagent, an allyl phosphate was chosen. Although both cationic groups (quaternary ammonium salts or amines under biological conditions) or anionic groups could provide the needed water solubility for this reagent, an anionic group was ultimately chosen for three reasons. The first is that the interior surface of the MS2 capsid is positively charged, which could result in an enhanced rate of thiol-maleimide coupling due to increased proximity afforded by electrostatic interactions between the capsid surface and the anionic solubilizing group. As thiol-maleimide reactions are known to be quite rapid, this effect may be quite small in improving the rate of reaction in this system. However, this reagent may tend to form micelles due to its surfactant-like qualities and any boost in reactivity could help drive the desired transformation towards completion. Secondly and perhaps more importantly, the palladium chemistry used in the hydrophobicity switch has been demonstrated with an anionic leaving group. Introduction of an amine or ammonium salt may alter the reactivity of the  $\pi$ -allylpalladium complex or prevent its formation in the first place through direct palladium binding through the nitrogen. The third and final reason, as will be discussed in Sections 2.3 and 2.4, is that the development of a polymeric analog of the reagent under current discussion is of great interest. Such a polymer could be used in disassembly-reassembly procedure followed by end-group cleavage, to result in a hydrophobic droplet within the capsid. As previously discussed, the main requirement of a macromolecule for successful encapsidation is that it has significant anionic character, such as that provided by phosphate monoesters. Although a very lengthy alkyl linker would likely lead to the greatest increase in surface hydrophobicity (after triggered hydrophobicity switching), the actual length of the linker was decided upon as a compromise between minimizing issues of water-solubility prior to cleavage, and availability and cost of suitable starting materials for the synthesis of such a bifunctional reagent.

Synthesis of bifunctional agent **2.1** began with commercially available 11-bromoundecanol, where the bromide would serve as a handle for the insertion of a maleimide, and the alcohol would be transformed into the terminal allyl phosphate (**Figure 2.3**). A three-stage strategy was used to convert the bifunctional starting material into the desired reagent. In the first stage, the primary alcohol would be converted to a terminal allylic alcohol, thus providing the desired carbon skeleton for the reagent. The thiol-reactive moiety would be installed in the second stage and, finally, the terminal allylic alcohol would be phosphorylated in the third stage, imparting water solubility to the reagent.



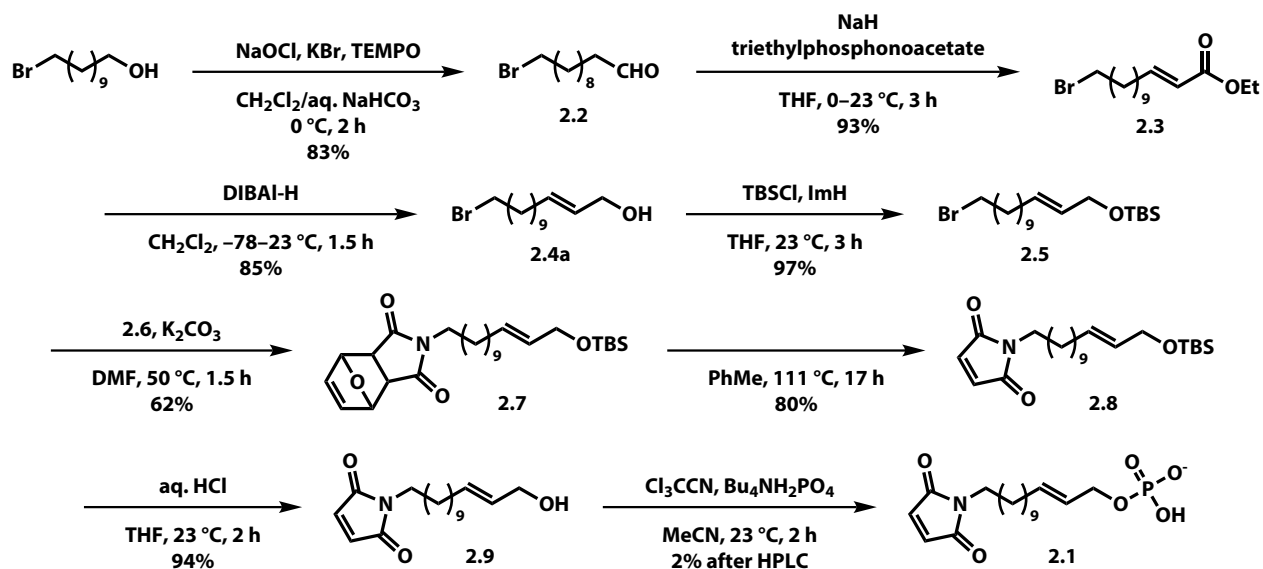


Figure 2.3. Synthesis of maleimide allyl phosphate 2.1.

The first stage of the synthesis started with the oxidation of 11-bromoundecanol with 2,2,6,6-tetramethyl-1-piperidinyloxy radical (TEMPO) and sodium hypochlorite under biphasic conditions to afford the corresponding aldehyde **2.2** in very good yield. Aldehyde **2.2** was then transformed to  $\alpha,\beta$ -unsaturated ester **2.3** in excellent yield by a Horner-Wadsworth-Emmons reaction with the sodium salt of triethyl phosphonoacetate.  $\alpha,\beta$ -Unsaturated ester **2.3** was then reduced with diisobutylaluminum hydride (DIBAL-H) to afford the corresponding allylic alcohol **2.4a** in very good yield, completing the central carbon scaffold of the bifunctional reagent and finalizing the first stage of its synthesis. On initial testing of this reaction, the reduction of  $\alpha,\beta$ -unsaturated ester **2.3** with 3 equivalents of DIBAL-H at –78 °C for 2 h afforded pure allylic alcohol **2.4a** after chromatography. However, upon scale-up, it was found that the same conditions resulted in over-reduction of a portion of the material to the fully saturated alcohol **2.4b**. The over-reduced material co-eluted with the desired allylic alcohol under all normal-phase chromatography conditions that were screened. A mixture of desired allylic alcohol **2.4a** containing roughly 5% of over-reduced alcohol **2.4b** was therefore carried forward and could be separated at a later stage. The reaction was later optimized to minimize reduction of the alkene. It was found that over-reduction could be avoided by decreasing the amount of DIBAL-H used to 2.2 equivalents, and by shortening the reaction time to 1.5 h.

With the completed scaffold of the hydrophobicity modulator in place, the thiol-reactive maleimide could then be installed. Allylic alcohol **2.4a** was first protected with *tert*-butyl dimethylsilyl chloride (TBSCl) to provide allylic silyl ether **2.5** in near quantitative yield. The primary bromide in **2.5** was then displaced with furan-maleimide adduct **2.6** under basic conditions to afford protected maleimide **2.7** in good yield. Retro-Diels-Alder deprotection of furan-maleimide adduct **2.7** was accomplished by refluxing the material in toluene for 17 h, affording maleimide **2.8** in very good yield and completing the second stage of the synthesis.

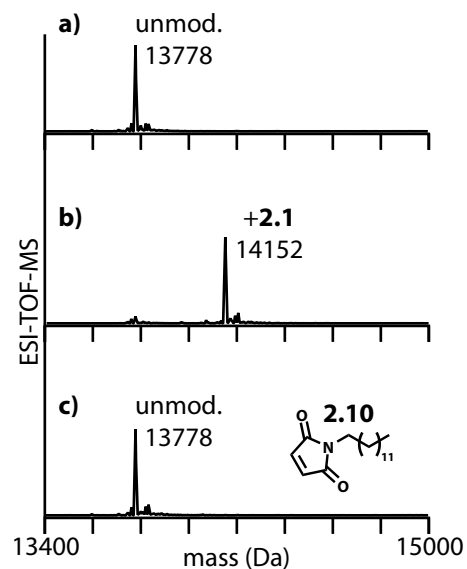
The final stage of the synthesis consisted of the deprotection of the allylic silyl ether and the installation of the water-solubilizing phosphate group. Treatment of allylic silyl ether **2.8** with 1:2 1 M aqueous HCl in THF afforded free allylic alcohol **2.9** in excellent yield. At this stage, allylic alcohol **2.9** became chromatographically distinct from the corresponding fully saturated alcohol

remaining from over-reduction of  $\alpha,\beta$ -unsaturated ester **2.3**. Allylic alcohol **2.9** was treated with phosphorylated trichloroacetimidate, formed *in situ* through the condensation of trichloroacetonitrile and tetrabutylammonium hydrogen phosphate, to afford the desired phosphate monoester **2.1** in moderate yield after purification by ion-exchange. Prior to use in bioconjugation reactions, the reagent was further purified by reversed-phase high-performance liquid chromatography (HPLC). Overall, maleimide allyl phosphate reagent **2.1** was synthesized in 8 linear steps from 11-bromo undecanol for a total yield of 13%, prior to the final HPLC purification.

## 2.2.2 Conjugation and hydrophobicity modulation with allyl phosphate **2.1**

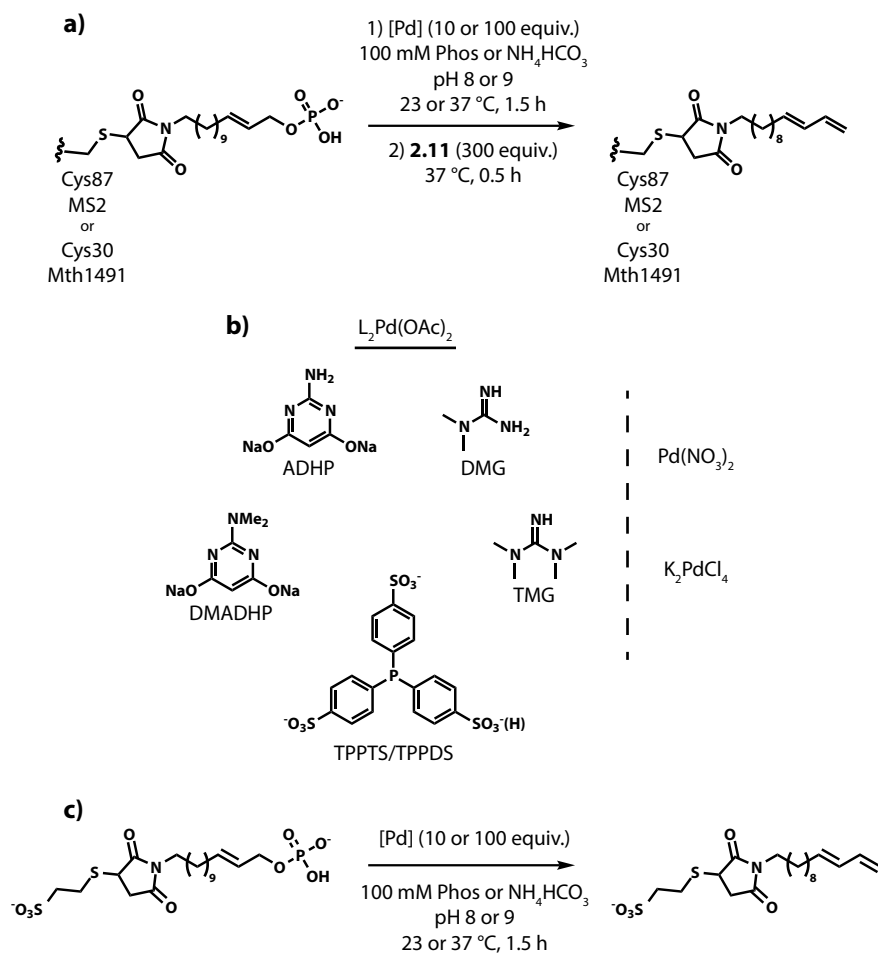
The conjugation of maleimide allyl phosphate **2.1** to the MS2 capsid protein at residue Cys87 was optimized with the intent of generating an MS2 construct with a hydrophobic interior surface. It was found that excellent labeling of MS2 could be achieved with 80 equivalents of maleimide **2.1** at pH 7.2 in phosphate buffer over 1 h at room temperature (**Figure 2.4**). For comparison, MS2 was treated with *N*-tridecylmaleimide **2.10**, prepared by a Mitsunobu reaction between maleimide and 1-tridecanol. Under the same coupling conditions, no modification of the capsid protein was observed. Increasing the concentration of DMF cosolvent to 10% still did not result in modification using this reagent. This experiment clearly demonstrated the necessity of the solubilizing group for the attachment of hydrophobic cargo to MS2, and validated the development of a switchable hydrophobicity modulator.

With maleimide allyl phosphate-labeled MS2 capsids prepared, the palladium-triggered cleavage could be investigated. Initially, conditions found to result in  $\beta$ -elimination in Tilley's tyrosine modification chemistry were used to attempt the desired cleavage.<sup>26</sup> Allyl phosphate-



**Figure 2.4.** ESI-TOF-MS characterization of MS2 Cys87 modified with hydrophobicity modulator **2.1**. MS2 a) before and b) after treatment with maleimide **2.1**. c) Cys87 could not be modified with *N*-tridecylmaleimide due to poor solubility of the reagent (expected: 14151 Da).

labeled MS2 was treated with a catalyst mixture composed of 1:12 Pd(OAc)<sub>2</sub>:triphenylphosphine trisulfonate (TPPTS) in phosphate buffer at pH 8.6. After standard work-up by repeated centrifugal ultrafiltration, very low intensity signal was observed by ESI-TOF-MS, even though the protein sample was of sufficient concentration as determined by absorbance at 280 nm, to obtain quality MS data under standard conditions. Additionally, significant levels of palladium adducts were observed by MS. Desalting was also attempted using Sephadex G-25 resin, however ESI-TOF-MS analysis still revealed suppressed ionization and palladium adducts. In order to obtain meaningful analytical data for the MS2 construct after attempted cleavage, it was essential to remove all of the palladium catalyst. Palladium scavenging was attempted with EDTA with minimal improvements. The Davis lab has reported similar issues of ESI-TOF-MS signal suppression while developing methodologies to carry out Suzuki-Miyaura couplings on proteins, and has developed effective palladium scavengers.<sup>29</sup> 3-Mercaptopropionic acid (3-MPrAc) was found to be highly effective at scavenging palladium from the MS2 constructs after attempted cleavage and facilitated ESI-TOF-MS analysis. However, 3-MPrAc was also found to



**Figure 2.5.** Screening conditions for palladium-induced hydrophobicity modulation of **2.1** conjugates. a) MS2- or Mth1491-**2.1** conjugates were treated with 10 or 100 equivalents of a b) Pd source, with or without ligand, under a variety of conditions. Pd was removed by treatment with scavenger **2.11**. No conversion to desired cleaved product was observed by ESI-TOF-MS. c) **2.1**-MESNa was not observed to undergo cleavage by LC-MS.

After removing palladium from the MS2 conjugate that had been exposed to the initial cleavage conditions, it was discovered that the expected terminal diene had not been formed and that the cleavage was unsuccessful, as determined by ESI-TOF-MS. Based on other conditions that Tilley had screened, the cleavage was reattempted in both triethylammonium carbonate and phosphate buffer at pH 8.0 and 9.0, and the reaction was carried out either at room temperature or 37 °C. Again, the starting material remained intact under these conditions. As the initially-chosen palladium catalyst mixture did not afford the desired cleavage, a screen of water soluble palladium sources and ligands was conducted (**Figure 2.5**). A total of three commercially available and air-stable palladium sources were surveyed:  $\text{Pd}(\text{OAc})_2$ ,  $\text{Pd}(\text{NO}_3)_2$ , and  $\text{K}_2\text{PdCl}_4$ . In addition to TPPTS, triphenylphosphine disulfonate (TPPDS), 2-amino-4,6-dihydropyrimidine (ADHP), 2-dimethylamino-4,6-dihydropyrimidine (DMADHP), 1,1-dimethylguanidine (DMG), and 1,1,3,3-tetramethylguanidine (TMG) were chosen as ligands for screening. Palladium(II) acetate was screened both with and without addition of ligand while  $\text{Pd}(\text{NO}_3)_2$  and  $\text{K}_2\text{PdCl}_4$  were used without ligands. A twelve-fold excess of ligand was used in ligand-containing catalyst mixtures. Relative to modified

be a highly efficient partner for thiol exchange, effectively reversing the modification of maleimide allyl phosphate-MS2 to afford MS2 with a free cysteine.<sup>30</sup> A dimeric form of 3-MPrAc (**2.11**) was prepared by symmetric bis-thioether formation between 3-MPrAc and 1,3-dibromopropane.<sup>29</sup> This dimer was also effective in removing palladium from the MS2 conjugate. As the dimeric scavenger did not contain any free thiols, thiol exchange between the scavenger and the cysteine on MS2 was no longer possible and the MS2 conjugate remained intact after purification. Purification with dimeric 3-MPrAc palladium scavenger **2.11**, in concert with centrifugal ultrafiltration, minimized ESI-TOF-MS signal suppression and allowed for straightforward analysis of palladium-catalyzed cleavage reactions.

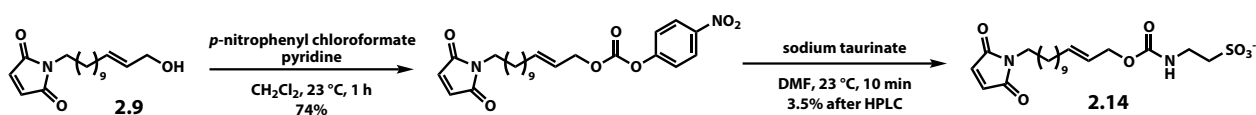
protein, either 10 or 100 equivalents of palladium were used and all reactions were carried out for 1.5 h at 37 °C in phosphate buffer at pH 8.0 or 9.0. Unfortunately, the desired  $\beta$ -elimination product was not detected for any of the conditions screened.

As no chemistries involving palladium-based reagents have been previously attempted on the interior surface of MS2, it was thought that the palladium catalysts may not be diffusing through the pores of the MS2 capsid. To test whether the desired cleavage chemistry would work on a more accessible protein surface, a model system was prepared using the Mth1491 scaffold. Maleimide allyl phosphate **2.1** was coupled to the cysteine of the Mth1491 A30C mutant, in high yield, by reaction with 80 equivalents of the reagent at pH 6.5. The full screen of palladium sources, ligands, and reaction conditions was carried out using this model system. Yet again, no cleavage was observed.

Given that all of the screened conditions produced negative results, the quality of the palladium catalysts was questioned. The activity of the catalyst mixtures was assayed using a different Mth1491-based model system, this time using well-known and proven Suzuki-Miyaura chemistry.<sup>31</sup> Mth1491 A30C was labeled quantitatively with 4-iodophenethylmaleimide (**2.12**). Roughly 10% double and triple modification was observed, suggesting that a portion of the protein sample no longer retained its native fold. However, the highest quality of the conjugate was not needed in this study, and was used as prepared. Coupling between the iodophenyl-labeled conjugate and phenylboronic acid was attempted with 20 equivalents of 1:12 Pd(OAc)<sub>2</sub>:DMG at pH 8.0 for 1 h at 37 °C. Analysis by ESI-TOF-MS revealed that coupling had occurred but that the major product was that from protodeiodination. Although the expected biphenyl product of the Suzuki-Miyaura reaction was found to be the minor product, these results demonstrated that the palladium catalyst mixture was indeed active.

In a final attempt to achieve successful cleavage, the  $\pi$ -allylpalladium chemistry was tested directly on the small molecule allyl phosphate **2.1**. Prior to screening, maleimide allyl phosphate **2.1** was first treated with 2-mercaptoethanesulfonate sodium salt (MESNa), to afford thioether **2.12**. Tagging with MESNa ensured that the cleaved product would remain water-soluble and allows for monitoring by LC-MS in negative mode. The screen of reaction conditions was repeated, yet no cleavage was detected.

With the failure of the palladium-catalyzed cleavage on the MS2 conjugate, the model Mth1491 conjugate, and the small molecule allyl phosphate **2.13**, it became very clear that the allyl phosphate was not suitable for this transformation. To more closely match the substrates used by Tilley *et al.*, the taurine carbamate **2.14** was prepared through activation of alcohol **2.9** with 4-nitrophenyl chloroformate followed by displacement with sodium taurinate (*Figure 2.6*). The maleimide of **2.14** was then reacted with MESNa and treated with a subset of the palladium sources and ligands screened above. Analysis of the reaction mixtures by negative mode LC-MS revealed that, in most cases, at least some of the allylic carbamate had been converted back to the corresponding alcohol, presumably via hydrolysis of the  $\pi$ -allylpalladium complex. However, not even a trace of terminal diene consistent with  $\beta$ -elimination of the  $\pi$ -allyl complex was observed.



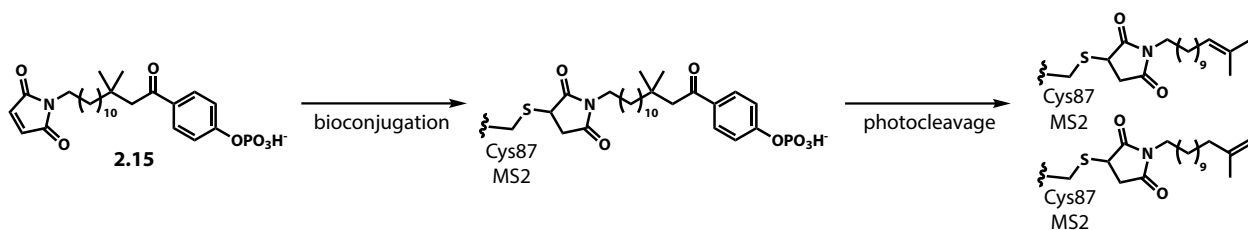
**Figure 2.6.** Synthesis of maleimide allyl taurine carbamate **2.14**. MESNa-**2.14** was not found to undergo  $\beta$ -elimination with Pd treatment, however partial hydrolysis to MESNa-**2.9** was observed by LC-MS.

While a maleimide allyl phosphate designed specifically for cysteine labeling and hydrophobicity modulation was synthesized and shown to effectively label the Cys87 on the interior surface of the MS2 capsid, it was found that the eliminative  $\pi$ -allylpalladium chemistry was unsuccessful. The question remains why this chemistry was ultimately fruitless on the current substrates. The finding that allyl taurine carbamate **2.14** underwent hydrolysis whereas allyl phosphate **2.1** remained intact under all conditions of the screen suggests that allyl phosphates are not great substrates for  $\pi$ -allylpalladium complex formation. Although strong precedent exists for the formation of dienes through a  $\beta$ -elimination mechanism from  $\pi$ -allyl complexes in organic solvent,<sup>27</sup> the best evidence of this transformation occurring under aqueous conditions is presented in Tilley's work on tyrosine modification using Tsuji-Trost chemistry.<sup>26</sup> In Tilley's studies, confirmation of diene formation was achieved through HPLC analysis of an unknown side product, presumed to be the diene, and an authentic sample. Based on the negative results in the current study, it seems quite possible that the similar HPLC retention times of presumed diene and authentic diene in the Tilley study may have been coincidental. Although  $\pi$ -allylpalladium chemistry offers an attractive bio-orthogonal methodology for hydrophobicity modulation, the possibility of a  $\beta$ -elimination pathway occurring under aqueous conditions must first be irrefutably confirmed.

### 2.2.3 Synthesis of a photo-induced hydrophobicity modulator

As conditions for effecting a hydrophobicity switch through  $\pi$ -allylpalladium chemistry were not found, efforts were turned to the development of a photoswitchable variant of the hydrophobicity modulating agent. Photocleavable protecting groups and caging groups generally reveal hydrophilic heteroatom containing functionalities after cleavage.<sup>20</sup> One of the few photocleavage reactions that does not reveal a heteroatom is the Norrish type II cleavage of aryl alkyl ketones (**Figure 2.7**).<sup>32</sup> In Norrish type II photocleavage, irradiation with photons around 300 nm results in excitation of the ketone to the 1,2-diradical. Intramolecular abstraction of a  $\gamma$ -hydrogen produces a 1,4-diradical as the primary photoproduct. Fragmentation of the 1,4-diradical results in an alkene and an enol, which tautomerizes to the methyl aryl ketone in aqueous solution. The 1,4-radical can also undergo Yang cyclization to form the cyclobutanol product.<sup>33</sup> The cyclization product can be avoided by incorporating gem-dimethyl groups  $\beta$  to the carbonyl.<sup>34</sup> The biocompatibility of this reaction has been demonstrated through the use of a photodegradable detergent for protein purification and characterization.<sup>35</sup> More recently, the utility of this reaction has been demonstrated for the light-induced controlled release of fragrance molecules.<sup>36</sup>

Similar considerations were given to the design of a thiol-reactive reagent (**2.15**) capable of undergoing photo-induced hydrophobicity switching as were given for the design of allyl phosphate



**Figure 2.7.** Envisioned surface hydrophobicity modulation using Norrish type II photochemistry. Cys87 on the interior surface of MS2 would first be modified with hydrophobicity modulator **2.15**. UVB irradiation of the conjugate should afford the hydrophobic disubstituted and trisubstituted alkene cleavage products through a radical mechanism.

2.1. Such a reagent would require 1) a thiol-reactive group, 2) a moiety suitable for Norrish type II photocleavage, and 3) a suitably long alkyl linker separating the two reactive groups. Again, a maleimide was chosen as the cysteine-targeting group and, again, a linker was chosen so that a C13 chain would remain after cleavage. As for the photoreactive moiety, a water-soluble alkyl phenyl ketone, derivatized with an anionic solubilizing group on the phenyl ring was required. Although an *ortho*-sulfophenyl ketone has been demonstrated to undergo Norrish type II photocleavage, the chemistry required to install the sulfonate is not particularly compatible with the maleimide end of the desired reagent or its precursors.<sup>35</sup> Instead, a phosphorylated *ortho*-hydroxyphenylketone was chosen, as it would provide the desired water solubility to the reagent, and because the synthesis of this group is compatible with chemistry required to insert the maleimide onto the opposite end of the reagent.

To initiate the synthesis of **2.15**, 11-bromoundecanol was protected with TBSCl to provide silyl ether **2.16** in excellent yield (Figure 2.8).  $\alpha,\beta$ -Unsaturated ketone **2.17** was prepared through Friedel-Crafts acylation of anisole in moderate yield, and then reacted with the Grignard reagent of alkyl halide **2.16** under copper-catalyzed conditions to afford the product of conjugate addition (**2.18**) in good yield. This conjugate addition completed the carbon skeleton of the bifunctional reagent, including the gem-dimethyl groups as inhibitors of Yang cyclization.

In order to deprotect the methyl phenyl ether and install the phosphate, a deprotection-reprotection sequence at the opposite end of the substrate was required. Removal of the silyl ether from **2.18**, which was required for protection during the Grignard coupling, was accomplished in near quantitative yield with 20% 1 M aqueous HCl in THF and the resultant alcohol **2.19** was then

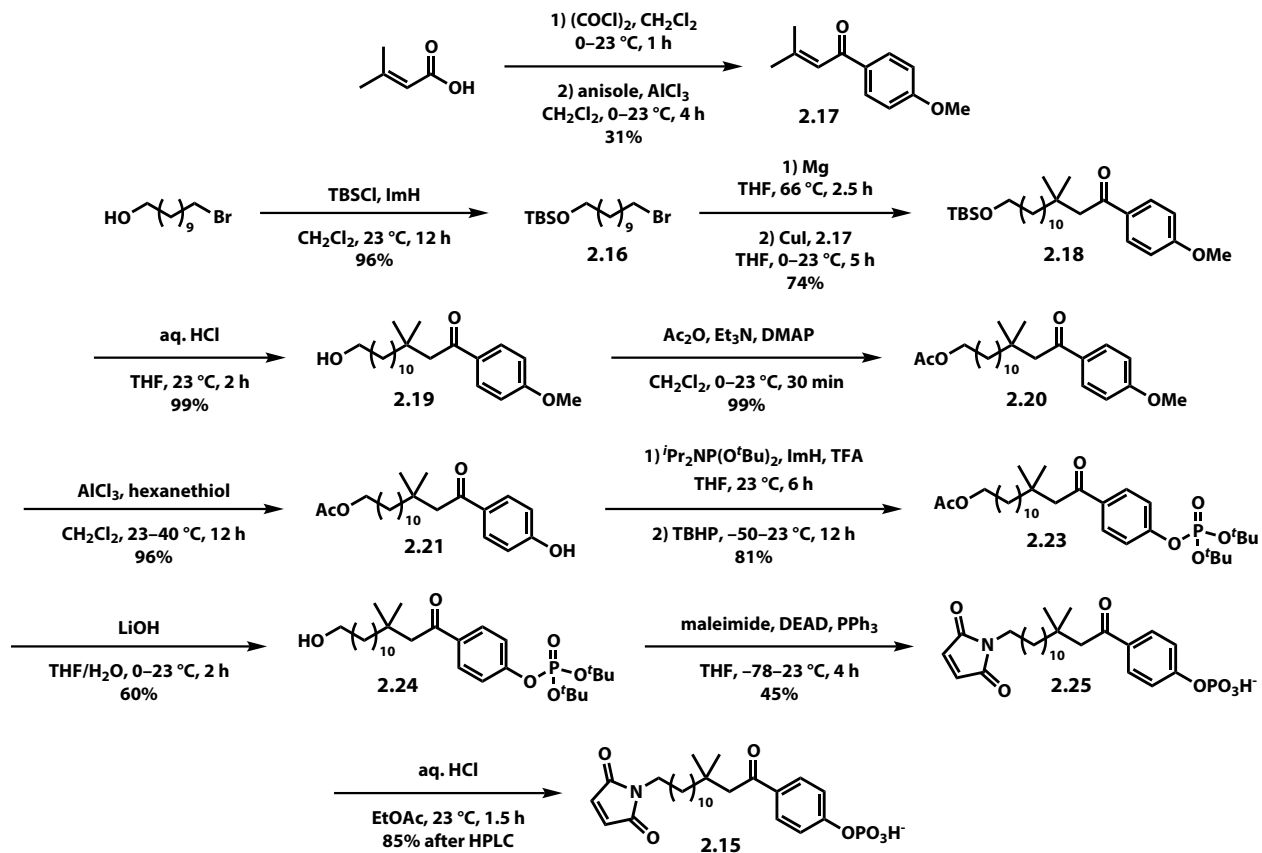


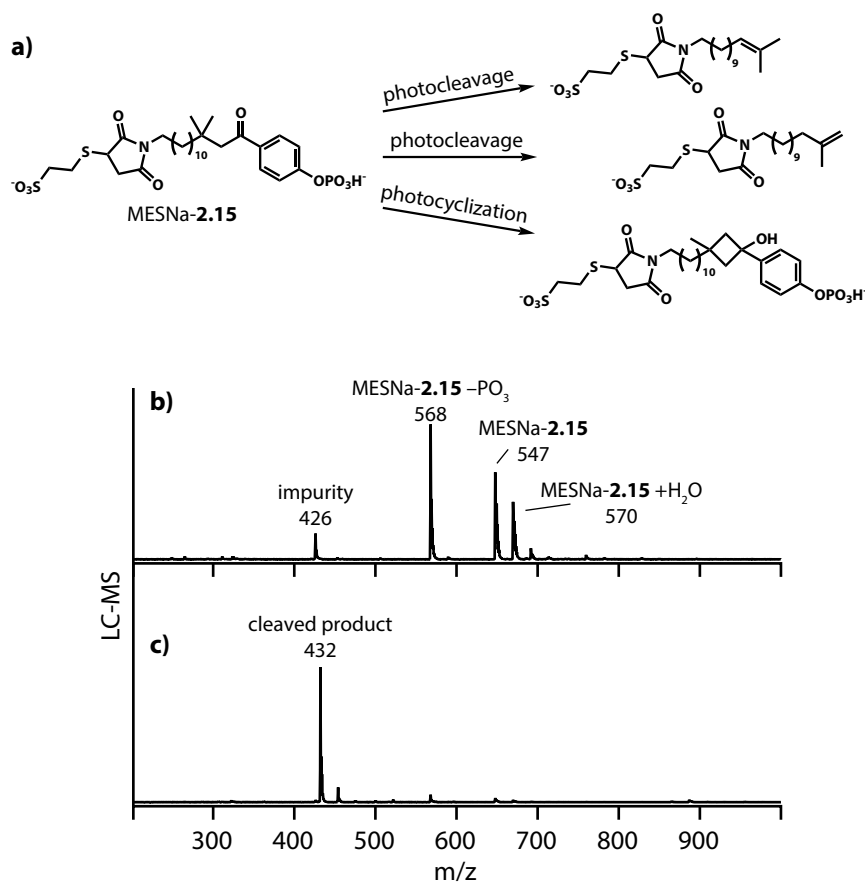
Figure 2.8. Synthesis of photocleavable reagent **2.15**.

reprotected by acetylation with acetic anhydride, again in near quantitative yield. With the primary alcohol suitably protected against conditions for demethylation of methyl phenyl ether **2.20**, a variety of conditions were screened for the preparation of phenol **2.21**. It was found that with  $\text{BBr}_3$  at low temperatures, deprotection of the methyl ether was not accomplished, and that higher temperatures resulted in global deprotection of the substrate. However, a high-yielding and selective deprotection was accomplished with  $\text{AlCl}_3$  and hexanethiol in refluxing methylene chloride.<sup>37</sup> The phosphate functionality was then installed through phosphoramidite chemistry. Phenol **2.21** was treated with di-*tert*-butyl *N,N*-diisopropylphosphoramidite, activated with imidazolium trifluoroacetate,<sup>38</sup> to afford phosphite **2.22**. Although intermediate phosphite **2.22** was isolable, it was found that oxidation of the crude material with *tert*-butyl hydroperoxide proceeded without issue. However, it was discovered that the resultant phosphate triester **2.23** was chromatographically indistinguishable from starting phenol **2.21** under normal-phase conditions on silica. Although it would be possible to separate unreacted phenol **2.21** at a later stage in the synthesis, it was deemed preferable to eliminate the problem at this stage. This problem was solved by driving the phosphite forming reaction to completion using 1.5 equivalents of the phosphoramidite reagent and extending the reaction time to 6 h. Oxidation was commenced only after all of the phenol had been consumed, as determined by thin layer chromatography.

With the phosphate ester in place, the thiol-reactive end of the reagent could then be installed. Hydrolysis of acetate **2.23** was achieved with  $\text{LiOH}$  in  $\text{THF}/\text{H}_2\text{O}$  to afford primary alcohol **2.24**. It was important to monitor this reaction closely as the aryl O-P bond is fairly labile under alkaline aqueous conditions. The primary alcohol was then substituted with maleimide through a Mitsunobu reaction using Walker's conditions.<sup>39</sup> Finally, di-*tert*-butylphosphate **2.25** was deprotected with a minimal amount of concentrated aqueous  $\text{HCl}$  in  $\text{EtOAc}$ , and purified by reversed-phase HPLC to afford desired bifunctional photocleavable reagent **2.15** in very good yield. As acid-catalyzed hydrolysis of the maleimide was found to occur to a greater extent under more pressing conditions, the milder deprotection conditions used did not result in complete conversion of phosphate triester **2.25** to phosphate monoester **2.15**. Overall, maleimide phenyl alkyl ketone reagent **2.15** was synthesized in 9 linear steps from 11-bromoundecanol in a total yield of 12%.

#### 2.2.4 Hydrophobicity modulation with photocleavable substrate **2.15**

Prior to attempting a conjugation of the photocleavable hydrophobicity modulator to the MS2 capsid, the Norrish type II reactivity of phenyl ketone **2.15** was first probed (**Figure 2.9**). Maleimide phenyl ketone **2.15** was treated with MESNa and the resulting thioether was purified by HPLC. As discussed above, the MESNa thioether imparts water solubility to the otherwise insoluble product of photocleavage, and allows for straightforward reaction monitoring by LC-MS. Characterization of phenyl ketone **2.15** by UV-Vis spectrophotometry revealed a  $\lambda_{\text{max}}$  of 274 nm. Although optimal cleavage rates would be achieved by irradiating with a spectral window centered around  $\lambda_{\text{max}}$ , irradiation was red-shifted to a central value of 300 nm to minimize photooxidation of photosensitive residues in protein conjugates. A sample of MESNa-tagged **2.15** in phosphate buffer at pH 7.2 was irradiated with  $300 \pm 10$  nm light from a xenon arc lamp for 1 h and then analyzed by LC-MS. After irradiation, the peak in the chromatogram corresponding to the starting material had all but disappeared, and one new peak consistent with a more hydrophobic compound was observed. Under these conditions, the Norrish type II chemistry was found to convert roughly 95% of the material to the photocleaved product. The observed masses corresponding to the starting phenyl ketone could



**Figure 2.9.** Photocleavage of 2.15-MESNa. a) Irradiation of 2.15-MESNa could result in the formation of disubstituted or trisubstituted alkenes through Norrish type II photocleavage or a cyclobutanol through Norrish-Yang cyclization. LC-MS analysis of 2.15-MESNa b) before and c) after irradiation at  $300\pm 10$  nm for 1 h revealed near quantitative conversion to the photocleaved products.

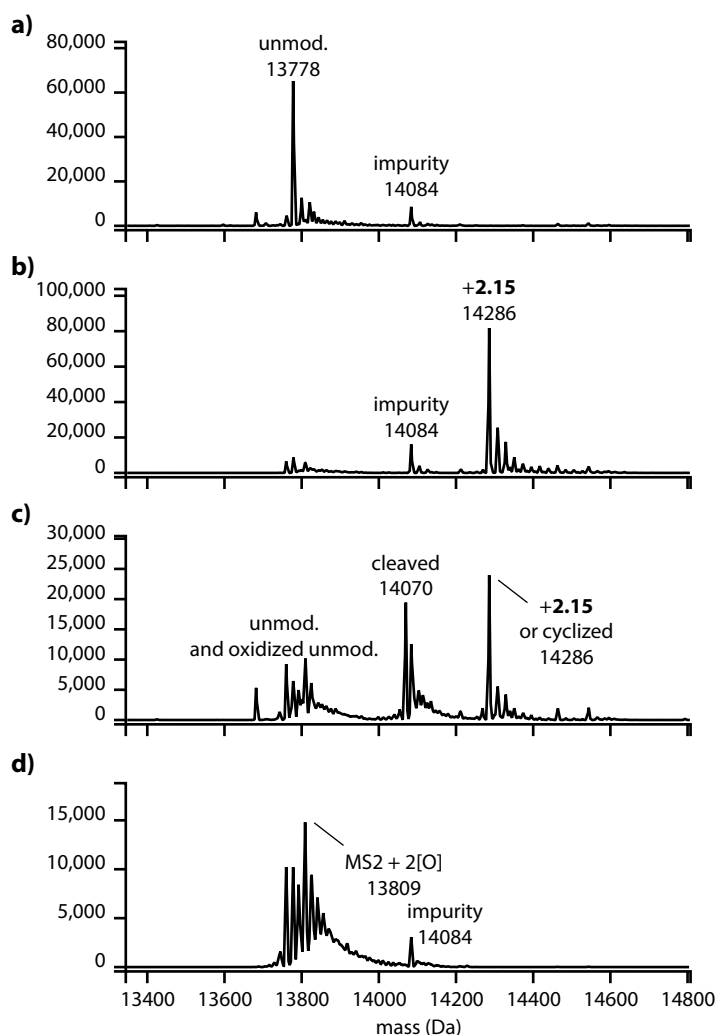
It was found that high levels of modification could be achieved through the reaction of MS2 with 40 equivalents of hydrophobicity modulator **2.15** at room temperature in pH 7.2 phosphate buffer (**Figure 2.10**). The reaction was found to be sluggish compared to other cysteine-maleimide couplings on MS2, perhaps due to a tendency of the reagent to form micelles in aqueous solution. Nevertheless, conjugation efficiencies of  $>80\%$  were achieved after 4 h, as determined by ESI-TOF-MS characterization. Analysis of the conjugate material revealed an SEC elution profile consistent with intact capsids. Likewise, particle diameters derived from DLS were consistent with intact capsids (unmodified capsids: 28.6(9) nm; modified capsids after cleavage: 26.8(8) nm).

Photocleavage of the cysteine-bound hydrophobicity modulator was attempted under the same conditions as that for MESNa-tagged **2.15**. Irradiation of the hydrophobicity modulator-MS2 conjugate proceeded with  $300\pm 10$  nm light from a xenon arc lamp for 1 h in phosphate buffer at pH 7.2, and the products were analyzed by ESI-TOF-MS. One hour of irradiation was found to be insufficient to photolyze the majority of the conjugate. The duration of irradiation was increased to 4 h, and significant conversion was achieved. The deconvoluted mass spectrum of the irradiated conjugate revealed that the cleaved conjugate did not ionize nearly as well as the unmodified protein nor the uncleaved conjugate. To determine the actual percent conversion, a subtractive method was used. Since the amount of material subjected to MS analysis was kept roughly constant and the

also be attributable to the cyclobutanol product formed via Norrish-Yang photocyclization; however this is unlikely due to the presence of the gem-dimethyl directing groups.<sup>34</sup> As the isomeric disubstituted and trisubstituted alkene products of the photocleavage have identical masses and would not necessarily be expected to be resolved during a short analytical HPLC run, it is impossible to determine the ratio of these two products. However, the isomeric cleavage products have very similar hydrophobic character, meaning that either would achieve the desired surface modulation on MS2.

With the photoreactivity of aryl ketone **2.15** demonstrated, modification of Cys87 of the MS2 capsid protein with the hydrophobicity modulator was attempted.





**Figure 2.10.** Conjugation of **2.15** to MS2 and photocleavage. ESI-TOF-MS characterization of MS2 a) before and b) after conjugation with **2.15**. c) Irradiation of MS2-**2.15** for 4 h at  $300 \pm 10$  nm resulted in significant Norrish type II photocleavage. d) Irradiation of unmodified MS2 resulted in significant photooxidation. Total ion count decreased significantly after irradiation.

applications, this photodamage could result in an undesired increase in immunogenicity if used in drug delivery. More immediately, this photodamage introduces batch-to-batch variation and uncertainty in protein concentration, due to changes in the photophysical properties of the aromatic residues of the protein.

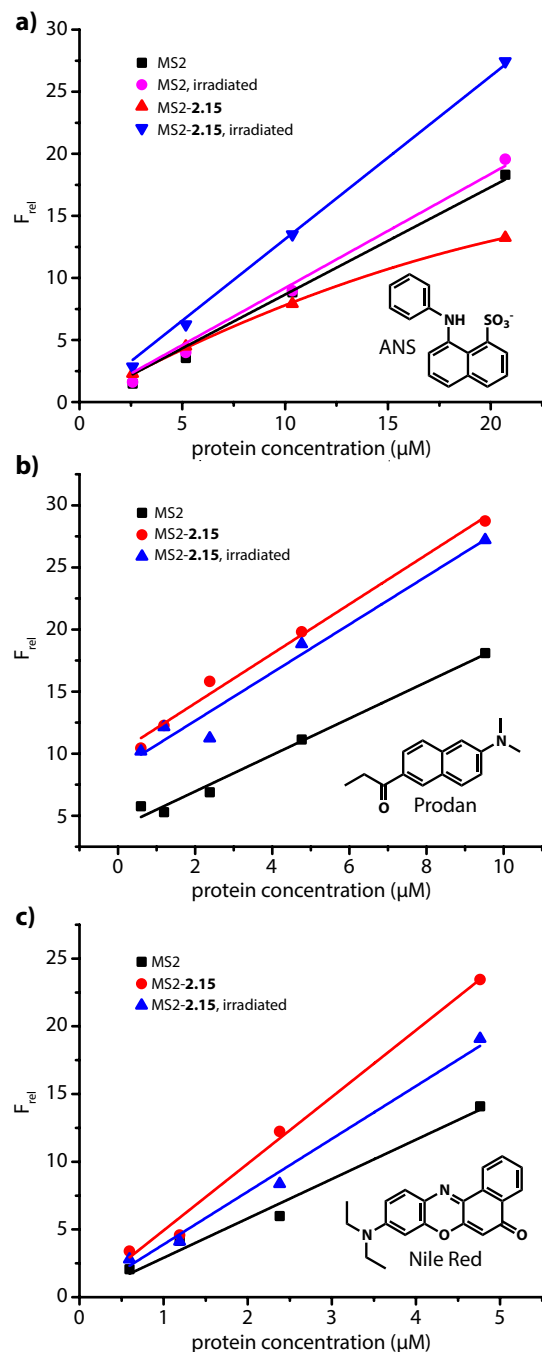
## 2.2.5 Characterization of surface hydrophobicity

Using a spectrofluorometric assessment of protein surface hydrophobicity based on the method developed by Kato and Nakai,<sup>40</sup> the effect of attachment of hydrophobicity modulator **2.15** on MS2 capsid surface hydrophobicity was examined. Incubation of a constant concentration of 8-aniline-1-naphthalenesulfonate (ANS) with increasing concentrations of MS2 capsids or modified MS2 capsids resulted in a significant hypsochromic shift in the emission spectrum of ANS,

deconvoluted mass signals corresponding to unmodified MS2 capsid protein were found to be roughly constant, the decrease in signal corresponding to modified MS2 after irradiation, relative to signal for unmodified protein, could be related directly to percent conversion to the desired photocleaved product. Although the photocleavage reaction was less efficient on the protein than on the free small molecule modulator, significant levels of conversion were achieved.

Modified capsids remained intact after photocleavage, as determined by SEC and DLS characterization. However, it was found that undesired photooxidation of native residues in the MS2 capsid protein had occurred under the photocleavage conditions. Irradiation of unmodified MS2 under these conditions, followed by MS characterization, clearly demonstrated the multiple +16 Da additions. These +16 Da additions, consistent with photooxidation, were also seen for unmodified protein, conjugate material, and cleaved conjugate material in the irradiated sample. While this oxidation did not have a noticeable effect on the protein's quaternary structure, it introduced increased heterogeneity to the material and may have an effect on the long-term stability of the material. Looking towards potential ap-

and an increase in quantum yield. Plotting the increase in emission intensity of ANS at the new blue-shifted emission maxima against protein concentration gives a curve with an initial slope equating to the surface hydrophobicity of the protein sample, by definition.<sup>41</sup> All experimentally determined values of surface hydrophobicity were normalized to that of a sample of untreated MS2 capsids.



**Figure 2.11.** Fluorometric surface hydrophobicity assays. Surface hydrophobicity curves generated for MS2 and MS2-2.15 conjugates, before and after photocleavage, using a) ANS, b) Prodan, and c) Nile Red as probe.

The surface hydrophobicity of untreated MS2 was first measured and normalized to give a value of 1.00(7) (**Figure 2.11a**). Analysis of MS2 labeled with hydrophobicity modulator 2.15 gave a surface hydrophobicity value of 1.00(5) prior to photocleavage. In contrast, MS2 labeled with hydrophobicity modulator 2.15 was found to have a surface hydrophobicity value of 1.51(7) after photocleavage and removal of the phosphoaryl methyl ketone byproduct. Irradiation of unmodified MS2 capsids resulted in, at most, a slight increase surface hydrophobicity relative to the untreated capsids, 1.06(6).

As a slight increase in surface hydrophobicity was found for irradiated unmodified capsids, a small portion ( $\leq 0.05$  normalized surface hydrophobicity units) of the increase in surface hydrophobicity may be credited to changes in the capsid not related to the surface hydrophobicity modulator. Increases in protein surface hydrophobicity caused by oxidative damage is not unprecedented. Oxidation of methionine, tyrosine, and tryptophan residues has been linked to an increase in surface hydrophobicity of rat liver protein extracts, as determined through an ANS spectrofluorometric assay.<sup>42</sup>

Based on the ANS spectrofluorometric assay, the hydrophobicity modulator after photocleavage imparted a 51% increase in surface hydrophobicity to the MS2 capsid. Considering that the attachment and cleavage of modulator 2.15 only amasses to a 2% increase in molecular weight of the conjugate, this was a remarkable achievement in surface property modulation. The finding that the uncleaved conjugate showed a similar hydrophobicity value to unmodified capsids suggested that removal of the solubilizing aryl phosphate moiety was necessary to increase the hydrophobic character of the interior surface of the MS2 capsid. Although it appeared that the surface hydrophobicity of the conjugates increased only after photocleavage, it seemed unlikely that this could be an accurate representation of the system as a long saturated alkane was introduced

onto the interior face of the majority of capsid protein monomers. The curve that was used to determine the surface hydrophobicity of the uncleaved conjugate deviated significantly from linearity. The vast majority of surface hydrophobicity curves, including those in this study, are linear for a range of protein concentrations up to and beyond the upper limit used in this experiment. The finding that the curve for the uncleaved conjugate was non-linear suggested that a factor other than surface hydrophobicity was affecting the spectral properties of the reporter molecule, ANS. Considering the possibility of electrostatic repulsion between the anionic phosphate of the surface hydrophobicity modulator and the anionic sulfate on the surface hydrophobicity reporter ANS, it seemed likely that this unfavorable interaction resulted in the nonlinearity of the curve. This hypothesis is strongly supported by the work of Matulis and Lovrien, which demonstrated the strong effect of ion pair formation on ANS fluorescence turn-on.<sup>43</sup>

Although ANS is the most widely-used fluorescent reporter of surface hydrophobicity, the strong influence of electrostatic interactions on the results of ANS spectrofluorometric assays raises some doubts as to the validity of its use in direct determination of surface hydrophobicity.<sup>43</sup> To probe the validity of this assay on the current system of interest, and to confirm the above surface hydrophobicity measurements, the assay was repeated with two alternative nonionic hydrophobicity probes: 6-propionyl-2-(dimethylamino)naphthalene (Prodan) and Nile Red.<sup>41,44</sup>

Spectrofluorometric analysis with Prodan resulted in a less-than-ideal set of hydrophobicity curves (**Figure 2.11b**). This assay was found to result reproducibly in regressions that did not share common y-intercepts between the three sample conditions analyzed (unmodified, modified but uncleaved, modified and cleaved). Nevertheless, this experiment suggested that both the uncleaved conjugate, 1.36(11), and the cleaved conjugate, 1.32(16), imparted significant and similar degrees of surface hydrophobicity to the MS2 capsid relative to the unmodified control, 1.00(9). These findings suggested that the nonlinear hydrophobicity curve and lack of hydrophobicity imparted by the uncleaved conjugate observed previously were a result of using the anionic probe ANS and not due solely to the nature of the uncleaved hydrophobicity modulator.

Compared to Prodan, spectrofluorometric analysis of surface hydrophobicity with Nile Red resulted in hydrophobicity curves of higher quality (**Figure 2.11c**). Nile Red gave relative surface hydrophobicity values of 1.00(6), 1.70(9), and 1.34(8) for unmodified MS2, uncleaved modified capsid protein, and cleaved and modified protein, respectively. Consistent with both the ANS and the Prodan assays, the modified protein showed a significant increase in surface hydrophobicity after cleavage, although to a lesser extent than determined with Nile Red as reporter. In agreement with the Prodan assay but inconsistent with the ANS assay is the finding that the uncleaved modified protein results in a significant increase in surface hydrophobicity. Furthermore, this assay suggests that intact hydrophobicity modulator imparts an even greater amount of hydrophobic character to the surface of the capsid than does the cleaved modulator.

To reconcile the finding that a long alkyl chain capped with a hydrophilic aryl phosphate imparts more hydrophobic character to the interior surface of the MS2 capsid than does alkene terminated cleaved modulator, two factors can be considered. First, specific  $\pi$ - $\pi$  stacking between the aryl ketone in the uncleaved conjugate and the aromatic naphthooxazine core of Nile Red could have a larger contribution to the fluorescence turn-on than the purely hydrophobic interactions between the dye and the alkyl chain. However, a more interesting interpretation of this result relates to the confirmation of the hydrophobicity modulator on the protein surface. The terminal phosphate group of the cysteine-bound hydrophobicity modulator has the potential to form salt bridges with cationic residues within roughly a 2–2.5 nm radius from Cys87. Specific residues that could form

salt bridges with this phosphate moiety include Arg38, Arg49, Arg56, Arg83, Lys43, Lys57, and Lys61. Depending on the relative location of the arginine or lysine residue to the modified cysteine, the alkyl linker of the hydrophobicity modulator may take on either an extended conformation or an arced conformation. In contrast, the photocleaved modulator cannot form salt bridges. It would be expected to minimize hydrophobic interactions with the solvent and polar residues by packing onto itself, with a nearby hydrophobicity modulator, or onto nearby hydrophobic regions of the protein surface. In effect, the uncleaved modulator may actually provide a larger solvent-exposed hydrophobic surface than does the cleaved modulator, as suggested by the Nile Red study.

Photoswitchable hydrophobicity modulator **2.15** was shown to significantly increase the surface hydrophobicity of the MS2 capsid, between 0% and 70% before cleavage, and between 32% and 51% after cleavage. Although this was a major achievement, this increase in surface hydrophobicity may not be sufficient to achieve the binding and retention of hydrophobic actives desired in this work. To determine whether these conjugates were capable of hydrophobic cargo retention, a size exclusion chromatography (SEC) study was undertaken. Unmodified MS2 capsids, capsids modified with the hydrophobicity modulator, and modified capsids after photocleavage were incubated with either ANS, Prodan, or Nile Red and then analyzed by SEC. Loading efficiency can be determined through analysis of the relative contributions of the protein (specifically, the contribution of tryptophan, tyrosine, and phenylalanine residues at 280 nm) and the probe to the UV/Vis spectrum of the fraction containing intact capsids. Unfortunately, it was found that neither of the constructs were able to bind and retain a detectable amount of the probe on the SEC timescale. Additionally, signal from the probe was undetectable across the entire chromatogram suggesting that any probe molecules within the capsid had migrated out and precipitated upon dilution. Inspection of the SEC column's guard column corroborated this hypothesis.

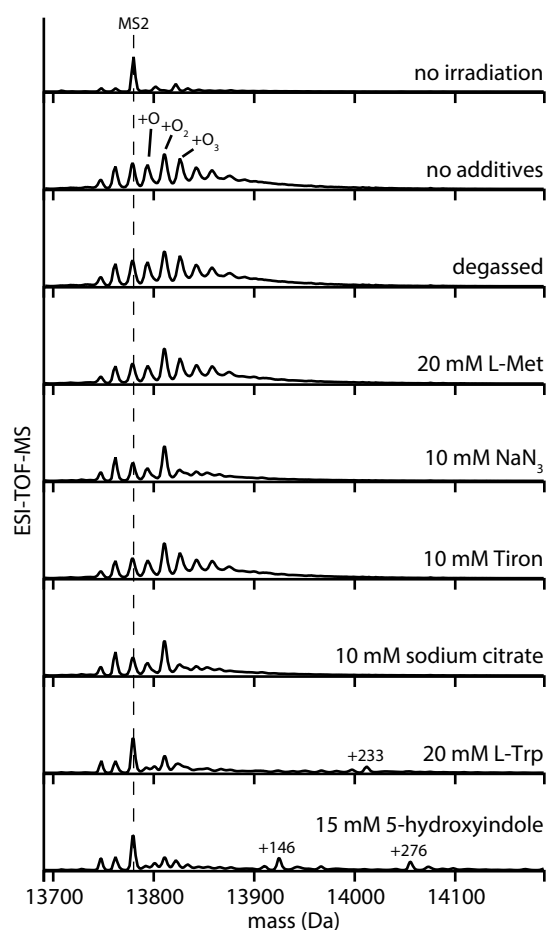
These findings, taken together with the results of the surface hydrophobicity studies, suggest that the hydrophobic probes are interacting with the hydrophobically modified interior surface of the capsid, but that no strong binding is occurring. As the modified capsids were not able to bind and retain the readily observable probes, efforts were not made to characterize their ability to bind hydrophobic actives of greater interest in the field of drug delivery. Rather, efforts were directed to increase the hydrophobic character of the capsids significantly further by moving away from surface modulation and moving towards the creation of hydrophobic droplet within the core of the capsid.

## 2.2.6 Improving the biocompatibility of Norrish type II chemistry

The utility of Norrish type II photocleavage as a new biocompatible bond cleavage reaction has been demonstrated on the MS2 capsid (Section 2.2.4). However, the conditions required for this reaction have been shown to have off-target effects. Irradiation of a protein or protein conjugate with photons in the UVB region (280–230 nm) can result in photooxidation of sensitive amino acids, as demonstrated by the series of M+16 Da species in *Figure 2.10d*. Although the specific surface modification discussed in Section 2.2.5 did not accomplish the desired ability to bind and retain hydrophobic actives for delivery applications, the optimization of this novel biocompatible cleavage chemistry was attempted, as this chemistry retains high potential for use in photoswitchable hydrophobicity-modulating polymers (Sections 2.3 and 2.4) and in other applications wherein controlled surface hydrophobicity enhancement would be of use. This section will begin with a brief overview of the photooxidation of proteins, followed by two strategies for improving the or-

thogonality of this cleavage reaction.

Significant efforts have been made to understand the products and pathways of photodamage to proteins.<sup>45</sup> Understanding the effects of exposure of proteins to light has become increasingly important, as growing numbers of protein biologics have entered the clinic in the past decades.<sup>46</sup> Direct damage to proteins due to UVB irradiation is limited to the aromatic residues, tryptophan, tyrosine, phenylalanine, histidine; as well as cystine. Upon irradiation, these residues are excited to their first excited singlet state, which, among other potential pathways, can then relax via intersystem crossing to the triplet state. The residue in its excited triplet state can then go on to form a radical cation by electron transfer to a suitable donor or, after a further deprotonation step, form an uncharged radical. The residue in its triplet state can also react with molecular oxygen to form reactive oxygen species such as the peroxy radical. Secondary photodamage occurs via radicals formed in primary photooxidation processes and reactive oxygen species. Taken together, primary



**Figure 2.12.** Effects of photoprotectants on MS2 photooxidation. Samples of MS2 in sodium phosphate buffer, with or without photoprotectant additive, were irradiated at  $300\pm 10$  nm for 2 h. ESI-TOF-MS analysis revealed significant photooxidation in each sample, except for those treated with L-Trp or 5-hydroxyindole. Larger mass additions, consistent with oxidative coupling, were observed in samples treated with L-Trp and 5-hydroxyindole.

and secondary photooxidation processes result in the formation of modified residues and crosslink formation within proteins. Most notably, oxidation of Trp residues results in formation of *N*-formylkynurenine and kynurenine; Tyr can form dityrosine crosslinks; Phe can form tyrosine isomers; Met is oxidized to methionine sulfoxide; and Cys has a number of oxidation products including cysteic acid. Ultimately, the damage induced by UVB irradiation can result in major changes in protein structure, aggregation, and loss of function.<sup>46</sup>

While complete protection of a protein sample from light is the only method of completely avoiding photodegradation, irradiation is unavoidable and necessary in the case of a photocleavage reaction. The simplest and most straightforward strategy for minimizing the photooxidation of proteins involves the removal of oxygen from the system. Total elimination of dissolved oxygen from a small volume sample is not a trivial task. A sample of MS2 was gently sonicated in an ultrasonic bath for 10 min, then transferred to a cuvette and blanketed with nitrogen prior to irradiation. Unfortunately, this simple procedure did not prevent photooxidation of the sample (**Figure 2.12**). The use of more rigorous oxygen-removal techniques, such as pump-freeze-thaw or sparging with nitrogen, are likely to result in denaturation of the protein samples and were thus avoided.

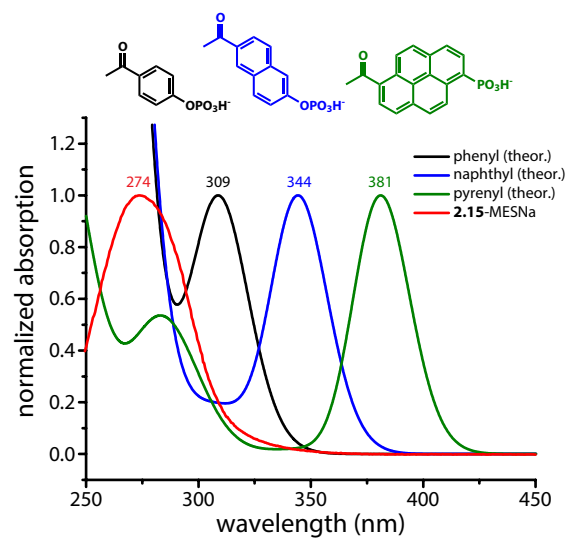
A number of agents have been reported as photoprotectants or scavengers for reactive oxygen species and a screen of such agents was conducted for the photoprotection of MS2. L-Methionine at 20 mM has been shown to reduce photooxidation in the HER2 recom-

binant human monoclonal antibody<sup>47</sup> but showed only slight photoprotective qualities in the present case. Sodium azide has been demonstrated as a specific scavenger of singlet oxygen, in a study of the specificity of chemiluminescence of luciferin analogues in response to various oxidants.<sup>48</sup> However, 10 mM sodium azide was not effective in minimizing photooxidation of MS2. Tiron (4,5-dihydroxybenzene-1,3-disulfonate), a known superoxide scavenger at 10 mM,<sup>49</sup> was also ineffective as a photoprotectant in this system. Sodium citrate, at 10 mM, was found to provide some protection against oxidation beyond the two oxidations, presumably via the chelation of iron. However, the M+32 species remained the major species after irradiation.

As tryptophan was the most likely site of photooxidation and the addition of L-methionine, a known protectant specific for methionine oxidation, did not reduce photodamage to MS2, tryptophan-specific ROS protectants were examined. The tryptophan-specific oxidative protection of L-tryptophan has been demonstrated on parathyroid hormone under a range of artificially-induced oxidative stress conditions.<sup>50,51</sup> However, L-Trp is also well-known to produce ROS under UV irradiation.<sup>52</sup> A recent screen for tryptophan-like antioxidants that do not generate additional ROS upon irradiation has revealed a collection of potential indole derivatives for use as excipients in protein formulations.<sup>53</sup> One of the more promising agents found in this study was 5-hydroxyindole, and this photoprotectant was screened alongside L-Trp for the photoprotection of MS2. 5-Hydroxyindole at 15 mM and L-Trp at 20 mM both significantly minimized photooxidation of MS2. The major M+32 Da species from the unprotected MS2 sample after irradiation had decreased significantly along with higher order oxidation products, with the addition of these indole derivatives. This finding supports the hypothesis that one or more of the tryptophan residues of MS2 are the primary targets of photooxidation, forming an *N*-formylkynurenine residue (M+32 Da), which can then act further as a photosensitizer. While L-Trp and 5-hydroxyindole prevented the formation of most M+(16n) Da species, larger mass additions were observed, corresponding to oxidative coupling between the protein and the indole derivatives. Although conditions were found that partially protected MS2 from photodamage, the use of excipients resultant in less than satisfactory photoprotection.

As conditions were not found to prevent photooxidation of the MS2 capsid protein under direct UVB irradiation, a more tailored strategy was developed. The issue of photooxidation in the present case stems from the requirement of irradiation with photons around 300 nm. This requirement is based on the absorption profile of phenylketone-based photoswitchable hydrophobicity modulator **2.15**. By shifting the wavelength required for photocleavage into the UVA (315–400 nm) or visible (390–700 nm) regions, direct excitation of aromatic residues can be largely avoided, eliminating primary photooxidation pathways from tryptophan. As the experiments above suggested that this pathway was the main factor in photooxidation of MS2, red-shifting the wavelength of irradiation for cleavage should greatly minimize photodamage to the protein.

While shifting the wavelength of irradiation to the far red tail (325–350 nm) of the absorption spectrum of hydrophobicity modulator **2.15** may help to alleviate the issue of photooxidation of native residues, the time required to achieve significant cleavage would become unreasonable. Alternatively, the aryl ketone moiety of the hydrophobicity modulator could be tuned to absorb more to the red than the native residues. Maintaining the same general phosphoaryl ketone motif as in modulator **2.15** but expanding the phenyl group into a naphthyl or pyrenyl system would provide a modulator with the desired red-shifted absorption spectrum. To get a sense of how these derivatives would compare, absorption spectra were calculated for 4-acetylphenyl hydrogen phosphate and its naphthyl and pyrenyl derivatives in water, using time-dependent density functional



**Figure 2.13.** Calculated UV-Vis absorption spectra for redshifted Norrish type II substrates. Spectra were obtained by TDDFT and plotted against the observed spectrum of **2.15**-MESNa in phosphate buffer. The calculated  $\lambda_{\text{max}}$  value for **2.15** was found to be 35 nm redshifted from the observed value.

photocleavable protecting groups have been shown to perform as expected. Naphthalene-based photocleavable protecting groups include 3-nitro-2-naphthalenemethanol and the 7-methoxy derivative for the protection of carboxylic acids,<sup>54,55</sup> 2-(bromoacetyl)naphthalene for carboxylic acids,<sup>56</sup> 2-naphthalenemethanol for the protection of phosphates,<sup>57</sup> 3-hydroxy-2-naphthalenemethanol for protection of acids<sup>58</sup> and for reversible thiol protection through the *o*-naphthoquinonemethide,<sup>59</sup> and 2-acetylnaphthalene and 2-formylnaphthalene for protection of alcohols as photolabile ketals and acetals.<sup>60</sup>

Given the success of other naphthyl protecting/caging groups, the synthesis of naphthyl hydrophobicity modulator **2.26** was undertaken (**Figure 2.14**). The naphthyl derivative was prepared in much the same way as the first generation reagent thiol-reactive photoswitchable reagent **2.15**.  $\alpha,\beta$ -Unsaturated naphthyl ketone **2.27** needed for the central C–C bond forming reaction in the synthesis of **2.26** was prepared through a three step synthesis. Commercially available 2-bromo-6-methoxynaphthalene was converted into the corresponding Grignard reagent and coupled to 3-methylcrotonaldehyde. Without purification, the resultant alcohol was then oxidized to the corresponding ketone **2.27** with  $\text{MnO}_2$ . A copper-catalyzed Grignard coupling between  $\alpha,\beta$ -unsaturated naphthyl ketone **2.27** and the Grignard reagent of alkyl bromide **2.16** afforded the desired alkyl naphthyl ketone **2.28** in very good yield.

In preparation for the deprotection of the methyl naphthyl ether, the silyl ether was deprotected under acidic conditions and reprotected by acetylation with acetic anhydride. Successful selective deprotection of ether **2.29** was achieved with  $\text{AlCl}_3$  and hexanethiol. Resulting naphthol **2.30** was then converted to phosphate triester **2.31** by reaction with di-*tert*-butyl *N,N*-diisopropylphosphoramidite, activated with imidazolium trifluoroacetate, followed by oxidation with *tert*-butyl hydroperoxide. As in the preparation of the phenyl phosphate analog **2.23**, it was critical to ensure that the naphthol had completely converted to the phosphite derivative prior to oxidation as naphthol **2.30** and phosphate triester **2.31** were chromatographically indistinguishable under

theory (TDDFT) (**Figure 2.13**). The calculated  $\lambda_{\text{max}}$  for the phenyl derivative was red-shifted 35 nm from the experimental  $\lambda_{\text{max}}$  of hydrophobicity modulator **2.15**. Nonetheless, the calculation provided reasonable shifts between the absorption maxima within the series of derivatives: 35 nm from phenyl to naphthyl and 37 nm from naphthyl to pyrenyl. Adjusting the calculated values to correspond with the observed absorption properties of modulator **2.15** gives predicted  $\lambda_{\text{max}}$  value of 309 nm and 346 nm for the naphthyl and pyrenyl derivatives, respectively. Given a hydrophobicity modulator with an absorption maxima of 309 nm, a significant portion of the absorption cross-section would not overlap with that of the native residues, allowing for photocleavage to be effected under selective irradiation.

Naphthyl ketones represent a novel substrate for Norrish type II photocleavage. Although there is no direct precedent for this class of photochemistry on naphthyl substrates, naphthyl analogs of other

normal-phase conditions.

Acetate **2.31** was deprotected with LiOH and, as with the phenyl analog **2.23**, care had to be taken to ensure the aryl phosphate ester remained intact. Surprisingly, installation of the maleimide functionality at the hydroxyl terminus of naphthyl substrate **2.32** was not straightforward. The Mitsunobu reaction between alcohol **2.32** and maleimide was unsuccessful, even under optimized Walker conditions, whereas it had been for phenyl analog **2.24**. Mitsunobu coupling was also attempted with more nucleophilic furan-maleimide Diels-Alder adduct **2.6**. However, this too was unsuccessful. Ultimately, a three-step activation-substitution-deprotection sequence was required to install the maleimide in this reagent. Alcohol **2.32** was converted to corresponding tosylate **2.33** under standard conditions. Direct displacement of the tosylate by maleimide was unsuccessful. However, substitution of tosylate **2.33** by furan-maleimide Diels-Alder adduct **2.6** afforded masked maleimide **2.34**, but required long incubation at elevated temperatures and was low-yielding. On attempted unmasking of adduct **2.34** through a thermal retro-Diels-Alder reaction, it was found that, in addition to the desired transformation, the di-*tert*-butyl phosphate triester had decomposed into desired naphthyl phosphate monoester **2.26**. Unfortunately, the high temperatures (111 °C) required for the retro-Diels-Alder reaction also resulted in significant dephosphorylation of the naphthol, resulting in a poor yield of the desired reagent after HPLC purification.

Overall, naphthyl derivative **2.26** of the thiol-reactive photocleavable reagent was prepared in 11 linear steps from 2-bromo-6-methoxynaphthalene in less than 0.1% total yield. Although the current synthetic route was highly inefficient, enough material was produced to allow for testing

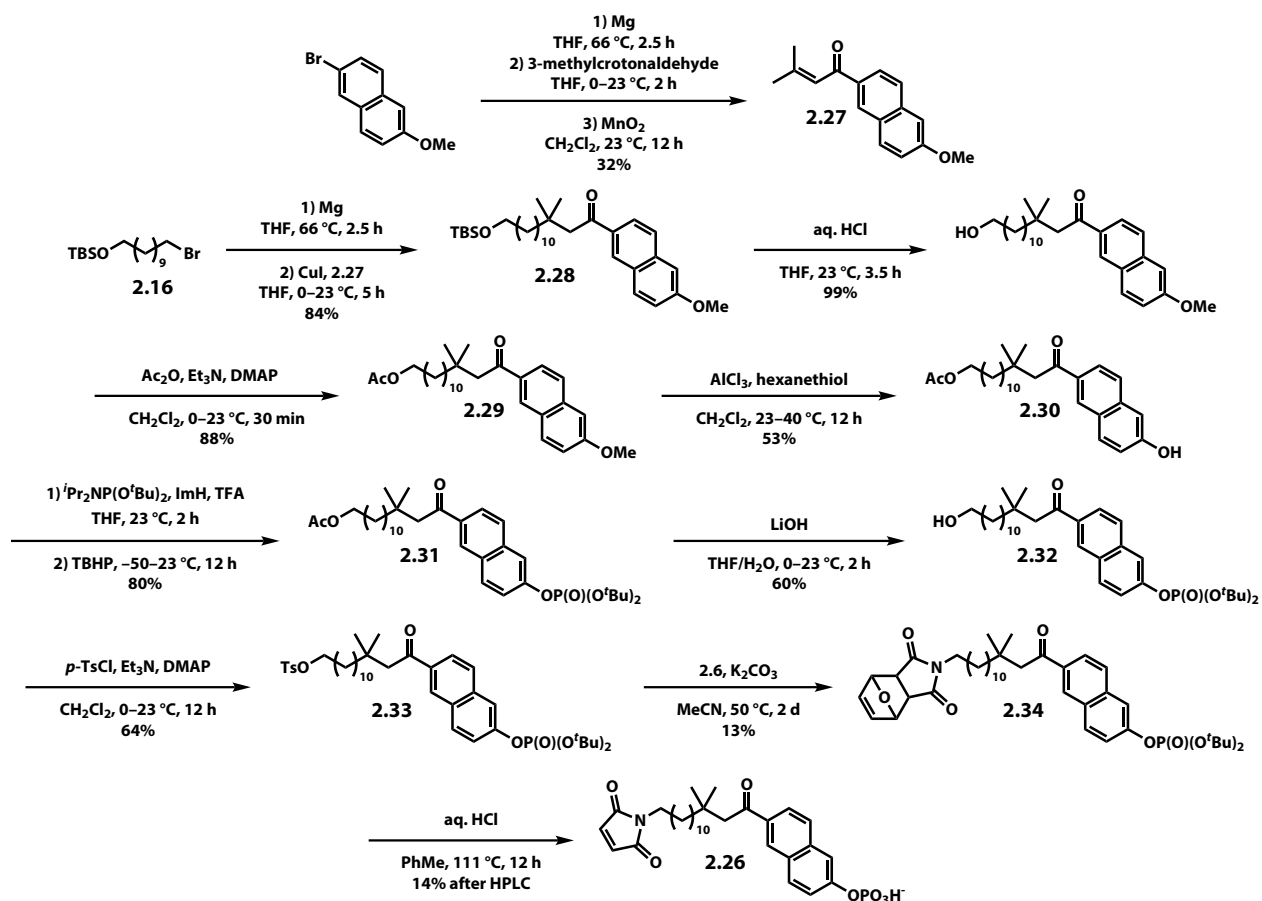
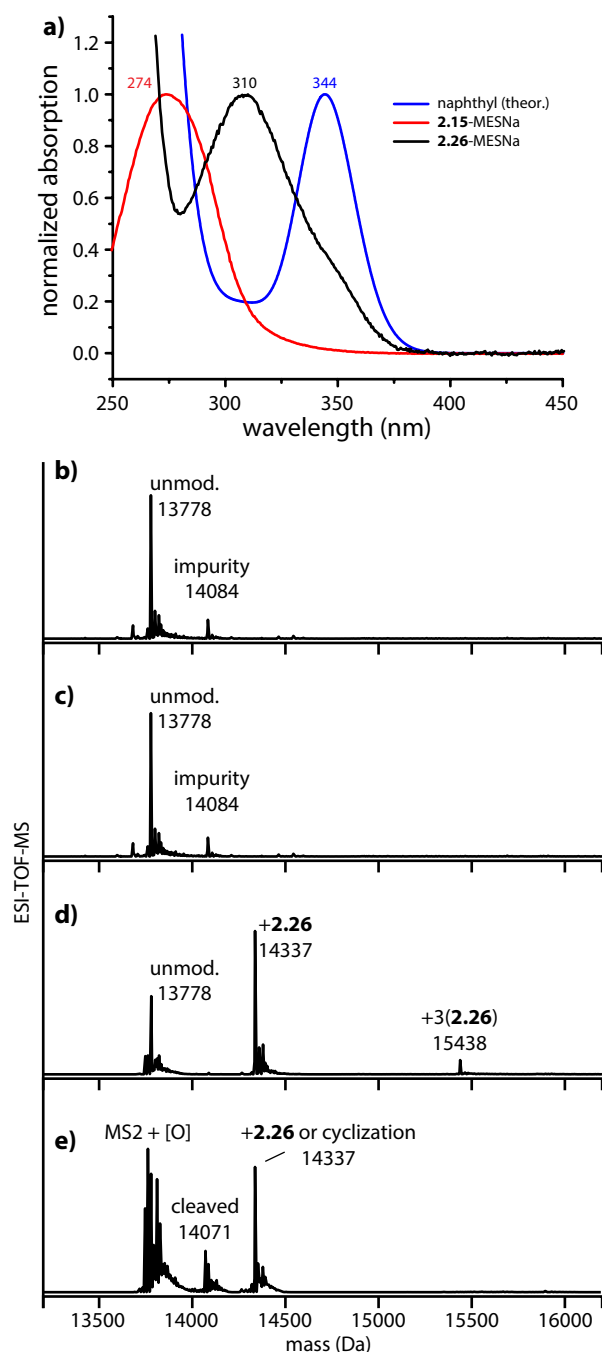


Figure 2.14. Synthesis of photocleavable reagent **2.26**.





**Figure 2.15.** Hydrophobicity modulator **2.26**. a) UV-Vis characterization of **2.26**-MESNa revealed an absorption maxima of 310 nm, consistent with the predicted value of 309 nm. ESI-TOF-MS analysis of MS2 b) before and c) after irradiation at  $340 \pm 10$  nm for 4 h revealed that no photodamage had occurred. d) Conjugation of **2.26** to MS2 did not go to completion and some triple modification was observed. e) Irradiation of MS2-**2.26** resulted in significant photocleavage. However, **2.26** acted as a sensitizer and resulted in the photooxidation of unmodified MS2 capsid proteins.

of its properties both on and off of protein scaffolds. The overall yield of this route is limited by the efficiency of the formation of  $\alpha,\beta$ -unsaturated naphthyl ketone **2.27** and the transformation of primary alcohol **2.32** into a maleimide. Improving the preparation of ketone **2.27** is largely unwarranted since all materials involved are commercially available and relatively inexpensive. However, there is room for improvement in the late stage transformation of alcohol **2.32** to maleimide **2.26**. Although direct substitution using Mitsunobu conditions would be optimal, this substrate does not appear to be a willing coupling partner in this reaction. Significant increases in yield and reaction rate may be achieved by going through the mesylate of **2.32** rather than the tosylate.

Characterization of naphthyl hydrophobicity modulator **2.26** revealed an absorption maxima of 310 nm. While the observed  $\lambda_{\max}$  of **2.26** differed from that determined computationally (*Figure 2.15a*), applying the adjustment factor from *Figure 2.13* puts the two values in excellent alignment. Photocleavage of modulator **2.26** was attempted in phosphate buffer after reaction of the maleimide with MESNa. Irradiation with a xenon arc lamp at  $340 \pm 10$  nm resulted in clean conversion to the cleaved product. However, significantly longer exposure (3 h compared to 1 h for the phenyl derivative) to the light source was required to achieve full conversion. As with the phenyl derivative, the product composition was not determined; that is, the two isomeric alkene products were not separated by HPLC nor were NMR studies conducted to determine their ratio.

With the photoreactivity of naphthyl ketone **2.26** demonstrated, conjugation of the reagent to Cys87 of the MS2 capsid protein was attempted (*Figure 2.15d*). Unlike the phenyl derivative of the hydrophobicity modulator, conjugation did not proceed as expected. Under the conditions found for optimal conjugation of the first generation hydrophobicity modulator—40 equivalents of maleimide reagent in phosphate buffer at pH 7.2 at room temperature for 4 h—only about 20% of the capsid proteins were modified. Extending the

reaction time to 12 h and doubling the amount of reagent increased the conjugation efficiency to roughly 60%. However, non-trivial quantities of triply modified protein was observed by ESI-TOF-MS. Although SEC and DLS characterization clearly demonstrated that the capsid remained intact, the presence of triple modification suggested that the reagent was able to disrupt the protein structure and modify both of the generally solvent inaccessible intrinsic cysteine residues (Cys46 and Cys101).

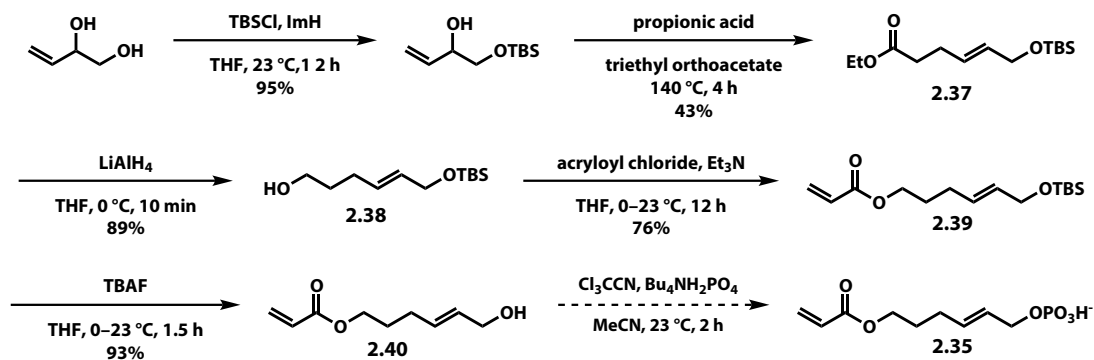
Looking past the potential structural changes of triply-modified subunits, photolysis of the modified capsids was attempted. Following 4 h of irradiation with a xenon arc lamp at  $340\pm 10$  nm, the conjugate sample was analyzed by SEC and ESI-TOF-MS. The SEC elution profile of the irradiated conjugate was consistent with that of untreated MS2 capsids. MS characterization revealed that the photolysis had been somewhat successful, in that a portion of the conjugate had been cleaved. However, the issue of photooxidation had not been resolved. Whereas unmodified MS2 capsid protein showed no oxidation after irradiation at  $340\pm 10$  nm, the characteristic +16 Da additions were seen for the conjugate after identical treatment. This finding suggested that while photooxidative pathways involving direct excitation of native aromatic residues had been eliminated, the naphthyl hydrophobicity modulator could act as a sensitizer, generating reactive oxygen species. The ROS could then result in indirect photodamage. *N*-Formylkynurenine would likely be generated, which, due to its red-shifted absorption profile compared to Trp, could then be directly excited under the irradiation conditions used for photocleavage resulting in further photodamage.

In summary, neither the use of excipients or use of a photoswitchable moiety with a red-shifted excitation profile resulted in significant reduction in photooxidative damage to the MS2 capsid protein. While the oxidative damage results in a less well-defined conjugate material and makes characterization of the material more difficult, these issues may or may not be of concern, depending on the ultimate application of such a protein-based formulation. For applications for which the oxidative damage is a major concern, rigorously anaerobic conditions for the photocleavage reaction could be achieved with the assistance of glucose oxidase as a scavenger of dissolved oxygen.<sup>61</sup>

### 2.3 Towards creating a hydrophobic droplet within MS2

At the onset of this project, it was unknown whether sufficient hydrophobic character could be imparted to the interior of the MS2 capsid by surface modification to enable the capsid to act as a vehicle for the delivery of hydrophobic actives. Based on the results from the previous sections on surface modification with a cysteine-reactive hydrophobicity modulator, it became evident that this strategy could modulate the surface hydrophobicity to a significant extent but not to the level needed for the binding and retention of hydrophobic actives. While developing the small molecule hydrophobicity modulators discussed above, two general strategies were developed to take advantage of the core volume of the MS2 capsid and create within it a hydrophobic droplet. The first strategy would take advantage of the natural self-assembly pathway of the MS2 capsid, but using an anionic switchable hydrophobicity-modulating polymer in place of the natural assembly-templating single stranded RNA. The second strategy envisioned the growth of switchable hydrophobicity-modulating polymer grafted from the interior surface of the MS2 capsid.

Prior to discovering that  $\pi$ -allylpalladium chemistry on an allyl phosphate substrate did not result in elimination of the water-solubilizing phosphate moiety (Section 2.2.2), efforts were made to create an acrylate monomer with an alkyl ester terminating in an allyl phosphate group. Unlike the preparation of the cysteine-reactive analog **2.15**, this reagent would need to be prepared on a larger



**Figure 2.16.** Preparation of allyl phosphate-containing acrylate monomer for polymerization.

scale to allow for the optimization at the polymerization stage. As such, a more efficient synthesis was envisioned, with the drawback being that a shorter alkyl linker would be present in the final reagent.

Acrylate monomer **2.35** was prepared in 6 steps from commercially available material (**Figure 2.16**). 3,4-Dihydroxy-1-butene was protected at the primary alcohol with TBSCl in excellent yield to afford mono-TBS ether **2.36**. Johnson-Claisen rearrangement of allylic alcohol **2.36** provided ester **2.37** in good yield. Reduction of ester **2.37** with lithium aluminum hydride afforded primary alcohol **2.38** in very good yield. Acylation of alcohol **2.38** with acryloyl chloride afforded acrylate **2.39** in moderate yield. Significantly higher yields would be expected for this simple acylation, but optimization was not performed for this step. Allylic alcohol **2.40** was obtained in excellent yield through mild deprotection of TBS ether **2.39** with tetrabutylammonium fluoride. Phosphorylation of allylic alcohol **2.40** was attempted with tetrabutylammonium hydrogen phosphate and trichloroacetonitrile. While desired phosphate monoester **2.35** was isolated, this phosphorylation was low yielding and purification was difficult, even by HPLC. At about this time, the realization was made that the palladium-based strategy for hydrophobicity modulation was not going to be successful and efforts towards the preparation of a palladium-switchable hydrophobicity-modulating polymer were redirected.

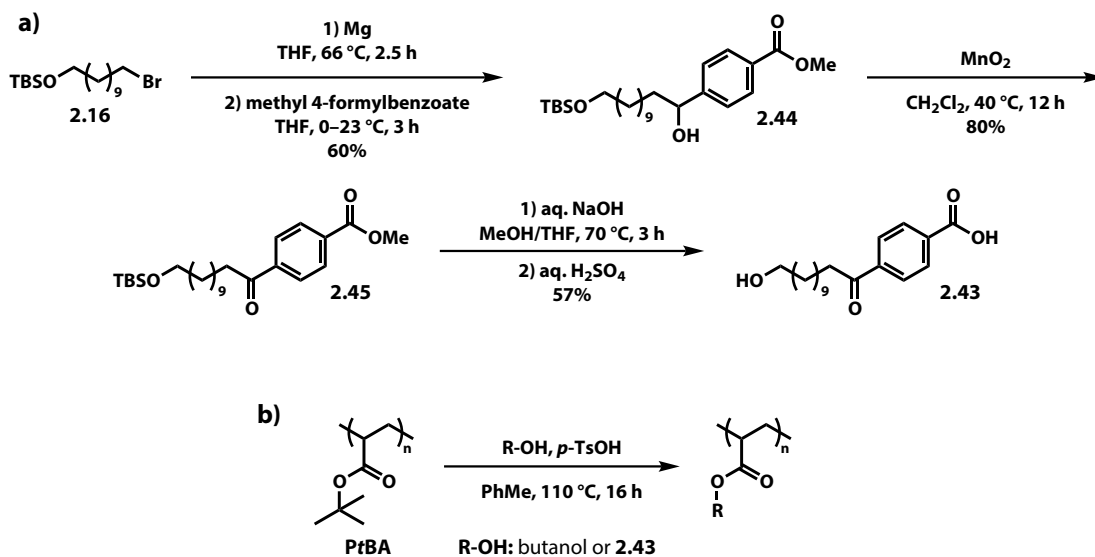
Along with the switch towards the development of a photoswitchable cysteine-reactive hydrophobicity modulator came a switch towards the preparation of a photoswitchable polymeric hydrophobicity modulator. One lesson learned from the attempted preparation of monomer **2.35** was the difficulty associated with the large-scale preparation of complex phosphate-containing molecules. Rather than working towards the preparation of an acrylate monomer capable of undergoing Norrish type II photocleavage, a reagent capable of modifying a pre-made polyacrylate through esterification was chosen. The advantage of a post-polymerization modification reagent over a monomer really comes down to scale. Preparation of a polymer from the acrylate monomer would require extensive modification to obtain polymers of appropriate size or sizes for MS2 reassembly and each polymerization reaction would require a significant mass of precious phosphate ester material. On the other hand, post-polymerization modification has the advantage of using pre-formed polymers with known sizes and can even be obtained from commercial sources. Significantly less reagent would be required for the optimization of PPM over polymerization from scratch.

A wide selection of low polydispersity poly(alkyl acrylates) and poly(alkyl methacrylates) of many molecular weights are commercially available. As such, finding conditions for the post-

polymerization modification of these readily-available materials would allow for the rapid screening of well-defined polymer variants on MS2 assembly and hydrophobicity switching. Precedent exists for the transesterification of poly(methyl methacrylate) (PMMA). Poly(ethylene oxide) has been grafted to PMMA through transesterification with potassium methoxide or potassium naphthalene.<sup>62</sup> For the polymers of interest in the present work, slightly less-demanding conditions were desired. Based on conversations with polymer chemists at BASF and Otera's excellent review of transesterification,<sup>63</sup> a selection of catalysts for poly(alkyl acrylate)/poly(alkyl methacrylate) modification was screened. For screening purposes, the transesterification of poly(methyl acrylate) (PMA) and PMMA was attempted with *n*-butanol in toluene (1:3 v/v) at 110 °C in sealed vials with the transesterification promoting agent (5 equivalents per monomer unit). Under these conditions, it was found that no transesterification had occurred after 16 h with DBU, sodium methoxide, titanium(IV) isopropoxide, trifluoroacetic acid, methanesulfonic acid, or *p*-toluenesulfonic acid (*p*-TsOH), as determined by NMR. As no conversion was observed under these already-pressing conditions, PMA and PMMA were replaced with the more reactive *tert*-butyl derivatives poly(*tert*-butyl acrylate) (*PtBA*) and poly(*tert*-butyl methacrylate) (*PtBMA*). As in the case of PMA and PMMA, no transesterification was observed for *PtBMA* under any of the conditions screened. However, full conversion of *PtBA* to poly(butyl acrylate) was found for reaction in the presence of *p*-TsOH. No noticeable transesterification of *PtBA* had occurred with any of the other reagents. The transesterification between *PtBA* and *n*-butanol was found to run to completion even when the amount of *n*-butanol was lowered to 5 equivalents per monomer unit, and the amount of *p*-TsOH was decreased to 1 equivalent per monomer unit.

Alcohol **2.41** was designed as a reagent for the post-polymerization modification of polyacrylates via transesterification. Synthesis of reagent **2.41** was attempted from phosphate triester **2.24**, a late-stage intermediate for the preparation of cysteine-reactive hydrophobicity modulator **2.15**. All that was required for preparation of the desired alcohol was deprotection of the *tert*-butyl phosphate ester to reveal the aryl phosphate monoester. It was quickly realized that phosphate triester **2.24** was not a willing substrate for this simple conversion. While the desired phosphate monoester was formed under standard deprotection conditions (TFA/DCM), significant quantities of an unexpected byproduct were observed. LC-MS and NMR analysis revealed the byproduct to be macrocyclic phosphate diester **2.42**, a completely unexpected product. A time screen of the reaction showed that the desired phosphate monoester **2.41** formed first but was rapidly converted to macrocyclic phosphate diester **2.42** as deprotection continued towards completion. In an attempt to minimize cyclization, water was added as a competing nucleophile under TFA/THF conditions but cyclization was not prevented. Deprotection with HCl in methanol also resulted in formation of the undesired phosphate macroester. While working on optimizing this deprotection, it was also realized that the aryl O–P bond in phosphate triester **2.24** was prone to hydrolysis under the storage conditions, neat at –20 °C. As phosphate triester **2.24** and its dephosphorylated analog could not be separated by silica gel chromatography, the preparation of transesterification partner **2.41** was deemed too complex of a target for this proof of principle study.

As reagent **2.41** proved to be a much more challenging target than expected, a simplified aryl ketone transesterification reagent (**2.43**) was proposed for a first-generation anionic hydrophobicity modulating polymer (**Figure 2.17**). The synthesis of reagent **2.43** was greatly simplified from that of the original target through the removal of the gem-dimethyl substituents and substitution of the phosphate solubilizing group for a carboxylic acid. Reagent **2.43** was prepared in five steps from commercially available materials. The core of the reagent was constructed through coupling



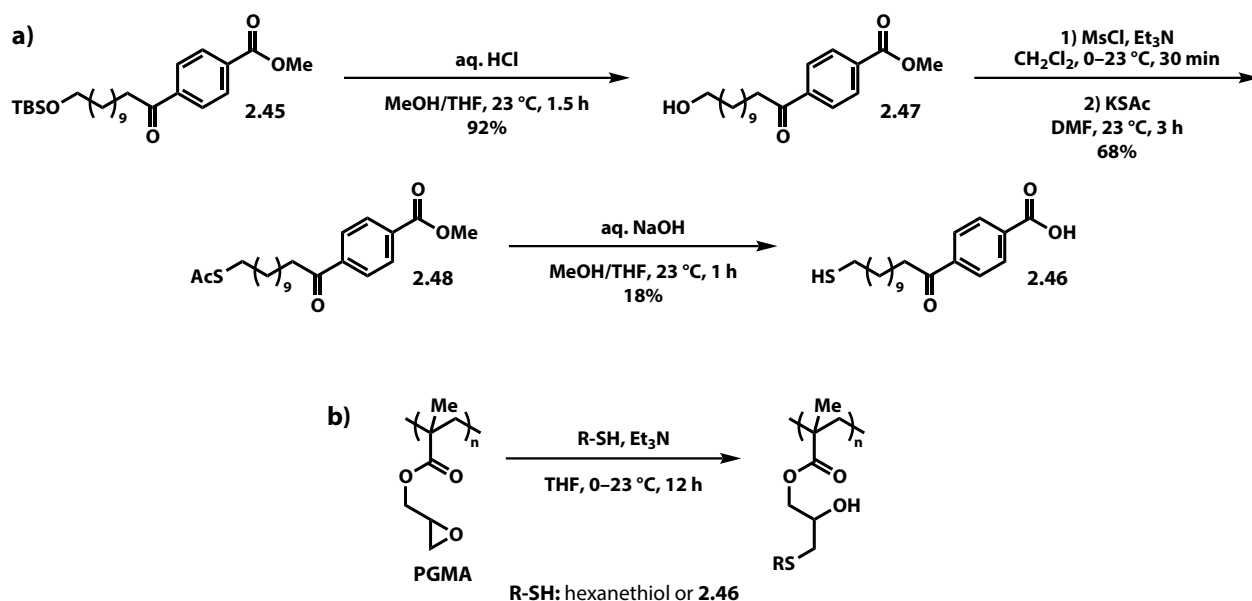
**Figure 2.17.** Post-polymerization modification via transesterification. a) Synthesis of simplified photo-cleavable reagent 2.43. b) General conditions for transesterification of poly(*tert*-butyl acrylate).

of the Grignard reagent of **2.16** and methyl 4-formylbenzoate. Oxidation of benzylic alcohol **2.44** with manganese dioxide afforded phenyl ketone **2.45**. Saponification of the ester followed by acidic workup afforded the desired reagent **2.42**.

Using the conditions found to be appropriate for transesterification of poly(*tert*-butyl acrylate) with *n*-butanol—*p*-TsOH in refluxing toluene—the post-polymerization modification reaction was carried out with phenyl ketone **2.42**. After refluxing overnight, a translucent solid was observed in the hot toluene mixture. After precipitation from ice cold ether and during sample preparation for NMR characterization it was found that the resultant material was insoluble in common NMR solvents (DMSO-*d*<sub>6</sub>, MeOD, CDCl<sub>3</sub>, and D<sub>2</sub>O). The material would swell in hot DMSO, but would not dissolve. Further attempts at dissolving the material in 1 M aqueous HCl, 1 M aqueous NaOH, concentrated ammonium hydroxide, acetonitrile, DMF, methanol, and isopropanol were unsuccessful. Given that this material was not formed in the test reactions with *n*-butanol nor without any nucleophile present, it was assumed that the desired modified polymer was indeed prepared. However, this could not be confirmed and the material was not useful towards the preparation of MS2 capsids with hydrophobic cores.

As an alternative to transesterification as a means to prepare a modified polyacrylate, thiol-epoxy opening as a method to modify poly(glycidyl methacrylate) was explored.<sup>64</sup> Nucleophilic ring opening of the epoxide with a thiol-containing reagent under basic conditions results in thioether formation and leaves a free alcohol near the backbone of the polymer (**Figure 2.18**). The free alcohol that is produced in this reaction could serve as a solubilizing group to eliminate the issue of total insolubility found for the transesterified polyacrylate discussed above. However, this free hydroxyl group could also decrease the hydrophobicity of the polymer after encapsidation and photocleavage. If the hydroxyl group was found to be useful for initial solubilization but too polar for creating a hydrophobic droplet within the MS2 capsid, tuning of this functionality would be possible through a second post-polymerization modification step: esterification of the alcohol.<sup>65</sup>

The thiol-containing reagent **2.46** was prepared from 11-bromoundecanol and methyl 4-formylbenzoate in much the same way as the hydroxyl analog **2.43** used for the attempted transesterification of PtBA. Silyl ether **2.45** was deprotected under acidic conditions to afford alcohol

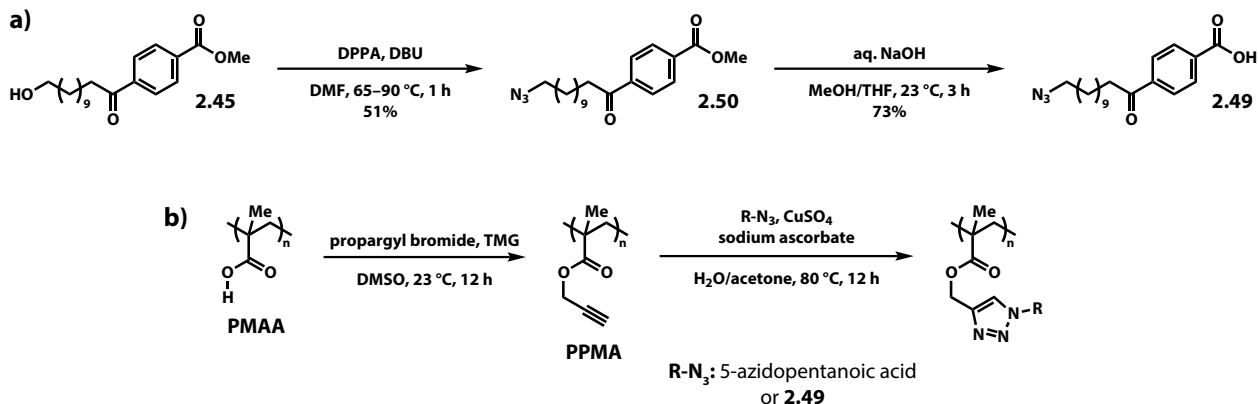


**Figure 2.18.** Post-polymerization modification via thiol-epoxide coupling. a) Synthesis of simplified photocleavable reagent 2.46. b) General conditions for transesterification of poly(glycidyl methacrylate).

2.47. Alcohol 2.47 was converted to the mesylate and then substituted with potassium thioacetate to afford the desired reagent in protected form. Thioacetate 2.48 was converted to the desired reagent through global deprotection of both the methyl ester and the thioacetate with sodium hydroxide in water/MeOH/THF.

Thiol-epoxy post-polymerization modification of PGMA with the thiol derivative of the photoswitchable carboxyphenyl ketone was attempted under the same conditions used for the test reactions. Much like in the case of transesterification of *PtBA* with reagent 2.43, a new material was isolated, which was again insoluble in water. Furthermore, the material was insoluble in 1 M aqueous HCl, 1 M aqueous NaOH, concentrated ammonium hydroxide, DMF, DMSO, acetonitrile, acetone, and isopropanol.

In a final effort to prepare a soluble hydrophobicity-modulating polymer, a copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) post-polymerization modification strategy was attempted (**Figure 2.19**). Preparation of azido-derivative 2.49 of the hydrophobicity modulation reagent was achieved in two steps from alcohol intermediate 2.47. Substitution of the primary alcohol



**Figure 2.19.** Post-polymerization modification via CuAAC. a) Synthesis of simplified photocleavable reagent 2.49. b) Preparation of poly(propargyl methacrylate) from poly(methacrylic acid) and general conditions for CuAAC.

in **2.47** with azide was achieved with diphenylphosphoryl azide and DBU in moderate yield. Treatment with NaOH in water/MeOH/THF afforded the desired reagent in good yield. Poly(propargyl methacrylate) was prepared via esterification of poly(methacrylic acid) with propargyl bromide and tetramethylguanidine, according to the protocol developed by Li *et al.*<sup>66</sup> Coupling of azide reagent **2.49** with the alkyne-containing polymer was attempted using conditions known to be compatible with ester-containing polymers.<sup>67</sup> Again, as in the cases of transesterification and thio-epoxy post-polymerization modification with this class of reagents, a new material was isolated and was found to be insoluble in every solvent tested.

Although *p*-carboxyphenyl ketone **2.43** and its derivatives **2.46** and **2.49** were found to be water soluble in the presence of 5% DMF cosolvent, polymers modified with these reagents through various post-polymerization modification strategies were found not only to be insoluble in water, but generally insoluble. While the prospect of using well-defined anionic polymers to template MS2 reassembly followed by switching to a hydrophobic state is highly appealing as a method to create a hydrophobic droplet within MS2 capsids, the preparation of appropriate polymeric materials has proven to be difficult. Although issues of solubility may be overcome by the elaboration of the *p*-carboxyphenyl ketone with water-solubility enhancing moieties (e.g., amide formation from the carboxyl group with taurine or bis(carboxymethyl)amine terminated amino-dPEG), this solution would come with the increased synthetic complexity of the required reagents. Perhaps the alternative strategy, of polymer grafting from the MS2 capsid protein using a minimally water-soluble and switchable monomer, is one more likely to produce materials with the desired properties.

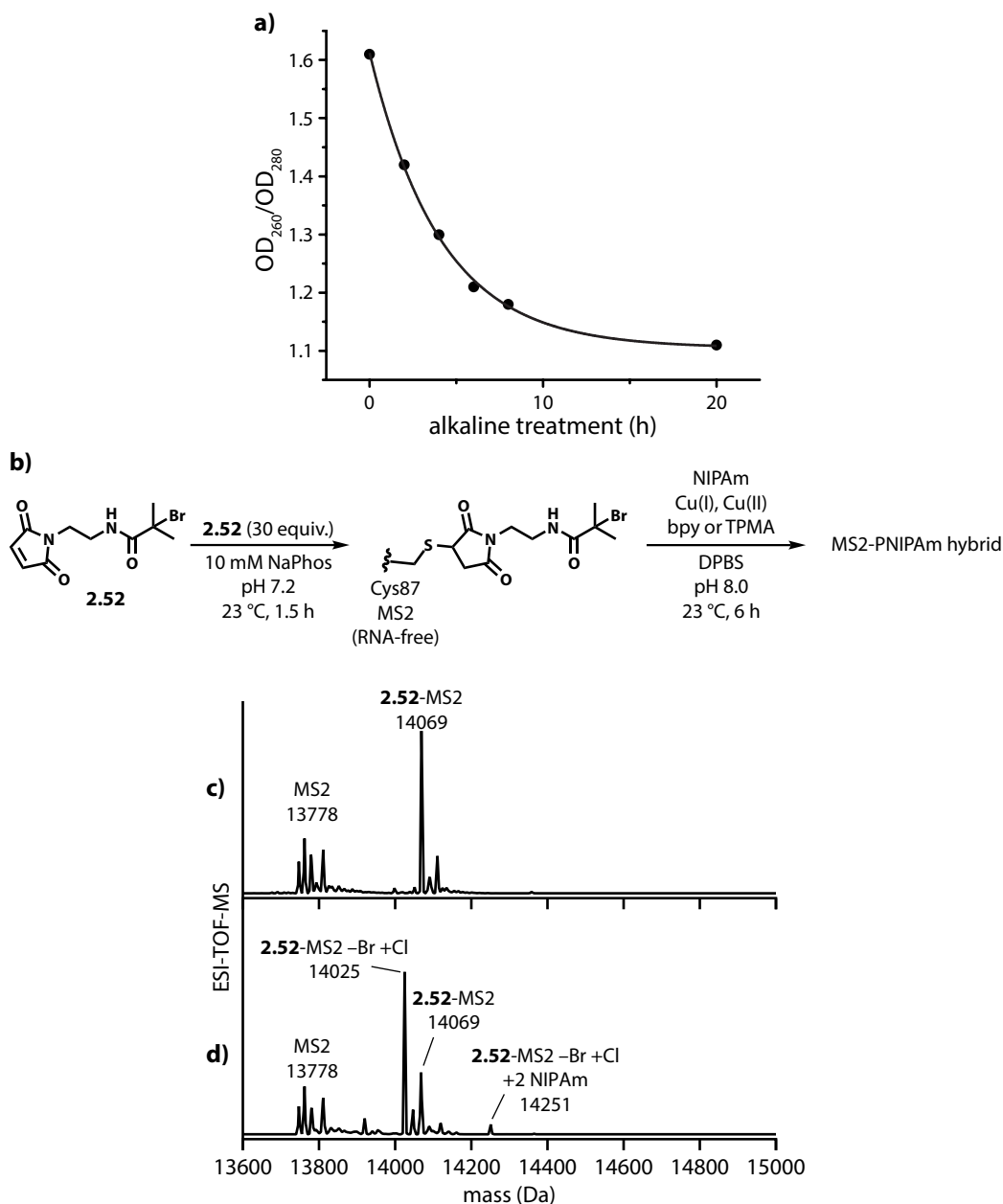
## 2.4 Towards hybrid hydrophobicity-modulating polymer-protein hybrids

Turning to the second general strategy for creating a hydrophobic droplet within the core of the MS2 capsid, attempts were made to graft a hydrophobicity modulating polymer from the interior surface of the capsid. Polymerization from the MS2 capsid offers a number of advantages over the encapsidation strategy discussed above. As no isolation nor purification of the polymer is required, only small quantities of switchable monomer are needed relative to standard polymerization, and the solubility of the pre-cleavage polymer is a non-issue as long as the hybrid material retains water solubility. Additionally, there is no loss of protein due to the inefficient disassembly-reassembly procedure.

Developments in the field of controlled radical polymerization in aqueous media have allowed for the preparation of novel protein-polymer hybrid materials through polymer growth from protein-based macroinitiators.<sup>68</sup> Numerous examples of protein-polymer conjugates prepared using a grafting-from strategy and either atom-transfer radical polymerization (ATRP) or reversible addition-fragmentation chain transfer polymerization (RAFT) as the specific method of controlled radical polymerization, have been reported. In line with the aims of this project, multiple examples of ATRP from the interior surface of VLPs, creating polymer-filled capsid core, have been reported. Lucon *et al.* have utilized ATRP to graft bisacrylamide-crosslinked poly(2-aminoethyl methacrylate) from the interior surface of P22.<sup>7</sup> High-density cargo loading was then achieved through acylation of the amine side-chains of the polymer-VLP hybrid with isocyanate-labeled gadolinium chelators for use as an MRI contrast agent. Using a similar strategy, photocatalytic hybrid P22 capsids were prepared by incorporating a Ru(methphen)<sub>3</sub><sup>2+</sup> tris-acrylamide into the polymer during the polymerization stage.<sup>69</sup> ATRP from the interior of VLPs on the same size scale

as MS2 has also been demonstrated. Hovlid *et al.* have prepared cationic polymer-filled Q $\beta$  hybrids from the polymerization of tertiary amine containing acrylate monomers and the resulting particles have been demonstrated to be more readily internalized by mammalian cells than do corresponding polymer-free VLPs.<sup>70</sup>

The proven utility of ATRP as a method for generating hybrid polymer-VLP materials provided sufficient precedent to attempt the preparation of a hybrid switchable hydrophobicity



**Figure 2.20.** Towards hydrophobicity modulating polymer-MS2 hybrid particles. a) Through cycles of alkaline treatment and precipitation, RNA was removed from MS2 capsids. b) MS2-macroinitiator was prepared by conjugation of maleimide **2.52** to Cys87. ATRP from the MS2-macroinitiator would result in a hybrid protein-polymer material. Representative ESI-TOF-MS characterization of the MS2-macroinitiator c) before and d) after ATRP. Halogen exchange and minimal oligomerization were observed.



modulating polymer-MS2 material. Preparation of such a material required 1) a water-soluble photoswitchable hydrophobicity-modulating acrylate monomer, 2) an MS2-macroinitiator, and 3) ATRP conditions optimized for this specific hybrid. Although this strategy requires significantly less monomer for preparation of sufficient polymeric material than does the disassembly-reassembly strategy discussed above, the monomer of choice should be synthetically tractable on the tens-of-milligrams scale. Therefore, a derivative of the simplified *p*-carboxyphenyl ketone was chosen over a derivative of the original Norrish type II substrate. Acylation of alcohol **2.43** with acryloyl chloride afforded the acrylate monomer **2.51** in sufficient yield. The monomer was found to be adequately water soluble in the presence of 5% DMSO cosolvent.

A cysteine-reactive ATRP initiator, 2-bromoisobutyryl aminoethyl maleimide (**2.52**), was readily prepared from *N*-(2-aminoethyl)maleimide (**1.7-Et**). Conjugation of maleimide **2.52** to Cys87 of the MS2 capsid protein proceeded smoothly with 30 equivalents of the maleimide at pH 7.2 over 2 h, resulting in roughly 75% modification by ESI-TOF-MS (**Figure 2.20**). To avoid wasting the semiprecious photoswitchable monomer **2.51**, test polymerizations were performed using *N*-isopropylacrylamide (NIPAM). NIPAM is a commonly-used monomer in protein-polymer hybrid formation using both ATRP and RAFT methods and is highly water-soluble. Additionally, due to the lower critical solution temperature (LCST) properties of poly(NIPAM), NIPAM-MS2 hybrids could have interesting properties. Lastly, inhibitor-free NIPAM can be readily prepared by recrystallization, and stores well.

Initially, polymerization from the MS2 macroinitiator was attempted with conditions found to be appropriate for P22-polymer hybrid construction.<sup>7</sup> These conditions, 482/1/43/43/177 monomer/macroinitiator/CuBr/CuBr<sub>2</sub>/2,2'-bpy in DPBS at pH 8.0, resulted in no discernable modification by ESI-TOF-MS or SDS-PAGE. Following this initial negative result, conditions identified from an in-depth screen performed in the Matyjaszewski lab of ATRP conditions for preparing protein-polymer hybrids were employed for grafting-from interior-directed MS2 macroinitiators.<sup>71</sup> A selection of the optimized conditions was screened, specifically using a monomer/macroinitiator/Cu(I)/Cu(II)/ligand ratio of 227/1/1/9/22 where the copper salts used were either CuCl and CuCl<sub>2</sub> or CuBr and CuBr<sub>2</sub> and the ligand used was either 2,2'-bipyridine or tris(2-pyridylmethyl)amine. Additionally, attempts were made to prepare the protein-polymer hybrid using a technique known as activator generated by electron transfer (AGET) ATRP, wherein Cu(I) is generated *in situ* with a reducing agent. In the case of AGET ATRP, a ratio of 227/1/10/11/0.1 monomer/macroinitiator/CuBr<sub>2</sub>/sodium ascorbate was used. Briefly, no gel shifts indicative of polymer formation were observed by SDS-PAGE for all conditions screened. However, some degree of bromide-chloride halogen exchange was observed by ESI-TOF-MS. This suggests that the catalysts were at least somewhat active.

In an attempt to improve the chances of successful polymerization, RNA was removed from the MS2 capsids by alkaline treatment. RNA was hydrolyzed within the MS2 capsids by incubation in high pH phosphate buffer (pH 11.8), and the capsids were purified by precipitation from ammonium sulfate solution.<sup>72</sup> The cycle was repeated until the decrease in the  $A_{260}/A_{280}$  ratio plateaued. RNA-free MS2 capsids were then modified with maleimide ATRP initiator **2.52**. After purification of the RNA-free macroinitiator, polymerization was attempted using the same conditions screened above. Again, some level of bromide-chloride exchange was observed by ESI-TOF-MS but no SDS-PAGE gel shifts were observed. However, in one case (polymerization with 227/1/1/9/22 NIPAM/RNA-free MS2-macroinitiator/CuCl/CuCl<sub>2</sub>/2,2'-bpy), initiation of chain growth was observed. In a small percentage of MS2 monomers (<2%), single and double addition of the NIPAM monomer

to the chain was observed by MS. This was an encouraging result, but optimization of this polymerization remains to be done.

While conditions were not found for the successful inward-directed polymerization of photoswitchable hydrophobicity-modulating acrylate **2.51**, let alone the model polymerization of NIPAM, from an MS2 macroinitiator, evidence for the eventual success of this strategy was gained. While optimization of the ratio of reaction components will be critical for forming polymer-protein hybrids of the highest quality, the main issue plaguing this initial screen may be the presence of dissolved oxygen. Oxygen is a known inhibitor of ATRP reactions, and attempts were made to eliminate it from the system prior to running polymerizations. Catalyst solutions were sparged with nitrogen for 0.5–1 h prior to addition into a solution of initiator and monomer, which had been put under an inert atmosphere by cycling between vacuum and nitrogen with gentle agitation. The most likely source of remaining oxygen in the system is the solution of monomer and protein macroinitiator, as only mild techniques can be used to remove dissolved gases due to the sensitive nature of the protein. One potential workaround to this problem is to use an enzymatic deoxygenation system such as that which was recently developed for screening RAFT conditions in open 96-well plates.<sup>61</sup> Further optimization in combination with improved deoxygenation techniques will lead to the preparation of this new class of MS2 hybrid materials for use as nanocarriers.

## 2.5 Conclusions

Two potential hydrophobicity modulating bond cleavage reactions were explored towards the development of general virus-like particle-based nanocarriers for hydrophobic actives. The first of these reactions, elimination of an allyl phosphate or allyl taurine carbamate through  $\pi$ -allylpalladium chemistry, was found to be ineffective. However, Norrish type II photocleavage of protein-bound aryl alkyl ketone reagents was successful. This transformation may be considered the first biocompatible bond cleavage reaction that does not result in the exposure of a water-solubilizing group. While this reaction was used to effect a significant increase in interior surface hydrophobicity, the resulting MS2 nanocarrier was not sufficiently hydrophobic to retain hydrophobic cargo.

In a push towards increasing the hydrophobic character of MS2 capsids even further, two strategies for creating a hydrophobic volume or droplet within the capsid were explored. Efforts were made to prepare hydrophobicity modulating polymers from derivatives of the Norrish type II substrate. However, solubility issues dampened progress in this area. The preparation of hydrophobicity modulating polymer-MS2 hybrid materials would circumvent such solubility issues, and initial studies towards the preparation of such materials were carried out. While significant optimization is still required, we are confident that these strategies will lead to new VLP-based formulations that could be used in applications ranging from drug delivery to agriculture.

## 2.6 Materials and Methods

### 2.6.1 General procedures and materials

Unless otherwise noted, all chemicals were purchased from commercial sources and used as received without further purification. Analytical thin layer chromatography (TLC) was performed

on EM Reagent 0.25 mm silica gel 60-F<sub>254</sub> plates and visualized by ultraviolet (UV) irradiation at 254 nm and staining with potassium permanganate. Purifications by flash silica gel chromatography were performed using EM silica gel 60 (230–400 mesh). Chromatography solvents were used without distillation. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H<sub>2</sub>O) used in all non-synthetic procedures was deionized using a NANO-pure<sup>TM</sup> purification system (Barnstead, USA). Polyacrylates were purchased from Polymer Source (Montreal, QC, Canada). Centrifugations were performed with an Eppendorf 5424 R (Eppendorf, Hauppauge, NY).

## 2.6.2 Instrumentation and sample analysis

**NMR.** <sup>1</sup>H, <sup>13</sup>C, and <sup>31</sup>P spectra were measured with a Bruker AVQ-400, AVB-400 (400 MHz), or AV-300 (300 MHz) spectrometer. <sup>1</sup>H NMR chemical shifts are reported as  $\delta$  in units of parts per million (ppm) relative to residual CHCl<sub>3</sub> ( $\delta$  7.26, singlet), H<sub>2</sub>O ( $\delta$  4.79, singlet), or DMSO ( $\delta$  2.50, pentet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (quintet), or br s (broad singlet). Coupling constants are reported as a *J* value in Hertz (Hz). The number of protons (*n*) for a given resonance is indicated as *n*H and is based on spectral integration values. <sup>13</sup>C NMR chemical shifts are reported as  $\delta$  in units of parts per million (ppm) relative to CDCl<sub>3</sub> ( $\delta$  77.16, triplet) or DMSO-d<sub>6</sub> ( $\delta$  39.52, septet).

**Mass Spectrometry.** Small molecules were analyzed on an AB Sciex 3200 Qtrap (AB Sciex, USA) equipped with an ESI Turbo V ion source connected in-line with an Agilent 1100 series high-performance liquid chromatograph (HPLC) (Agilent Technologies, USA). Chromatography was performed using a Chromolith RP-18 column (Merck, Darmstadt, Germany) with a H<sub>2</sub>O/MeCN gradient mobile phase containing 0.1% formic acid. Proteins and protein conjugates were analyzed on an Agilent 6224 Time-of-Flight (TOF) mass spectrometer with a dual electrospray source connected in-line with an Agilent 1200 series HPLC (Agilent Technologies, USA). Chromatography was performed using either a Poroshell 5 $\mu$  C18 300 Å (Agilent Technologies, USA) or Proswift RP-4H (Thermo Scientific, USA) column with a H<sub>2</sub>O/MeCN gradient mobile phase containing 0.1% formic acid. Mass spectra of proteins and protein conjugates were deconvoluted with the MassHunter Qualitative Analysis Suite B.05 (Agilent Technologies, USA).

**High Performance Liquid Chromatography (HPLC).** HPLC was performed on Agilent 1100 series HPLC systems (Agilent Technologies, USA) equipped with in-line diode array detector (DAD) and fluorescence detector (FLD). Analytical and preparative reverse-phase HPLC of small molecules was accomplished using Gemini C18 columns (Phenomenex, Torrance, CA) and a H<sub>2</sub>O/MeCN gradient mobile phase containing 0.1% TFA. Size exclusion chromatography was accomplished on a Phenomenex PolySep-GFC-P 5000 or BioSep SEC-S-4000 column using an aqueous mobile phase (50 mM Tris-HCl, 100 mM NaCl, pH 7.2).

**Ultraviolet-Visible (UV-Vis) Spectrometry.** UV-Vis spectrometry was performed using 3x3 mm Suprasil quartz cuvettes (Hellma Analytics, Germany) with a Varian Cary 50 spectrophotometer (Varian Inc., Palo Alto, CA). All measurements were collected with a scan rate of 80 nm/s and were baseline corrected. Small-scale UV-Vis spectrometry was performed using a Nanodrop 1000 (Thermo Scientific, USA).

**Spectrofluorometry.** Fluorescence measurements were performed using 3x3 mm Suprasil quartz cuvettes (Hellma Analytics, Germany) with a Fluoromax-4 spectrofluorometer equipped with a Peltier temperature controller (HORIBA Jobin Yvon Inc., Edison, NJ). Samples were diluted to an OD  $\leq$  0.03 and filtered through a 0.22  $\mu$ m centrifugal filter prior to analysis. All measurements were collected with an integration time of 0.1 s.

**Dynamic Light Scattering (DLS).** DLS measurements were obtained using a Malvern Instruments Zetasizer Nano ZS (Malvern Instruments, United Kingdom). Samples were filtered through 0.22  $\mu$ m centrifugal filter units (Millipore Corporation, Billerica, MA) prior to data collection.

### 2.6.3 Experimental procedures

**Expression and purification of MS2 mutants.** Bacteriophage MS2 T19pAF N87C plasmid production, expression, and purification has been previously reported.<sup>73,74</sup> Bacteriophage MS2 T19Y N87C plasmid production, expression, and purification has been previously reported.<sup>75</sup>

**Expression and purification of Mth1491 A30C.** Mth1491 was expressed and purified according to a modified literature procedure.<sup>76</sup> BL21(Gold $\lambda$ DE3) cells were transformed with a pJexpress404 plasmid containing the ampicillin-resistant gene ampR and the Mth1491 A30C gene with an N-terminal His<sub>6</sub>-tag followed by a thrombin cleavage site (DNA2.0 Inc., Menlo Park, CA). Colonies were selected for inoculation of lysogeny broth cultures with 100  $\mu$ g/mL ampicillin. When cultures reached mid-log phase as determined by OD<sub>600</sub>, expression was induced by addition of 10  $\mu$ M IPTG. Cultures were grown for 14–18 h at 37 °C and cells were isolated by centrifugation. Cells were resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0). Cells were lysed by sonication (Fisher Scientific, USA) and cleared by centrifugation. The His<sub>6</sub>-Mth1491 mutant was purified from the cleared lysate by affinity chromatography on Ni-NTA agarose, following the manufacturer's protocol (Qiagen, USA). The purified protein was buffer-exchanged into 50 mM Tris-HCl, pH 8.0 and then into cleavage buffer (50 mM Tris-HCl, pH 8.0 with 10 mM CaCl<sub>2</sub>). The His<sub>6</sub>-tag was cleaved using Thrombin CleanCleave (Sigma-Aldrich, USA), incubating for 24 h at 4 °C. Mth1491 was isolated from the cleaved His<sub>6</sub>-tag and residual His<sub>6</sub>-Mth1491 by buffer exchange into 50 mM Tris-HCl, pH 8.0 and then into lysis buffer and performing a pull-down with Ni-NTA agarose. Collection of the supernatant yielded 20–40 mg of protein per liter of culture. Isolated protein was exchanged into 25 mM potassium phosphate, pH 6.5 with 0.01% NaN<sub>3</sub> and stored at 4 °C.

**Cysteine alkylation of MS2 with maleimide 2.1.** An Eppendorf tube was charged with a solution of (T19pAF N87C or T19Y N87C) MS2 (25  $\mu$ M final concentration, 10 mM sodium phosphate, pH 7.2). To this solution was added maleimide 2.1 (50 mM DMF stock solution, 2 mM final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 100 kDa MWCO filters with 10 mM sodium phosphate, pH 7.2.

**Cysteine alkylation of Mth1491 A30C with maleimide 2.1.** An Eppendorf tube was charged with a solution of Mth1491 A30C (25  $\mu$ M final concentration, 25 mM potassium phosphate, pH 6.5). To this solution was added maleimide 2.1 (50 mM DMF stock solution, 2 mM final concentration).

The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 10 kDa MWCO filters with 25 mM potassium phosphate, pH 6.5.

**General procedure for screening  $\pi$ -allylpalladium conditions.** MS2-2.1 or Mth1491-2.1 (diluted to 10  $\mu$ M final concentration with 100 mM sodium phosphate, pH 8.0 or 9.0, or 100 mM  $\text{NH}_4\text{HCO}_3$ , pH 8 or 9) was treated an aliquot of catalyst solution (10 or 100 equiv. of 1:12  $\text{Pd}(\text{OAc})_2/\text{ADHP}$ , 1:12  $\text{Pd}(\text{OAc})_2/\text{DMADHP}$ , 1:12  $\text{Pd}(\text{OAc})_2/\text{DMG}$ , 1:12  $\text{Pd}(\text{OAc})_2/\text{TMG}$ , 1:12  $\text{Pd}(\text{OAc})_2/\text{TPPTS}$ , 1:12  $\text{Pd}(\text{OAc})_2/\text{TPPDS}$ ,  $\text{Pd}(\text{NO}_3)_2$ , or  $\text{K}_2\text{PdCl}_4$  in  $\text{ddH}_2\text{O}$ ) at room temperature or 37  $^\circ\text{C}$ . After 1.5 h, a solution of MPrAc diad **2.11** (500 equiv.) was added and the reaction was incubated at 37  $^\circ\text{C}$  for 0.5 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 100 kDa MWCO filters with 10 mM potassium phosphate, pH 7.2. Samples were analyzed by ESI-TOF-MS.

**Suzuki coupling on Mth1491 A30C.** An Eppendorf tube was charged with a solution of Mth1491 A30C (25  $\mu$ M final concentration, 25 mM potassium phosphate, pH 6.5). To this solution was added *p*-iodophenethyl maleimide **2.12** (50 mM DMF stock solution, 2 mM final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 1 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 10 kDa MWCO filters with 100 mM sodium phosphate, pH 8.0.

To a solution of Mth1491 A30C-**2.12** (10  $\mu$ M final concentration, 100 mM sodium phosphate, pH 8.0) and phenylboronic acid (100 equiv.) was added 1:12  $\text{Pd}(\text{OAc})_2/\text{DMG}$  catalyst solution (50 equiv. Pd). The resulting mixture was briefly vortexed and the reaction was incubated at 37  $^\circ\text{C}$  for 1 h. Bis(thioether) **2.11** (500 equiv.) was added and the reaction was incubated at 37  $^\circ\text{C}$  for 0.5 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 10 kDa MWCO filters with 25 mM potassium phosphate, pH 6.5.

**Cysteine alkylation of MS2 with maleimide 2.15.** An Eppendorf tube was charged with a solution of (T19Y N87C) MS2 (25  $\mu$ M final concentration, 10 mM sodium phosphate, pH 7.2). To this solution was added maleimide **2.15** (50 mM DMF stock solution, 1 mM final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 4 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 100 kDa MWCO filters with 10 mM sodium phosphate, pH 7.2.

**Cysteine alkylation of MS2 with maleimide 2.26.** An Eppendorf tube was charged with a solution of (T19Y N87C) MS2 (25  $\mu$ M final concentration, 10 mM sodium phosphate, pH 7.2). To this solution was added maleimide **2.26** (50 mM DMF stock solution, 1 mM final concentration). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 12 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 100 kDa MWCO filters with 10 mM sodium phosphate, pH 7.2.

**Photocleavage of Norrish type II substrates.** Photocleavage was performed in 3x3 mm Supracil quartz cuvettes (Hellma Analytics, Germany) using the 150 W ozone-free xenon arc lamp of a Fluoromax-4 spectrofluorometer equipped with a Peltier temperature controller (HORIBA Jobin Yvon Inc., Edison, NJ).

Solutions of unconjugated hydrophobicity modulators (**2.15**, **2.26**) (25  $\mu\text{M}$  final concentration, 10 mM sodium phosphate, pH 7.2) were treated with 2 equivalents of MESNa in 10 mM for 30 min prior to irradiation. Modulator **2.15** was irradiated at 300 nm with a 20 nm excitation slit width at 23  $^{\circ}\text{C}$  for 1 h. Modulator **2.26** was irradiated at 340 nm with a 20 nm excitation slit width at 23  $^{\circ}\text{C}$  for 3 h. Reaction progress was monitored by LC-MS in negative mode.

Unmodified MS2 controls and MS2-hydrophobicity modulator conjugates were adjusted to 25  $\mu\text{M}$  monomer in 10 mM sodium phosphate, pH 7.2. MS2-**2.15** conjugates were irradiated at 300 nm with a 20 nm excitation slit width at 23  $^{\circ}\text{C}$  for 4 h. MS2-**2.26** conjugates were irradiated at 340 nm with a 20 nm excitation slit width at 23  $^{\circ}\text{C}$  for 4 h. After irradiation, aryl ketone byproducts were removed through successive rounds of centrifugal filtration in 100 kDa MWCO filters with 10 mM sodium phosphate, pH 7.2. Irradiated samples were then characterized by ESI-TOF-MS, SEC, and DLS.

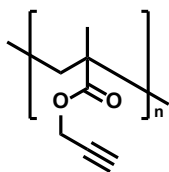
**Calculation of theoretical UV-Vis spectra.** Theoretical UV-Vis spectra of 4-acetylphenyl hydrogen phosphate, 6-acetylnaphthalen-2-yl hydrogen phosphate, and 6-acetylpyren-1-yl hydrogen phosphate were calculated using MacroModel (Schrödinger, LLC, USA). Time-dependent density functional theory (TDDFT) calculations were performed for the ground state and 20 excited states using the B3LYP functional and 6-31G\*\* basis set. The PBF solvation model was used.

**Spectrofluorometric determination surface hydrophobicity.** Stock solutions of ANS (8 mM, 10 mM sodium phosphate, pH 7.2), Prodan (3  $\mu\text{M}$ , 10 mM sodium phosphate, pH 7.2), and Nile Red (20  $\mu\text{M}$  in MeOH) were prepared. For ANS, 5  $\mu\text{L}$  of ANS stock solution was added to successive samples containing 0, 2.5, 5, 10, and 20  $\mu\text{M}$  protein (T19Y N87C MS2, or T19Y N87C MS2 irradiated at 300 $\pm$ 10 nm for 4 h) or protein conjugate (MS2-**2.15** or MS2-**2.15** after photocleavage) in 10 mM sodium phosphate, pH 7.2 (100  $\mu\text{L}$ ). For Prodan and Nile Red, 5  $\mu\text{L}$  of Prodan or Nile Red stock solution was added to successive samples containing 0, 0.625, 1.25, 2.5, and 5  $\mu\text{M}$  protein (T19Y N87C MS2) or protein conjugate (MS2-**2.15** or MS2-**2.15** after photocleavage) in 10 mM sodium phosphate, pH 7.2 (100  $\mu\text{L}$ ). Samples were briefly vortexed and incubated for 10 min at room temperature prior to analysis. Emission spectra were collected for all samples, exciting at 390, 365, and 560 nm for ANS, Prodan, and Nile Red, respectively, at 25  $^{\circ}\text{C}$ .

Emission intensities at 470, 465, and 623 nm for ANS, Prodan, and Nile Red, respectively, were used for surface hydrophobicity determination. The emission intensity of the control samples (0  $\mu\text{M}$  protein) for each probe were subtracted from the observed emission intensities to give relative fluorescence intensity (RFI) values. These RFI values were then plotted against final protein concentration. Linear regression was then conducted to determine the slopes of these curves, which corresponded to the surface hydrophobicity ( $S_0$ ) of the samples.  $S_0$  values were then normalized to that of unmodified protein for each probe. The reported uncertainty represents the standard error of the slope from the linear regression operation.

**Tranesterification of poly(*tert*-butyl acrylate).** A screw cap vial was charged with poly(*tert*-butyl acrylate) ( $M_n = 15$  kDa,  $M_w = 19$  kDa, 50 mg, 0.39 unit mmol) and butanol (0.18 mL, 2.0 mmol, 5 equiv.) or **2.43** (125 mg, 2.0 mmol, 5 equiv.). Toluene (1 mL) and *p*-TsOH monohydrate (74 mg, 0.39 mmol, 1 equiv.) were added, the vial was sealed, and the mixture was heated to 110  $^{\circ}\text{C}$  for 16 h. After cooling to room temperature, the polymer was precipitated into ice cold Et<sub>2</sub>O (45 mL) and isolated by centrifugation.

**Post-polymerization modification of poly(glycidyl methacrylate).** To a solution of poly(glycidyl methacrylate) ( $M_n = 15$  kDa,  $M_w = 24$  kDa, 20 mg, 0.14 unit mmol) and hexanethiol (0.04 mL, 0.28 mmol, 1.25 equiv.) or **2.46** (94 mg, 0.28 mmol, 1.25 equiv.) in THF (1 mL) was added  $\text{Et}_3\text{N}$  (0.02 mL, 0.07 mmol, 0.5 equiv.) at 0 °C. The reaction was warmed to room temperature and stirred for 12 h. The polymer was precipitated into ice cold  $\text{Et}_2\text{O}$  (45 mL) and collected by centrifugation.



**Synthesis of poly(propargyl methacrylate).** To a solution of PMAA (syndiotactic rich,  $M_n = 30$  kDa,  $M_w = 32.5$  kDa, 100 mg, 1.16 unit mmol) in dry DMSO (2 mL) was added 1,1,3,3-tetramethylguanidine (0.29 mL, 2.32 mmol, 2 equiv.) and propargyl bromide (80% in toluene, 0.2 mL, 1.74 mmol, 1.5 equiv.) at room temperature. After 2 h, the reaction was neutralized with 1 M aqueous acetic acid and the resulting polymer was precipitated into cold MeOH (40 mL). The crude polymer was dissolved in THF and precipitated into cold MeOH (40 mL) to give pure poly(propargyl methacrylate). Near quantitative conversion (>97.5%) to the desired ester was determined by  $^1\text{H}$  NMR integration of the ester protons ( $\delta$  4.61, 1.95H) and alkynyl proton ( $\delta$  2.62–2.46, 0.99H) relative to backbone methyl protons ( $\delta$  1.39–0.78, 3H).

**CuAAC coupling on polymer.** To a solution of poly(propargyl methacrylate) (29 mg, 0.14 unit mmol) in 2:1  $\text{H}_2\text{O}$ /acetone (500  $\mu\text{L}$ ) was added **2.49** (94 mg, 0.28 mmol, 2 equiv.),  $\text{CuSO}_4$  (11 mg, 0.07 mmol, 0.5 equiv.), and sodium ascorbate (28 mg, 0.14 mmol, 1 equiv.) at room temperature. The reaction was heated to 80 °C in a sealed vial and stirred for 12 h. The polymer was precipitated into ice cold MeOH (45 mL) and collected by centrifugation.

**ATRP grafting-from MS2.** MS2-macroinitiator was prepared by removing encapsidated RNA followed by coupling of maleimide **2.52** to Cys87. Encapsidated RNA was hydrolyzed by treatment of T19Y N87C MS2 capsids (72  $\mu\text{M}$ , 10 mM sodium phosphate, pH 7.2) with  $\frac{1}{4}$  volume of 500 mM sodium phosphate, pH 11.8 buffer at room temperature. After 2 h, capsids were isolated by precipitation through a 2x dilution with saturated aqueous  $\text{NH}_4\text{SO}_4$ . After incubating for 30 min, precipitated capsids were collected by centrifugation and resuspended in 10 mM sodium phosphate, pH 7.2 to a final concentration of 72  $\mu\text{M}$ . The hydrolysis/precipitation cycle was repeated 4 times followed by a 12 h hydrolysis step and final precipitation. Progress was monitored by the decrease in  $A_{260}:A_{280}$  ratio by Nanodrop.

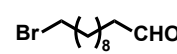
RNA-free capsids were resuspended in 10 mM sodium phosphate, pH 7.2 (25  $\mu\text{M}$  final concentration). To this solution was added maleimide **2.52** (50 mM DMF stock solution, 30 equiv.). The resulting mixture was briefly vortexed and the reaction was incubated at room temperature for 2 h. The reaction mixture was then purified through successive rounds of centrifugal filtration in 100 kDa MWCO filters with DPBS, pH 8.0 to yield MS2-macroinitiator.

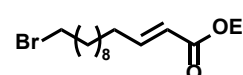
MS2-macroinitiator in DPBS, pH 8.0 (50  $\mu\text{M}$ , 250  $\mu\text{L}$ ) was added to a screw cap vial with septum charged with NIPAM (0.3 mg, 227 equiv.). This solution was cycled between gentle vacuum and  $\text{N}_2$  a minimum of 5 times. Stock catalyst solution was prepared by addition of dd- $\text{H}_2\text{O}$  (10 mL), which had been sparged with  $\text{N}_2$  for a minimum of 30 min, to a vial containing Cu(I) ( $\text{CuCl}$  or  $\text{CuBr}$ , 0.05 mmol), Cu(II) ( $\text{CuCl}_2$  or  $\text{CuBr}_2$ , 0.45 mmol), and ligand (bpy or TPMA, 1.1 mmol) under  $\text{N}_2$ . The stock catalyst solution was sonicated under  $\text{N}_2$  for 5 min prior to addition to the reaction vessel. To the initiator solution was added catalyst solution (5  $\mu\text{L}$ ) under  $\text{N}_2$ , reaching a final ratio of 227/1/1/9/22 monomer/initiator/Cu(I)/Cu(II)/ligand. After incubating for 6 h

at room temperature, the reaction was quenched by exposure to air and the reaction mixture was purified through successive rounds of centrifugal filtration in 10 kDa MWCO filters with 10 mM sodium phosphate, pH 7.2 and analyzed by ESI-TOF-MS and SDS-PAGE.

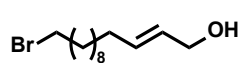
**AGET ATRP grafting-from MS2.** The same general procedure was followed as for ATRP grafting-from MS2. Pre-catalyst solution was prepared by addition of dd-H<sub>2</sub>O (10 mL), which had been sparged with N<sub>2</sub> for a minimum of 30 min, to a vial containing CuBr<sub>2</sub> (112 mg, 0.5 mmol) and bpy (86 mg, 0.55 mmol) under N<sub>2</sub>. The stock pre-catalyst solution was sonicated under N<sub>2</sub> for 5 min prior to addition to the reaction vessel. To the initiator solution was added pre-catalyst solution (5  $\mu$ L) under N<sub>2</sub>, followed by a solution of sodium ascorbate (500  $\mu$ M, 5  $\mu$ L) in N<sub>2</sub>-sparged dd-H<sub>2</sub>O reaching a final ratio of 227/1/10/11/0.1 monomer/initiator/Cu(II)/bpy/ascorbate. After incubating for 6 h at room temperature, the reaction was quenched by exposure to air and the reaction mixture was purified through successive rounds of centrifugal filtration in 10 kDa MWCO filters with 10 mM sodium phosphate, pH 7.2 and analyzed by ESI-TOF-MS and SDS-PAGE.

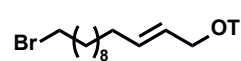
## 2.6.4 Small molecule synthesis

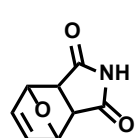
 **Synthesis of 11-bromoundecanal (2.2).** To a solution of 11-bromoundecanol (5 g, 19.9 mmol, 1 equiv.) and TEMPO (31 mg, 0.20 mmol, 0.01 equiv.) in CH<sub>2</sub>Cl<sub>2</sub> (30 mL) was added a solution of KBr (2.61 g, 21.9 mmol, 1.1 equiv.) in 5% aqueous NaHCO<sub>3</sub> (30 mL) at 0 °C. To this mixture was added 1 M aqueous NaOCl (39.8 mL, 39.8 mmol, 2 equiv.) dropwise with vigorous stirring. After 2 h, the reaction was warmed to room temperature and the layers were separated. The aqueous layer was extracted twice with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with saturated aqueous NaHCO<sub>3</sub>, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was removed under reduced pressure. The crude liquid was then purified via flash silica gel chromatography, eluting with 10% EtOAc/hexanes to afford a clear colorless liquid (4.41 g, 89%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  9.76 (t,  $J$  = 1.9 Hz, 1H), 3.40 (t,  $J$  = 6.9 Hz, 2H), 2.42 (td,  $J$  = 7.4, 1.9 Hz, 2H), 1.84 (p,  $J$  = 6.9 Hz, 2H), 1.62 (p,  $J$  = 7.2 Hz, 2H), 1.47–1.36 (m, 2H), 1.31–1.27 (m, 10H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  203.05, 44.04, 34.19, 32.93, 29.47, 29.42, 29.26, 28.85, 28.27, 22.18.

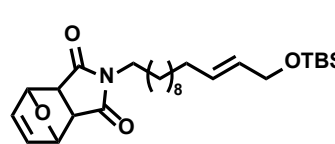
 **Synthesis of ethyl (*E*)-13-bromotridec-2-enoate (2.3).** Triethylphosphonoacetate (1.88 mL, 9.47 mmol, 1 equiv.) was slowly added to a suspension of NaH (60% dispersion in mineral oil, 379 mg, 9.47 mmol, 1 equiv.) in dry THF (25 mL) at 0 °C under N<sub>2</sub> and stirred for 10 min. A solution of 11-bromoundecanal (2.36 g, 9.47 mmol, 1 equiv.) in dry THF (10 mL) was slowly added to the mixture at 0 °C and the reaction was warmed to room temperature and stirred for 3 h. The reaction was cooled to 0 °C and quenched with ice cold H<sub>2</sub>O (35 mL) and extracted in Et<sub>2</sub>O three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was removed under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10% EtOAc/hexanes to afford a clear colorless liquid (2.8 g, 93%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  6.95 (dt,  $J$  = 15.6, 6.9 Hz, 1H), 5.80 (dd,  $J$  = 15.7, 1.7 Hz, 1H), 4.17 (q,  $J$  = 7.1 Hz, 2H), 3.40 (t,  $J$  = 6.9 Hz, 2H), 2.18 (qd,  $J$  = 7.1, 1.6 Hz, 2H), 1.84 (p,  $J$  = 7.0 Hz, 2H), 1.48–1.36 (m, 4H), 1.35–1.21 (m, 13H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  166.90, 149.57, 121.34, 60.25, 34.18, 32.94, 32.31, 29.51, 29.46, 29.23, 28.86, 28.28, 28.12, 14.42.



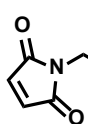
 **Synthesis of (E)-13-bromotridec-2-en-1-ol (2.4a).** To a solution of ethyl (E)-13-bromotridec-2-enoate (2.80 g, 8.77 mmol, 1 equiv.) in dry CH<sub>2</sub>Cl<sub>2</sub> (25 mL) was added a solution of DIBAL-H in hexanes (1.0 M, 26.3 mL, 26.3 mmol, 3 equiv.) at -78 °C under N<sub>2</sub>. The reaction was warmed to room temperature and stirred for 1.5 h. The solution was cooled to 0 °C and carefully quenched with an aqueous solution of potassium sodium tartrate (25 w/v%, 2.6 mL). After warming to room temperature, additional tartrate solution (15 mL) was added and the mixture was diluted with Et<sub>2</sub>O, filtered through a pad of Celite, and the layers were separated. The organic layer was washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was removed under reduced pressure. The crude product was then purified via flash silica gel chromatography, eluting with 20% EtOAc/hexanes to afford a greasy white solid (2.07 g, 85%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.76–5.57 (m, 2H), 4.11–4.06 (m, 2H), 3.40 (t, *J* = 6.9 Hz, 2H), 2.03 (q, *J* = 6.7 Hz, 2H), 1.85 (p, *J* = 7.0 Hz, 2H), 1.49–1.19 (m, 14H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 133.71, 128.94, 64.01, 34.23, 32.97, 32.35, 29.60, 29.57, 29.56, 29.29, 29.25, 28.90, 28.31.

 **Synthesis of (E)-((13-bromotridec-2-en-1-yl)oxy)(tert-butyl)dimethylsilane (2.5).** *tert*-Butyldimethylsilyl chloride (1.69 g, 11.20 mmol, 1.5 equiv.) was added to a solution of imidazole (1.27 g, 18.67 mmol, 2.5 equiv.) and (E)-13-bromotridec-2-en-1-ol (2.07 g, 7.47 mmol, 1 equiv.) in dry THF (50 mL) and stirred for 3 h at room temperature under N<sub>2</sub>. The mixture was cooled to 0 °C and quenched with saturated aqueous solution of NaHCO<sub>3</sub> (50 mL), warmed to room temperature and stirred for 30 min. The quenched reaction mixture was extracted three times into Et<sub>2</sub>O. The organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was removed under reduced pressure. The crude material was then purified via flash silica gel chromatography, eluting with 10% EtOAc/hexanes to afford a clear colorless oil (2.83 g, 97%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.63 (dtt, *J* = 15.3, 6.5, 1.4 Hz, 1H), 5.52 (dtt, *J* = 15.3, 5.3, 1.3 Hz, 1H), 4.12 (dd, *J* = 5.3, 1.2 Hz, 2H), 3.40 (t, *J* = 6.9 Hz, 2H), 2.01 (q, *J* = 6.9 Hz, 2H), 1.85 (p, *J* = 6.9 Hz, 2H), 1.47–1.20 (m, 14H), 0.90 (s, 9H), 0.07 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 131.70, 129.21, 64.27, 34.21, 32.99, 32.33, 29.63, 29.59, 29.57, 29.32, 28.91, 28.32, 26.15, 18.60, -4.93.

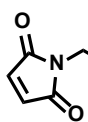
 **Synthesis of 3a,4,7,7a-tetrahydro-1H-4,7-epoxyisoindole-1,3(2H)-dione (2.6).** To a solution of maleimide (2.0 g, 20.6 mmol, 1 equiv.) in Et<sub>2</sub>O was added furan (5.0 mL, 68.7 mmol, 3.33 equiv.) at room temperature under N<sub>2</sub>. After 2 d, the reaction was filtered to afford the desired product as a white solid (3:1 *exo/endo*, 2.37 g, 70%). Isolation of the *endo* and *exo* isomers was possible via flash silica gel chromatography, eluting with 20–33% EtOAc/hexanes, however the mixture of isomers was used without further purification. *Endo* isomer: *R*<sub>f</sub> = 0.48 (50% EtOAc/Hexanes) (2.89 g, 85%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.87 (br s, 1H), 6.50 (t, *J* = 0.9 Hz, 2H), 5.34–5.32 (m, 2H), 3.56 (dd, *J* = 3.6, 1.7 Hz, 2H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 176.8, 134.6, 78.7, 47.4. *Exo* isomer: *R*<sub>f</sub> = 0.41 (50% EtOAc/Hexanes). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.27 (br s, 1H), 6.52 (t, *J* = 0.9 Hz, 2H), 5.31 (t, *J* = 0.8 Hz, 2H), 2.89 (s, 2H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 177.9, 136.7, 80.6, 48.6.

 **Synthesis of (E)-2-(13-((tert-butyl)dimethylsilyloxy)tridec-11-en-1-yl)-3a,4,7,7a-tetrahydro-1H-4,7-epoxyisoindole-1,3(2H)-dione (2.7).** Solid K<sub>2</sub>CO<sub>3</sub> (200 mg, 2.55 mmol, 5 equiv.) was added to a mixture of 3a,4,7,7a-tetrahydro-1H-4,7-epoxyisoindole-1,3(2H)-dione (mixture of *endo* and *exo* isomers, 84 mg, 0.511 mmol, 1 equiv.) and (E)-((13-bromotridec-2-en-1-

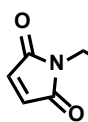
yl)oxy)(*tert*-butyl)dimethylsilane (200 mg, 0.511 mmol, 1 equiv.) in dry DMF (5 mL) at room temperature under N<sub>2</sub> and the reaction was then heated to 50 °C for 1.5 h. Volatiles were removed by rotary evaporation and the crude mixture was taken up in CHCl<sub>3</sub> and purified via flash silica gel chromatography, eluting with 20–33% EtOAc/hexanes to afford the isolated *exo* (106 mg, 44%) and *endo* (43 mg, 18%) isomers of the desired product as clear colorless oils (62% overall yield). *Endo* isomer:  $R_f = 0.30$  (20% EtOAc/Hexanes) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.39 (t,  $J = 1.0$  Hz, 2H), 5.62 (dtt,  $J = 15.3, 6.4, 1.3$  Hz, 1H), 5.52 (dtt,  $J = 15.3, 5.0, 0.9$  Hz, 1H), 5.36–5.27 (m, 2H), 4.16–4.07 (m, 2H), 3.55–3.45 (m, 2H), 3.29 (t,  $J = 7.5$  Hz, 2H), 2.01 (q,  $J = 7.0$  Hz, 2H), 1.46–1.15 (m, 16H), 0.90 (s, 9H), 0.06 (s, 6H). *Exo* isomer:  $R_f = 0.26$  (20% EtOAc/Hexanes) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.51 (t,  $J = 1.0$  Hz, 2H), 5.62 (dtt,  $J = 15.3, 6.4, 1.3$  Hz, 1H), 5.52 (dtt,  $J = 15.3, 5.0, 0.9$  Hz, 1H), 5.26 (t,  $J = 1.0$  Hz, 2H), 4.16–4.05 (m, 2H), 3.45 (t,  $J = 7.4$  Hz, 2H), 2.83 (s, 2H), 2.00 (q,  $J = 7.0$  Hz, 2H), 1.61–1.45 (m, 2H), 1.41–1.17 (m, 14H), 0.90 (s, 9H), 0.06 (s, 6H).



**Synthesis of (*E*)-1-(13-((*tert*-butyl)dimethylsilyloxy)tridec-11-en-1-yl)-1*H*-pyrrole-2,5-dione (2.8).** A solution of (*E*)-2-(13-((*tert*-butyl)dimethylsilyloxy)tridec-11-en-1-yl)-3a,4,7,7a-tetrahydro-1*H*-4,7-epoxyisoindole-1,3(2*H*)-dione (mixture of isomers, 1.19 g, 2.50 mmol) in dry toluene (20 mL) was heated to reflux under N<sub>2</sub> for 17 h. Volatiles were removed by rotary evaporation and the crude mixture was purified via flash silica gel chromatography, eluting with 5% EtOAc/hexanes to afford a clear colorless oil (810 mg, 80%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.67 (s, 2H), 5.63 (dtt,  $J = 15.3, 6.5, 1.3$  Hz, 1H), 5.51 (dtt,  $J = 15.3, 5.1, 1.1$  Hz, 1H), 4.11 (dd,  $J = 5.2, 1.4$  Hz, 2H), 3.50 (t,  $J = 7.3$  Hz, 2H), 2.00 (q,  $J = 7.8, 7.4$  Hz, 2H), 1.56 (p,  $J = 7.2$  Hz, 2H), 1.39–1.19 (m, 14H), 0.90 (s, 9H), 0.06 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 171.02, 134.16, 131.73, 129.17, 64.26, 38.07, 32.33, 29.62, 29.60, 29.58, 29.32, 29.26, 28.69, 26.88, 26.14, 18.59, –4.93.

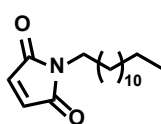


**Synthesis of (*E*)-1-(13-hydroxytridec-11-en-1-yl)-1*H*-pyrrole-2,5-dione (2.9).** To a solution of (*E*)-1-(13-((*tert*-butyl)dimethylsilyloxy)tridec-11-en-1-yl)-1*H*-pyrrole-2,5-dione (693 mg, 1.70 mmol) in dry THF (20 mL) was added a 1 M aqueous solution of HCl (10 mL) at room temperature under N<sub>2</sub>. After 2 h, THF was removed under reduced pressure and the remaining aqueous solution was extracted into CH<sub>2</sub>Cl<sub>2</sub> three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was removed under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 20–26.5–30% EtOAc/hexanes to afford a white solid (469 mg, 94%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.67 (s, 2H), 5.74–5.56 (m, 2H), 4.07 (d,  $J = 5.2$  Hz, 2H), 3.49 (t,  $J = 7.3$  Hz, 2H), 2.02 (q,  $J = 6.7$  Hz, 2H), 1.55 (p,  $J = 7.2$  Hz, 2H), 1.43 (s, 1H), 1.39–1.30 (m, 2H), 1.30–1.19 (m, 12H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 171.03, 134.14, 133.63, 128.94, 63.95, 38.04, 32.31, 29.53, 29.22, 29.20, 28.64, 26.83.



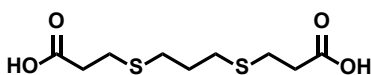
**Synthesis of (*E*)-1-(13-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)tridec-2-en-1-yl) dihydrogen phosphate (2.1).** To a solution of (*E*)-1-(13-hydroxytridec-11-en-1-yl)-1*H*-pyrrole-2,5-dione (50 mg, 0.17 mmol, 1 equiv.) in dry MeCN (0.5 mL) was added trichloroacetonitrile (0.04 mL, 0.41 mmol, 2.4 equiv.) followed by dropwise addition of a solution of Bu<sub>4</sub>NH<sub>2</sub>PO<sub>4</sub> in dry MeCN (2.5 mL) at room temperature under N<sub>2</sub>. After 2 h, the solvent was removed under reduced pressure and the crude mixture was purified via flash silica gel chromatography, eluting with 7:2:1 <sup>1</sup>PrOH/concen-

trated aqueous  $\text{NH}_4\text{OH}/\text{H}_2\text{O}$ . Pure fractions containing phosphate (as determined by silica gel thin layer chromatography, visualized by treatment with  $(\text{NH}_4)_2\text{MoO}_4/\text{HNO}_3$  solution followed by acidic aqueous  $\text{SnCl}_2$ ) were combined and concentrated under reduced pressure. The product was converted to the ammonium hydrogen phosphate salt via ion exchange with Dowex 50WX8 resin, eluting with 25 mM  $\text{NH}_4\text{HCO}_3$  buffer. Fractions containing phosphate were lyophilized to afford the ammonium salt of the desired product as a white solid. This material was further purified via reverse phase HPLC (C18, 5–95% MeCN/ $\text{H}_2\text{O}$  with 0.1% TFA) and isolated as the TFA salt (1.3 mg, 2%).  $^1\text{H}$  NMR (400 MHz, Deuterium Oxide)  $\delta$  6.27 (s, 2H), 5.82 (dt,  $J = 14.8, 6.4$  Hz, 1H), 5.60 (dt,  $J = 14.8, 6.1$  Hz, 1H), 4.28 (t,  $J = 6.8$  Hz, 2H), 3.20 (t,  $J = 6.4$  Hz, 3H), 2.14–1.94 (m, 2H), 1.62–1.06 (m, 16H).  $^{31}\text{P}$  NMR (162 MHz,  $\text{D}_2\text{O}$ )  $\delta$  0.49. HRMS (ESI) calculated for  $\text{C}_{17}\text{H}_{27}\text{NO}_6\text{P}^-$  ( $[\text{M}-\text{H}]^-$ ) 372.1581, found 372.1582.



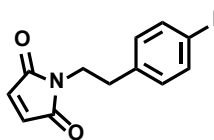
**Synthesis of 1-tridecyl-1H-pyrrole-2,5-dione (2.10).** To a solution of  $\text{PPh}_3$  (175 mg, 0.665 mmol, 1 equiv.) in dry THF (5 mL) was added a solution of DEAD in toluene (40 wt%, 0.3 mL, 0.665 mmol, 1 equiv.) dropwise at  $-78^\circ\text{C}$  under  $\text{N}_2$ . After 5 min, 1-tridecanol (200 mg, 1.00 mmol, 1.5 equiv.) was added as a solid in one portion.

After 5 min, maleimide (65 mg, 0.665 mmol, 1 equiv.) was added as a solid in one portion and the reaction was warmed slowly to room temperature and stirred for 12 h. Volatiles were removed under reduced pressure and the crude mixture was purified via flash silica gel chromatography, eluting with 20% EtOAc/hexanes to afford a lustrous white solid (136 mg, 73%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.67 (s, 2H), 3.49 (t,  $J = 7.3$  Hz, 2H), 1.56 (p,  $J = 7.2$  Hz, 2H), 1.34–1.16 (m, 20H), 0.87 (t,  $J = 6.7$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.02, 134.14, 38.07, 32.04, 29.78, 29.77, 29.75, 29.68, 29.62, 29.48, 29.26, 28.68, 26.88, 22.82, 14.26.



**Synthesis of 3,3'-(propane-1,3-diylbis(sulfaneydiyl))dipropionic acid (2.11).** To a mixture of 3-mercaptopropionic acid (3.6 mL, 41.0 mmol, 2.16 equiv.) and NaOH (3.12 g, 78.0 mmol, 4.11 equiv.)

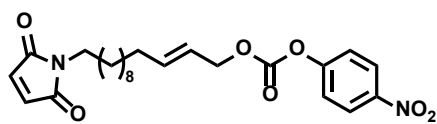
in MeOH (70 mL) was added a solution of 1,3-dibromopropane (2.0 mL, 19.0 mmol, 1 equiv.) in MeOH (10 mL) dropwise at room temperature under  $\text{N}_2$ . After 2 h, the reaction was heated to  $65^\circ\text{C}$  for 12 h. After cooling to room temperature, volatiles were removed under reduced pressure. The crude material was dissolved in  $\text{H}_2\text{O}$  (80 mL) and precipitated by the addition of concentrated aqueous HCl. The product was collected by filtration, affording a white solid after drying under reduced pressure (3.3 g, 69%).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$  12.14 (s, 2H), 2.62 (t,  $J = 7.1$  Hz, 4H), 2.53 (t,  $J = 7.2$  Hz, 4H), 2.45 (t,  $J = 7.0$  Hz, 4H), 1.70 (p,  $J = 7.2$  Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz, DMSO)  $\delta$  173.01, 34.51, 29.82, 29.04, 26.30.



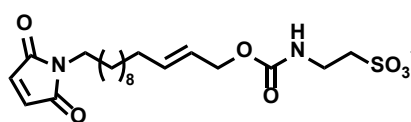
**Synthesis of 1-(4-iodophenethyl)-1H-pyrrole-2,5-dione (2.12).** To a solution of  $\text{PPh}_3$  (175 mg, 0.665 mmol, 1 equiv.) in dry THF (5 mL) was added a solution of DEAD in toluene (40 wt%, 0.3 mL, 0.665 mmol, 1 equiv.) dropwise at  $-78^\circ\text{C}$  under  $\text{N}_2$ . After 5 min, 4-iodobenzyl alcohol (234 mg, 1.00 mmol, 1.5 equiv.) was

added as a solid in one portion. After 5 min, maleimide (65 mg, 0.665 mmol, 1 equiv.) was added as a solid in one portion and the reaction was warmed slowly to room temperature and stirred for 12 h. Volatiles were removed under reduced pressure and the crude mixture was purified via flash silica gel chromatography, eluting with 10–20% EtOAc/hexanes to afford an off-white solid (97 mg, 31%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.69–7.60 (m, 2H), 7.14–7.05 (m, 2H), 6.71 (s, 2H),

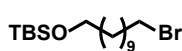
4.61 (s, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  162.79, 138.57, 134.95, 131.67, 113.20, 39.27, 34.55.



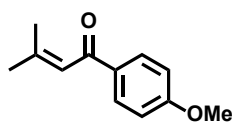
**Synthesis of (*E*)-13-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)tridec-2-en-1-yl (4-nitrophenyl) carbonate.** To a solution of *p*-nitrophenyl chloroformate (38 mg, 0.19 mmol, 1.1 equiv.) and pyridine (0.03 mL, 0.34 mmol, 2 equiv.) in  $\text{CH}_2\text{Cl}_2$  (2 mL) was added (*E*)-1-(13-hydroxytridec-11-en-1-yl)-1*H*-pyrrole-2,5-dione (50 mg, 0.17 mmol, 1 equiv.) at room temperature under  $\text{N}_2$ . After 1 h, the reaction mixture was concentrated under reduced pressure and purified via flash silica gel chromatography, eluting with 20% EtOAc/hexanes to afford a white solid (58 mg, 74%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.35–8.23 (m, 2H), 7.39 (dd,  $J = 10.0, 2.9$  Hz, 2H), 6.68 (s, 2H), 5.91 (dt,  $J = 14.4, 6.6$  Hz, 1H), 5.65 (dt,  $J = 14.7, 6.7$  Hz, 1H), 4.71 (d,  $J = 6.8$  Hz, 2H), 3.50 (t,  $J = 7.3$  Hz, 2H), 2.09 (q,  $J = 7.2$  Hz, 2H), 1.55 (s, 2H), 1.39 (s, 2H), 1.26 (s, 12H).



**Synthesis of (*E*)-2-((((13-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)tridec-2-en-1-yl)oxy)carbonyl)amino)ethane-1-sulfonate (2.14).** To a solution of (*E*)-13-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)tridec-2-en-1-yl (4-nitrophenyl) carbonate (58 mg, 0.13 mmol, 1 equiv.) in dry DMF (2 mL) was added freshly prepared sodium taurinate (38 mg, 0.26 mmol, 2 equiv.) at room temperature under  $\text{N}_2$ . The reaction was monitored by negative mode LC-MS and determined to be complete after 10 min. DMF was removed under reduced pressure and the crude material was purified by reverse phase HPLC to afford the desired taurine carbamate as a white solid after lyophilization (2 mg, 3.5% assuming the  $\text{Na}^+$  salt). LRMS (ESI) calculated for  $\text{C}_{20}\text{H}_{31}\text{N}_2\text{O}_7\text{S}^-$  ( $\text{M}^-$ ) 443.2, found 443.2.

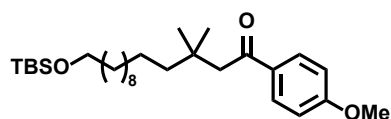


**Synthesis of ((11-bromoundecyl)oxy)(*tert*-butyl)dimethylsilane (2.16).** To a solution of imidazole (8.13 g, 119.4 mmol, 1.5 equiv.) and 11-bromo-1-undecanol (20.0 g, 79.6 mmol, 1 equiv.) was added *tert*-butyldimethylsilyl chloride portion-wise at room temperature under  $\text{N}_2$ . After 12 h, the reaction mixture was filtered. The filtrate was washed three times with water, once with brine, dried over  $\text{Na}_2\text{SO}_4$ , and the solvent was removed under reduced pressure. The crude material was then purified via flash silica gel chromatography, eluting with 2.5% EtOAc/hexanes to afford a clear colorless liquid (28.05 g, 96%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.59 (t,  $J = 6.6$  Hz, 2H), 3.40 (t,  $J = 6.9$  Hz, 2H), 1.85 (p,  $J = 6.9$  Hz, 2H), 1.50 (p,  $J = 7.5$  Hz, 2H), 1.40 (p,  $J = 7.0$  Hz, 2H), 1.34–1.23 (m, 12H), 0.89 (s, 9H), 0.04 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  63.47, 34.21, 33.02, 32.99, 29.73, 29.62, 29.57, 28.92, 28.32, 26.14, 25.94, 18.53, -5.10.

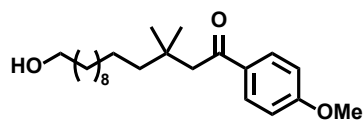


**Synthesis of 1-(4-methoxyphenyl)-3-methylbut-2-en-1-one (2.17).** To a solution of 3,3-dimethylacrylic acid (1.20 g, 12.0 mmol, 1 equiv.) in dry  $\text{CH}_2\text{Cl}_2$  (25 mL) was added oxalyl chloride (1.12 mL, 13.2 mmol, 1.1 equiv.) slowly at 0 °C under  $\text{N}_2$ . The reaction was warmed to room temperature and stirred for 1 h. The reaction mixture was concentrated under reduced pressure, taken up in  $\text{CHCl}_3$ , and concentrated again to remove solvent and excess oxalyl chloride. The crude acid chloride was dissolved in dry  $\text{CH}_2\text{Cl}_2$  (25 mL) and anisole (1.63 mL, 15.0 mmol, 1.25 equiv.) was added under  $\text{N}_2$ . The solution was cooled to 0 °C and solid  $\text{AlCl}_3$  (2.0 g, 15.0 mmol, 1.25 equiv.) was added portion-wise over 15 min. The reaction was warmed to room temperature and stirred for 4 h. The reaction was quenched by rapid addition of ice cold  $\text{H}_2\text{O}$  (50 mL) with vigorous stirring. The layers were sepa-

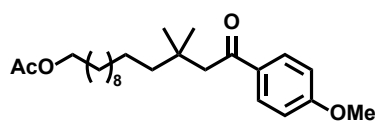
rated and the organic layer was washed with 1 M HCl, brine, dried over  $\text{Na}_2\text{SO}_4$ , and the solvent was removed under reduced pressure. The crude material was then purified via flash silica gel chromatography, eluting with 5–10% EtOAc/hexanes to afford a pale yellow oil (0.70 g, 31%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.98–7.88 (m, 2H), 6.95–6.88 (m, 2H), 6.74–6.68 (m, 1H), 3.84 (s, 3H), 2.17 (s, 3H), 1.99 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  190.40, 163.01, 155.36, 132.20, 130.53, 121.26, 113.66, 55.51, 27.98, 21.15.



**Synthesis of 14-((*tert*-butyldimethylsilyl)oxy)-1-(4-methoxyphenyl)-3,3-dimethyltetradecan-1-one (2.18).** In a flame-dried flask, ((11-bromoundecyl)oxy)(*tert*-butyl)dimethylsilane (12.68 g, 34.7 mmol, 2.2 equiv.) was added dropwise to freshly etched Mg turnings (1.69 g, 69.4 mmol, 4.4 equiv.) in dry THF (50 mL). The mixture was heated to reflux for 2.5 h under  $\text{N}_2$  and then cooled to room temperature. The resultant solution of (11-((*tert*-butyldimethylsilyl)oxy)undecyl)magnesium bromide was cannulated into a fresh flame-dried flask charged with CuI (300 mg, 1.58 mmol, 0.1 equiv.) at 0 °C. After stirring for 5 min, neat 1-(4-methoxyphenyl)-3-methylbut-2-en-1-one (3.00 g, 15.8 mmol, 1 equiv.) was added dropwise and the reaction was warmed slowly to room temperature for 5 h. The reaction mixture was cooled to 0 °C and quenched with saturated aqueous  $\text{NH}_4\text{Cl}$  (25 mL) and diluted with  $\text{Et}_2\text{O}$  (50 mL). After stirring vigorously for 1 h, the layers were separated and the aqueous layer was extracted into  $\text{Et}_2\text{O}$  three times. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 5% EtOAc/hexanes to afford a colorless oil (5.6 g, 74%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.96–7.85 (m, 2H), 6.93–6.84 (m, 2H), 3.82 (s, 3H), 3.57 (t,  $J = 6.6$  Hz, 2H), 2.76 (s, 2H), 1.48 (p,  $J = 6.7$  Hz, 2H), 1.40–1.18 (m, 18H), 0.98 (s, 6H), 0.87 (s, 9H), 0.02 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  199.05, 163.19, 131.86, 130.50, 113.57, 63.36, 55.40, 47.75, 42.90, 34.04, 32.96, 30.52, 29.78, 29.74, 29.70, 29.53, 27.80, 26.06, 25.88, 24.30, 18.42, -5.18.

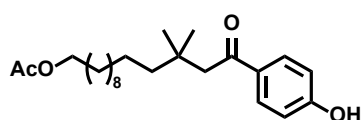


**Synthesis of 14-hydroxy-1-(4-methoxyphenyl)-3,3-dimethyltetradecan-1-one (2.19).** To a solution of 14-((*tert*-butyldimethylsilyl)oxy)-1-(4-methoxyphenyl)-3,3-dimethyltetradecan-1-one (5.6 g, 11.7 mmol) in THF (20 mL) was added 1 M aqueous HCl (5 mL) at room temperature and stirred for 2 h. The THF was removed under reduced pressure and the remaining aqueous solution was extracted three times with EtOAc. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 20–33% EtOAc/hexanes to afford a pale yellow oil (4.24 g, 99%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.96–7.88 (m, 2H), 6.95–6.88 (m, 2H), 3.86 (s, 3H), 3.63 (t,  $J = 6.7$  Hz, 2H), 2.78 (s, 2H), 1.56 (p,  $J = 6.6$  Hz, 2H), 1.40–1.18 (m, 18H), 0.99 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  199.38, 163.27, 131.97, 130.62, 113.67, 63.23, 55.58, 47.90, 42.97, 34.17, 32.95, 30.55, 29.81, 29.77, 29.72, 29.71, 29.56, 27.87, 25.87, 24.35.



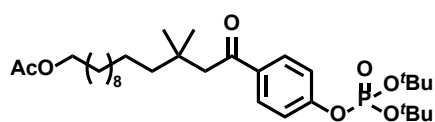
**Synthesis of 14-(4-methoxyphenyl)-12,12-dimethyl-14-oxotetradecyl acetate (2.20).** To a mixture of 14-hydroxy-1-(4-methoxyphenyl)-3,3-dimethyltetradecan-1-one (1.75 g, 4.83 mmol, 1 equiv.), DMAP (59 mg, 0.48 mmol, 0.1 equiv.), and  $\text{Et}_3\text{N}$  (2.0 mL, 14.5 mmol, 3 equiv.) in dry  $\text{CH}_2\text{Cl}_2$  (20 mL) was added  $\text{Ac}_2\text{O}$  (0.55 mL, 5.79 mmol, 1.2 equiv.) at 0 °C under  $\text{N}_2$ ,

The reaction was warmed to room temperature. After 30 min, saturated aqueous NaHCO<sub>3</sub> (20 mL) was added, thoroughly mixed, and the layers were separated. The aqueous solution was extracted three times into CH<sub>2</sub>Cl<sub>2</sub> and the combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10–20% EtOAc/hexanes to afford a pale yellow oil (1.95 g, 99%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.95–7.88 (m, 2H), 6.95–6.87 (m, 2H), 4.04 (t, *J* = 6.8 Hz, 2H), 3.86 (s, 3H), 2.78 (s, 2H), 2.04 (s, 3H), 1.61 (p, *J* = 7.0 Hz, 2H), 1.39–1.18 (m, 18H), 0.99 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 199.31, 171.38, 163.26, 131.96, 130.60, 113.65, 64.80, 55.56, 47.88, 42.95, 34.14, 30.55, 29.81, 29.76, 29.69, 29.64, 29.39, 28.72, 27.86, 26.04, 24.35, 21.17.



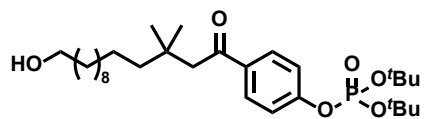
**Synthesis of 14-(4-hydroxyphenyl)-12,12-dimethyl-14-oxotetradecyl acetate (2.21).** To a solution of 14-(4-methoxyphenyl)-12,12-dimethyl-14-oxotetradecyl acetate (4.23 g, 10.46 mmol, 1 equiv.) in dry CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was added a suspension of AlCl<sub>3</sub> (4.18 g,

31.37 mmol, 3 equiv.) in dry CH<sub>2</sub>Cl<sub>2</sub> (50 mL) at room temperature under N<sub>2</sub>. Hexanethiol (2.97 mL, 20.91 mmol, 2 equiv.) was added dropwise and the reaction mixture was stirred for 30 min at room temperature and heated to reflux for 12 h. After cooling, the reaction was quenched by pouring over ice cold 1 M aqueous HCl (50 mL). The layers were separated and the aqueous solution was extracted into CH<sub>2</sub>Cl<sub>2</sub> three times. The combined organic layers were washed with 1 M aqueous HCl, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 20–50% EtOAc/hexanes to afford a pale yellow oil (3.9 g, 96%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.92–7.82 (m, 2H), 7.41 (s, 1H), 6.94–6.86 (m, 2H), 4.06 (t, *J* = 6.8 Hz, 2H), 2.79 (s, 2H), 2.06 (s, 3H), 1.61 (p, *J* = 6.9 Hz, 2H), 1.39–1.16 (m, 18H), 0.99 (s, 6H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 200.38, 172.12, 162.43, 160.85, 131.47, 131.08, 115.41, 65.11, 47.98, 42.97, 34.31, 30.47, 29.72, 29.69, 29.64, 29.59, 29.33, 28.67, 27.92, 25.98, 24.31, 21.25.



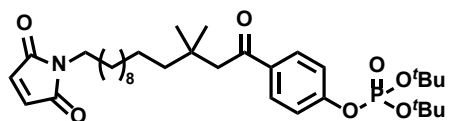
**Synthesis of 14-(4-((di-tert-butoxyphosphoryl)oxy)phenyl)-12,12-dimethyl-14-oxotetradecyl acetate (2.23).** Trifluoroacetic acid (3.52 mL, 45.9 mmol, 4.6 equiv.) was added to a solution of imidazole (3.13 g, 45.9 mmol, 4.6 equiv.) in dry THF

(30 mL) and stirred for 10 min under N<sub>2</sub>. To the resultant imidazolium trifluoroacetate solution was added di-tert-butyl *N,N*-diisopropylphosphoramidite (4.73 mL, 14.9 mmol, 1.5 equiv.) dropwise at room temperature. After 10 min, a solution of 14-(4-hydroxyphenyl)-12,12-dimethyl-14-oxotetradecyl acetate (3.90 g, 9.99 mmol, 1 equiv.) in dry THF (50 mL) was added dropwise at room temperature and the reaction was stirred for 6 h. The solution of crude phosphite was cooled to –50 °C and *tert*-butyl hydroperoxide (80% in di-*tert*-butyl peroxide/water 3:2, 3.25 mL, 26.0 mmol, 2.6 equiv.) was added dropwise. The reaction was warmed to room temperature and stirred for 12 h. Saturated aqueous NaHCO<sub>3</sub> (50 mL) was added, thoroughly mixed, and then diluted with H<sub>2</sub>O (25 mL). The mixture was extracted three times into EtOAc and the combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10–20% EtOAc/hexanes to afford a colorless oil (4.69 g, 81%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.98–7.86 (m, 2H), 7.30–7.20 (m, 2H), 4.03 (t, *J* = 6.7 Hz, 2H), 2.80 (s, 2H), 2.03 (s, 3H), 1.60 (p, *J* = 7.0 Hz, 2H), 1.50 (s, 18H), 1.39–1.18 (m, 18H), 0.99 (s, 6H). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ –16.05.



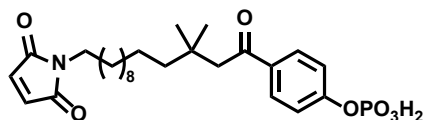
**Synthesis of di-tert-butyl (4-(14-hydroxy-3,3-dimethyltetradecanoyl)phenyl) phosphate (2.24).** To a solution of 14-(4-((di-tert-butoxyphosphoryl)oxy)phenyl)-12,12-dimethyl-14-oxotetradecyl acetate (4.69 g, 8.05 mmol, 1 equiv.) in 2.5:1

THF/H<sub>2</sub>O (100 mL) was added a solution of LiOH (579 mg, 24.2 mmol, 3 equiv.) in H<sub>2</sub>O (20 mL) dropwise at 0 °C. The reaction was warmed to room temperature and additional LiOH (579 mg, 24.2 mmol, 3 equiv.) in H<sub>2</sub>O (20 mL) was added in one portion. After 12 h, the reaction mixture was extracted into EtOAc three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 33–50% EtOAc/hexanes to afford a colorless oil (2.63 g, 60%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.98–7.86 (m, 2H), 7.30–7.20 (m, 2H), 3.58 (t, *J* = 6.4 Hz, 2H), 2.80 (s, 2H), 1.80 (p, *J* = 6.9 Hz, 2H), 1.50 (s, 18H), 1.39–1.18 (m, 18H), 0.99 (s, 6H). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ –16.35.



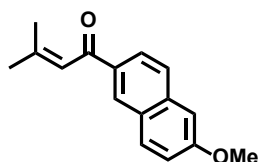
**Synthesis of di-tert-butyl (4-(14-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-3,3-dimethyltetradecanoyl)phenyl) phosphate (2.25).** To a solution of PPh<sub>3</sub> (23 mg, 0.086 mmol, 0.667 equiv.) in dry THF (3 mL) was added a solution of

diethyl azodicarboxylate in toluene (40 wt%, 0.24 mL, 0.516 mmol, 5.73 equiv.) dropwise at –78 °C under N<sub>2</sub>. After 5 min, di-tert-butyl (4-(14-hydroxy-3,3-dimethyltetradecanoyl)phenyl) phosphate (70 mg, 0.129 mmol, 1.0 equiv.) was added as a solid in one portion. After 5 min, maleimide (8.6 mg, 0.086 mmol, 0.667 equiv.) was added as a solid in one portion and the reaction was warmed slowly to room temperature and stirred for 12 h. Volatiles were removed under reduced pressure and the crude mixture was purified via flash silica gel chromatography, eluting with 33–50% EtOAc/hexanes to afford a white solid (36 mg, 45%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.97–7.86 (m, 2H), 7.33–7.20 (m, 2H), 6.67 (s, 2H), 3.49 (t, *J* = 7.3 Hz, 2H), 2.79 (s, 2H), 1.61–1.53 (m, 2H), 1.50 (s, 18H), 1.39–1.18 (m, 18H), 0.98 (s, 6H). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ –16.07.



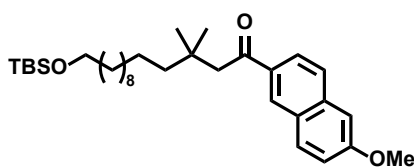
**Synthesis of 4-(14-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-3,3-dimethyltetradecanoyl)phenyl dihydrogen phosphate (2.15).** To a solution of di-tert-butyl (4-(14-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-3,3-dimethyltetradecanoyl)phenyl)

phosphate (5 mg, 0.008 mmol) in EtOAc (1 mL) was added concentrated aqueous HCl (0.1 mL) at room temperature and stirred for 1.5 h. Volatiles were removed under reduced pressure and the crude material was purified via reverse phase HPLC (C18, 5–95% MeCN/H<sub>2</sub>O with 0.1% TFA) to afford a white solid (3.5 mg, 85%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 7.92 (d, *J* = 8.5 Hz, 2H), 7.23 (d, *J* = 8.5 Hz, 2H), 7.00 (s, 2H), 3.37 (t, *J* = 7.0 Hz, 2H), 2.82 (s, 2H), 1.46 (p, *J* = 6.8 Hz, 2H), 1.22 (d, *J* = 7.6 Hz, 18H), 0.93 (s, 6H). <sup>31</sup>P NMR (162 MHz, DMSO) δ –6.09. HRMS (ESI) calculated for C<sub>26</sub>H<sub>37</sub>NO<sub>7</sub>P<sup>–</sup> ([M–H]<sup>–</sup>) 506.2313, found 506.2311.



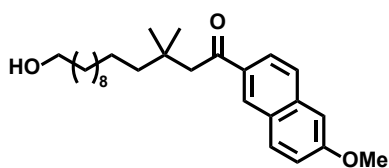
**Synthesis of 1-(6-methoxynaphthalen-2-yl)-3-methylbut-2-en-1-one (2.27).** In a flame-dried flask, 2-bromo-6-methoxynaphthalene (10.0 g, 42.2 mmol, 1.5 equiv.) was added dropwise to freshly etched Mg turnings (1.37 g, 56.2 mmol, 2 equiv.) in dry THF (50 mL) under N<sub>2</sub> and the reaction was heated to reflux for 2.5 h. The resultant solution of (6-methoxy-

naphthalen-2-yl)magnesium bromide was cooled to room temperature and added slowly to a solution of 3-methylcrotonaldehyde (2.7 mL, 28.1 mmol, 1 equiv.) in dry THF (50 mL) at 0 °C. The reaction was warmed to room temperature and stirred for 2 h. After cooling to 0 °C, saturated aqueous NH<sub>4</sub>Cl (40 mL) was added and the mixture was extracted into EtOAc three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude alcohol was then dissolved in CH<sub>2</sub>Cl<sub>2</sub> (100 mL) and MnO<sub>2</sub> (19.6 g, 225 mmol, 8 equiv.) was added. The mixture was stirred at room temperature overnight and then filtered through Celite. The filtered solution was concentrated under reduced pressure and the residue was dissolved in CHCl<sub>3</sub> and purified via flash silica gel chromatography, eluting with 5–10% EtOAc/hexanes to afford an off-white solid (3.2 g, 32%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.35 (d, *J* = 1.7 Hz, 1H), 8.00 (dd, *J* = 8.6, 1.8 Hz, 1H), 7.81 (d, *J* = 8.9 Hz, 1H), 7.73 (d, *J* = 8.6 Hz, 1H), 7.16 (dd, *J* = 8.9, 2.6 Hz, 1H), 7.11 (d, *J* = 2.5 Hz, 1H), 6.90–6.82 (m, 1H), 3.90 (s, 3H), 2.24 (d, *J* = 1.3 Hz, 3H), 2.03 (d, *J* = 1.4 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 191.20, 159.48, 155.93, 136.90, 134.57, 131.04, 129.45, 127.90, 127.05, 125.08, 121.32, 119.54, 105.71, 55.38, 28.01, 21.22.



**Synthesis of 14-((*tert*-butyldimethylsilyloxy)-1-(6-methoxynaphthalen-2-yl)-3,3-dimethyltetradecan-1-one (2.28).** In a flame-dried flask, ((11-bromoundecyl)oxy)(*tert*-butyl)dimethylsilane (10.04 g, 27.5 mmol, 2.2 equiv.) was added dropwise to freshly etched Mg turnings (1.34 g, 55.0 mmol, 4.4 equiv.) in dry

THF (50 mL). The mixture was heated to reflux for 2.5 h under N<sub>2</sub> and then cooled to room temperature. The resultant solution of (11-((*tert*-butyldimethylsilyloxy)undecyl)magnesium bromide was cannulated into a fresh flame-dried flask charged with CuI (238 mg, 1.25 mmol, 0.1 equiv.) at 0 °C. After stirring for 5 min, neat 1-(6-methoxynaphthalen-2-yl)-3-methylbut-2-en-1-one (3.00 g, 12.5 mmol, 1 equiv.) was added dropwise and the reaction was warmed slowly to room temperature for 5 h. The reaction mixture was cooled to 0 °C and quenched with saturated aqueous NH<sub>4</sub>Cl (25 mL) and diluted with Et<sub>2</sub>O (50 mL). After stirring vigorously for 1 h, the layers were separated and the aqueous layer was extracted into Et<sub>2</sub>O three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 5% EtOAc/hexanes to afford a yellow oil (5.53 g, 84%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.37 (d, *J* = 1.7 Hz, 1H), 7.99 (dd, *J* = 8.6, 1.8 Hz, 1H), 7.85 (d, *J* = 8.9 Hz, 1H), 7.76 (d, *J* = 8.6 Hz, 1H), 7.20 (dd, *J* = 8.9, 2.5 Hz, 1H), 7.15 (d, *J* = 2.6 Hz, 1H), 3.94 (s, 3H), 3.59 (t, *J* = 6.6 Hz, 2H), 2.94 (s, 2H), 1.49 (p, *J* = 6.6 Hz, 2H), 1.44–1.36 (m, 2H), 1.25 (d, *J* = 5.8 Hz, 16H), 1.04 (s, 6H), 0.89 (s, 9H), 0.05 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 200.47, 159.74, 137.16, 134.36, 134.33, 131.27, 129.77, 127.95, 127.13, 125.01, 119.75, 105.79, 63.50, 60.53, 55.55, 48.16, 42.99, 34.28, 33.04, 30.59, 29.86, 29.81, 29.76, 29.60, 27.92, 26.14, 25.95, 24.42, 18.53, 14.35, –5.10.

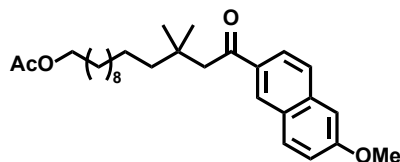


**Synthesis of 14-hydroxy-1-(6-methoxynaphthalen-2-yl)-3,3-dimethyltetradecan-1-one.** To a solution of 14-((*tert*-butyldimethylsilyloxy)-1-(6-methoxynaphthalen-2-yl)-3,3-dimethyltetradecan-1-one (5.53 g, 10.5 mmol) in THF (40 mL) was added 1 M aqueous HCl (10 mL) at room temperature and the mixture was stirred

for 3.5 h. The THF was removed under reduced pressure and the remaining aqueous solution was extracted three times with EtOAc. The combined organic layers were washed with brine, dried over

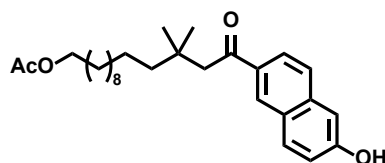


$\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 20–33–50% EtOAc/hexanes to afford a yellow oil (4.33 g, 99%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.36 (d,  $J$  = 1.7 Hz, 1H), 7.98 (dd,  $J$  = 8.6, 1.8 Hz, 1H), 7.84 (d,  $J$  = 9.0 Hz, 1H), 7.75 (d,  $J$  = 8.6 Hz, 1H), 7.19 (dd,  $J$  = 8.9, 2.5 Hz, 1H), 7.14 (d,  $J$  = 2.5 Hz, 1H), 3.94 (s, 3H), 3.63 (t,  $J$  = 6.6 Hz, 2H), 2.94 (s, 2H), 1.55 (p,  $J$  = 6.5 Hz, 2H), 1.46–1.36 (m, 2H), 1.36–1.18 (m, 16H), 1.03 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  200.48, 159.71, 137.13, 134.28, 131.24, 129.76, 127.92, 127.11, 124.97, 119.72, 105.76, 63.18, 60.53, 55.53, 48.12, 42.90, 34.26, 32.93, 30.54, 29.79, 29.75, 29.70, 29.55, 27.92, 25.85, 24.38, 21.18, 14.32.



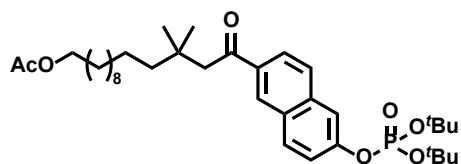
**Synthesis of 14-(6-methoxynaphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (2.29).** To a mixture of 14-(6-methoxynaphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (4.33 g, 10.5 mmol, 1 equiv.), DMAP (128 mg, 1.05 mmol, 0.1 equiv.), and  $\text{Et}_3\text{N}$  (4.39 mL, 31.5 mmol, 3 equiv.) in dry  $\text{CH}_2\text{Cl}_2$

(50 mL) was added  $\text{Ac}_2\text{O}$  (1.19 mL, 12.6 mmol, 1.2 equiv.) at 0 °C under  $\text{N}_2$ . The reaction was warmed to room temperature. After 30 min, saturated aqueous  $\text{NaHCO}_3$  (50 mL) was added, the mixture was thoroughly mixed, and the layers were separated. The aqueous solution was extracted three times into  $\text{CH}_2\text{Cl}_2$  and the combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10–20% EtOAc/hexanes to afford a yellow oil (3.85 g, 88%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.36 (d,  $J$  = 1.7 Hz, 1H), 7.99 (dd,  $J$  = 8.6, 1.8 Hz, 1H), 7.84 (d,  $J$  = 8.9 Hz, 1H), 7.75 (d,  $J$  = 8.6 Hz, 1H), 7.19 (dd,  $J$  = 8.9, 2.5 Hz, 1H), 7.14 (d,  $J$  = 2.6 Hz, 1H), 4.04 (t,  $J$  = 6.7 Hz, 2H), 3.94 (s, 3H), 2.94 (s, 2H), 2.04 (s, 3H), 1.61 (p,  $J$  = 6.9 Hz, 2H), 1.45–1.38 (m, 2H), 1.37–1.20 (m, 16H), 1.04 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  200.38, 171.32, 159.75, 137.15, 134.34, 131.23, 129.72, 127.96, 127.11, 124.98, 119.72, 105.82, 64.78, 55.52, 48.15, 42.91, 34.25, 30.55, 29.79, 29.75, 29.67, 29.62, 29.38, 28.74, 27.92, 26.03, 24.39, 21.13.



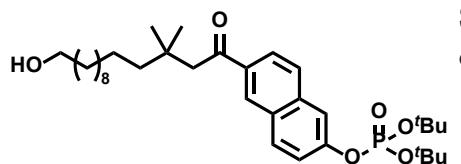
**Synthesis of 14-(6-hydroxynaphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (2.30).** To a solution of 14-(6-methoxynaphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (2.85 g, 6.27 mmol, 1 equiv.) in dry  $\text{CH}_2\text{Cl}_2$  (20 mL) was added a suspension of  $\text{AlCl}_3$  (3.34 g, 25.1 mmol, 3 equiv.) in dry  $\text{CH}_2\text{Cl}_2$

(20 mL) at room temperature under  $\text{N}_2$ . Hexanethiol (1.78 mL, 12.5 mmol, 2 equiv.) was added dropwise and the reaction mixture was stirred for 30 min at room temperature and heated to reflux for 12 h. After cooling, the reaction was quenched by pouring over ice cold 1 M aqueous HCl (20 mL). The layers were separated and the aqueous solution was extracted into  $\text{CH}_2\text{Cl}_2$  three times. The combined organic layers were washed with 1 M aqueous HCl, brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10–20–33% EtOAc/hexanes to afford a yellow oil (1.46 g, 53%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.37 (s, 1H), 7.97 (dd,  $J$  = 8.6, 1.9 Hz, 1H), 7.87 (d,  $J$  = 9.4 Hz, 1H), 7.69 (d,  $J$  = 8.7 Hz, 1H), 7.24–7.08 (m, 2H), 6.28 (s, 1H), 4.07 (dd,  $J$  = 7.7, 5.9 Hz, 2H), 2.95 (s, 2H), 2.07 (d,  $J$  = 1.8 Hz, 3H), 1.60 (p,  $J$  = 7.0 Hz, 2H), 1.46–1.12 (m, 18H), 1.05 (s, 6H).



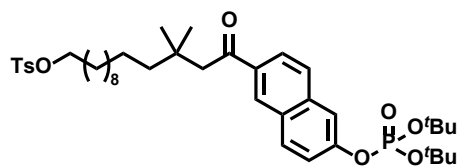
**Synthesis of 14-(6-((di-*tert*-butoxyphosphoryl)oxy)naphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (2.31).**

Trifluoroacetic acid (1.17 mL, 15.2 mmol, 4.6 equiv.) was added to a solution of imidazole (1.04 g, 15.2 mmol, 4.6 equiv.) in dry THF (10 mL) and stirred for 10 min under  $N_2$ . To the resultant imidazolium trifluoroacetate solution was added di-*tert*-butyl *N,N*-diisopropylphosphoramidite (1.57 mL, 4.97 mmol, 1.5 equiv.) dropwise at room temperature. After 10 min, a solution of 14-(6-hydroxynaphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (1.46 g, 3.31 mmol, 1 equiv.) in dry THF (10 mL) was added dropwise at room temperature and the reaction was stirred for 2 h. The solution of crude phosphite was cooled to  $-50\text{ }^\circ\text{C}$  and *tert*-butyl hydroperoxide (80% in di-*tert*-butyl peroxide/water 3:2, 0.95 mL, 8.62 mmol, 2.6 equiv.) was added dropwise. The reaction was warmed to room temperature and stirred for 12 h. Saturated aqueous  $\text{NaHCO}_3$  (10 mL) was added, thoroughly mixed, and then diluted with  $\text{H}_2\text{O}$  (10 mL). The mixture was extracted three times into EtOAc and the combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10–20% EtOAc/hexanes to afford a yellow oil (1.68 g, 80%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.38 (s, 1H), 7.98 (dd,  $J = 8.7, 1.7$  Hz, 1H), 7.90 (d,  $J = 8.9$  Hz, 1H), 7.78 (d,  $J = 8.7$  Hz, 1H), 7.68 (d,  $J = 2.3$  Hz, 1H), 7.41 (dd,  $J = 8.9, 2.4$  Hz, 1H), 4.02 (t,  $J = 6.7$  Hz, 2H), 2.93 (s, 2H), 2.01 (s, 3H), 1.58 (p,  $J = 6.7$  Hz, 2H), 1.51 (s, 18H), 1.48–1.43 (m, 2H), 1.42–1.34 (m, 2H), 1.34–1.17 (m, 14H), 1.02 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  200.28, 171.22, 151.12 (d,  $J = 7.9$  Hz), 136.27, 135.57, 131.29, 129.64, 129.47, 127.88, 124.87, 121.30 (d,  $J = 5.9$  Hz), 116.06 (d,  $J = 5.1$  Hz), 84.12 (d,  $J = 7.5$  Hz), 64.69, 48.18, 42.81, 34.19, 30.48, 29.94 (d,  $J = 4.3$  Hz), 29.72, 29.67, 29.60, 29.55, 29.30, 28.67, 27.83, 25.96, 24.32, 21.06.  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  -15.68.



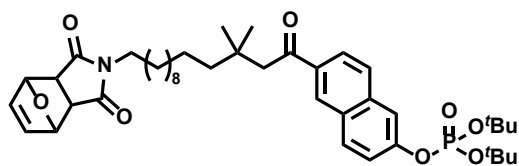
**Synthesis of di-*tert*-butyl (6-(14-hydroxy-3,3-dimethyltetradecanoyl)naphthalen-2-yl) phosphate (2.32).**

To a solution of 14-(6-((di-*tert*-butoxyphosphoryl)oxy)naphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl acetate (1.68 g, 2.66 mmol, 1 equiv.) in 2.5:1 THF/ $\text{H}_2\text{O}$  (47 mL) was added a solution of LiOH (191 mg, 7.96 mmol, 3 equiv.) in  $\text{H}_2\text{O}$  (25 mL) dropwise at  $0\text{ }^\circ\text{C}$ . The reaction was warmed to room temperature and additional LiOH (191 mg, 7.96 mmol, 3 equiv.) in  $\text{H}_2\text{O}$  (25 mL) was added in one portion. After 12 h, the reaction mixture was extracted into EtOAc three times. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 33–50% EtOAc/hexanes to afford a pale yellow oil (1.04 g, 66%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.37 (d,  $J = 1.6$  Hz, 1H), 7.97 (dd,  $J = 8.6, 1.8$  Hz, 1H), 7.90 (d,  $J = 8.8$  Hz, 1H), 7.78 (d,  $J = 8.6$  Hz, 1H), 7.67 (d,  $J = 2.3$  Hz, 1H), 7.41 (dd,  $J = 8.9, 2.5$  Hz, 1H), 3.58 (t,  $J = 6.7$  Hz, 2H), 2.92 (s, 2H), 2.15 (br s, 1H), 1.87–1.77 (m, 2H), 1.51 (s, 18H), 1.43–1.33 (m, 2H), 1.33–1.15 (m, 16H), 1.01 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  200.36, 151.07 (d,  $J = 7.2$  Hz), 136.25, 135.56, 131.30, 129.62, 129.47, 127.87, 124.86, 121.26 (d,  $J = 5.8$  Hz), 116.05 (d,  $J = 5.1$  Hz), 84.21 (d,  $J = 7.5$  Hz), 62.88, 48.15, 42.76, 34.19, 32.85, 30.43, 29.92 (d,  $J = 4.3$  Hz), 29.67, 29.64, 29.60, 29.49, 27.84, 25.83, 25.64, 24.28.  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  -15.80.



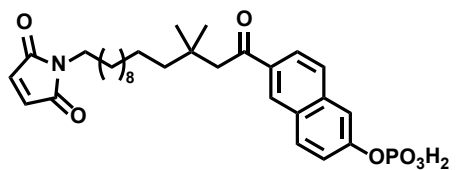
**Synthesis of 14-(6-((di-tert-butoxyphosphoryl)oxy)naphthalen-2-yl)-12,12-dimethyl-14-oxotetradecyl 4-methylbenzenesulfonate (2.33).** To a solution of di-tert-butyl (6-(14-hydroxy-3,3-dimethyltetradecanoyl)naphthalen-2-yl) phosphate (100 mg, 0.169 mmol, 1 equiv.), DMAP (2 mg,

0.017 mmol, 0.1 equiv.), and Et<sub>3</sub>N (0.02 mL, 0.169 mmol, 1 equiv.) in dry CH<sub>2</sub>Cl<sub>2</sub> (5 mL) was added *p*-toluenesulfonyl chloride (64 mg, 0.338 mmol, 2 equiv.) in one portion at 0 °C under N<sub>2</sub>. The reaction was warmed to room temperature and stirred overnight. Volatiles were removed under reduced pressure and the crude material was purified via flash silica gel chromatography, eluting with 20–33% EtOAc/hexanes to afford a colorless oil (80 mg, 64%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.38 (s, 1H), 7.98 (dd, *J* = 8.7, 1.7 Hz, 1H), 7.93 (d, *J* = 7.1 Hz, 2H), 7.90 (d, *J* = 8.9 Hz, 1H), 7.78 (d, *J* = 8.7 Hz, 1H), 7.68 (d, *J* = 2.3 Hz, 1H), 7.44 (d, *J* = 7.1 Hz, 2H), 7.41 (dd, *J* = 8.9, 2.4 Hz, 1H), 3.98 (t, *J* = 6.6 Hz, 2H), 3.64 (s, 3H), 2.93 (s, 2H), 2.01 (s, 3H), 1.57 (p, *J* = 6.6 Hz, 2H), 1.51 (s, 18H), 1.48–1.43 (m, 2H), 1.42–1.34 (m, 2H), 1.34–1.17 (m, 14H), 1.02 (s, 6H).



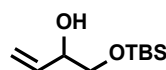
**Synthesis of di-tert-butyl (6-(14-(1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-epoxyisoindol-2-yl)-3,3-dimethyltetradecanoyl)naphthalen-2-yl) phosphate (2.34).** To a solution of 14-(6-((di-tert-butoxyphosphoryl)oxy)naphthalen-2-yl)-12,12-dimethyl-

14-oxotetradecyl 4-methylbenzenesulfonate (80 mg, 0.107 mmol, 1 equiv.) and 3a,4,7,7a-tetrahydro-1*H*-4,7-epoxyisoindole-1,3(2*H*)-dione (mixture of *endo* and *exo* isomers, 40 mg, 0.236 mmol, 2.2 equiv.) in dry MeCN (4 mL) was added K<sub>2</sub>CO<sub>3</sub> in one portion at room temperature. The reaction was heated to 50 °C for 2 d. After diluting with EtOAc (10 mL), the reaction mixture was filtered and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with a gradient of 20–70% EtOAc/hexanes to afford a colorless oil (3:1 *endo/exo*, 10 mg, 13%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.40 (s, 1H), 8.01 (d, *J* = 8.3 Hz, 1H), 7.93 (d, *J* = 9.1 Hz, 1H), 7.82 (d, *J* = 8.4 Hz, 1H), 7.70 (s, 1H), 7.44 (d, *J* = 9.0 Hz, 1H), 6.52–6.48 (*exo*, m, 1.5H), 6.39 (*endo*, s, 0.5H), 5.34–5.30 (*endo*, m, 0.5H), 5.28–5.24 (*exo*, m, 1.5H), 3.46 (t, *J* = 7.3 Hz, 2H), 3.32–3.26 (*endo*, m, 0.5H), 2.95 (s, 2H), 2.89 (*exo*, s, 1.5H), 1.54 (s, 18H), 1.51–1.47 (m, 2H), 1.45–1.36 (m, 2H), 1.36–1.17 (m, 16H), 1.04 (s, 6H). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ –15.63.

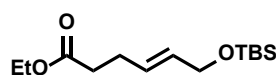


**Synthesis of 6-(14-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)-3,3-dimethyltetradecanoyl)naphthalen-2-yl dihydrogen phosphate (2.26).** A solution of di-tert-butyl (6-(14-(1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2*H*-4,7-epoxyisoindol-2-yl)-3,3-dimethyltetradecanoyl)naphthalen-2-yl) phosphate (10 mg,

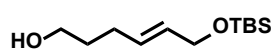
0.014 mmol) in dry toluene (15 mL) was heated to reflux for 12 h under N<sub>2</sub>. Volatiles were removed under reduced pressure and the residue was purified via reverse phase HPLC (C18, 5–95% MeCN/H<sub>2</sub>O with 0.1% TFA) to give an off-white solid (1.2 mg, 14%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.67 (s, 1H), 8.15 (d, *J* = 9.1 Hz, 1H), 8.03–7.89 (m, 2H), 7.73 (s, 1H), 7.45 (d, *J* = 9.2 Hz, 1H), 7.00 (s, 2H), 3.92 (t, *J* = 7.2 Hz, 2H), 3.00 (s, 2H), 1.54–1.40 (m, 2H), 1.40–1.10 (m, 18H), 0.98 (s, 6H). <sup>31</sup>P NMR (162 MHz, DMSO) δ –6.33. HRMS (ESI) calculated for C<sub>30</sub>H<sub>39</sub>NO<sub>7</sub>P<sup>−</sup> ([*M*–*H*]<sup>−</sup>) 556.2470, found 556.2473.



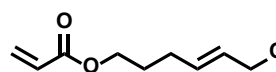
**Synthesis of 1-((*tert*-butyldimethylsilyl)oxy)but-3-en-2-ol (2.36).** To a solution of 3-buten-1,2-diol (8.01 mL, 95.2 mmol, 1 equiv.) and imidazole (25.94 g, 381 mmol, 4 equiv.) in dry THF (100 mL) was added a solution of TBSCl in THF (1.0 M, 100 mL, 100 mmol, 1.05 equiv.) by cannula at room temperature under N<sub>2</sub>. After 12 h, H<sub>2</sub>O (100 mL) was added and the mixture was extracted into Et<sub>2</sub>O (100 mL) three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 5% EtOAc/hexanes to afford a colorless oil (18.3 g, 95%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.81 (ddd, *J* = 17.4, 10.6, 5.8 Hz, 1H), 5.34 (dt, *J* = 17.2, 1.6 Hz, 1H), 5.19 (dt, *J* = 10.5, 1.5 Hz, 1H), 4.20–4.13 (m, 1H), 3.66 (dd, *J* = 9.9, 3.7 Hz, 1H), 3.44 (dd, *J* = 10.0, 7.7 Hz, 1H), 2.58 (d, *J* = 3.5 Hz, 1H), 0.90 (s, 9H), 0.08 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 136.70, 116.38, 73.03, 66.87, 25.81, 18.32, –5.38.



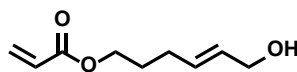
**Synthesis of ethyl (*E*)-6-((*tert*-butyldimethylsilyl)oxy)hex-4-enoate (2.37).** To a solution of 1-((*tert*-butyldimethylsilyl)oxy)but-3-en-2-ol (2.15 g, 10.6 mmol, 1 equiv.) in triethyl orthoacetate (13.6 mL, 74.4 mmol, 7 equiv.) was added propionic acid (0.075 mL, 1.0 mmol, 0.094 equiv.) and the mixture was heated to 140 °C under N<sub>2</sub>. After 4 h, unreacted triethyl orthoacetate was removed by vacuum distillation. The crude material was purified via flash silica gel chromatography, eluting with 2.5% EtOAc/hexanes to afford a light yellow oil (1.23 g, 43%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.70–5.52 (m, 2H), 4.16–4.08 (m, 4H), 2.40–2.33 (m, 4H), 1.24 (t, *J* = 7.1 Hz, 3H), 0.89 (s, 9H), 0.05 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 173.22, 130.51, 128.94, 63.86, 60.44, 34.06, 27.60, 26.10, 18.56, 14.39, –5.02.



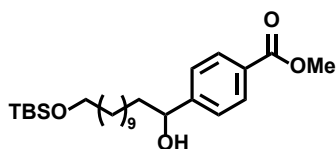
**Synthesis of (*E*)-6-((*tert*-butyldimethylsilyl)oxy)hex-4-en-1-ol (2.38).** To a solution of (*E*)-6-((*tert*-butyldimethylsilyl)oxy)hex-4-enoate (1.23 g, 4.51 mmol, 1 equiv.) in dry THF (20 mL) was added LiAlH<sub>4</sub> as a solution in THF (1.0 M, 11.3 mL, 11.29 mmol, 2.5 equiv.) at 0 °C under N<sub>2</sub>. After 10 min, the reaction was quenched with saturated aqueous K<sub>2</sub>CO<sub>3</sub> (10 mL) and stirred for 30 min at room temperature. The mixture was filtered through Celite and passed through a short plug of silica gel, eluting with 33% EtOAc/hexanes to afford a clear colorless oil (921 mg, 89%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.75–5.48 (m, 2H), 4.15–4.07 (m, 2H), 3.65 (t, *J* = 6.5 Hz, 2H), 2.16–2.08 (m, 2H), 1.70–1.61 (m, 2H), 0.90 (s, 9H), 0.06 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 130.56, 129.94, 64.02, 62.61, 32.24, 28.64, 26.12, 18.58, –4.98.



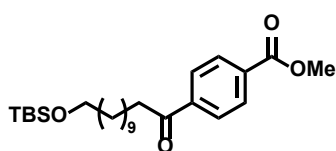
**Synthesis of (*E*)-6-((*tert*-butyldimethylsilyl)oxy)hex-4-en-1-yl acrylate (2.39).** To a solution of (*E*)-6-((*tert*-butyldimethylsilyl)oxy)hex-4-en-1-ol (3.9 g, 16.9 mmol, 1 equiv.) and Et<sub>3</sub>N (7.08 mL, 50.8 mmol, 3 equiv.) in dry THF (50 mL) was added acryloyl chloride (2.74 mL, 33.9 mmol, 2 equiv.) slowly at 0 °C under N<sub>2</sub>. The solution was warmed to room temperature and stirred for 12 h. After cooling to at 0 °C, H<sub>2</sub>O (20 mL) was added and the mixture was extracted into Et<sub>2</sub>O three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10% EtOAc/hexanes to afford a pale yellow oil (3.63 g, 76%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.40 (dd, *J* = 17.4, 1.5 Hz, 1H), 6.12 (dd, *J* = 17.3, 10.4 Hz, 1H), 5.82 (dd, *J* = 10.4, 1.6 Hz, 1H), 5.70–5.52 (m, 2H), 4.16 (t, *J* = 6.6 Hz, 2H), 4.14–4.10 (m, 2H), 2.20–2.06 (m, 2H), 1.84–1.71 (m, 2H), 0.90 (s, 9H), 0.06 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 166.40, 130.69, 130.34, 129.70, 128.69, 64.15, 63.96, 28.66, 28.26, 26.13, 18.59, –4.98.



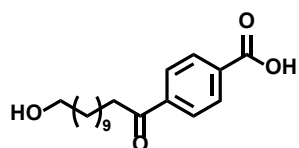
**Synthesis of (*E*)-6-hydroxyhex-4-en-1-yl acrylate (2.40).** To a solution of (*E*)-6-((*tert*-butyldimethylsilyl)oxy)hex-4-en-1-yl acrylate (547 mg, 1.92 mmol, 1 equiv.) in dry THF (20 mL) was added a solution of TBAF in THF (1.0 M, 2.12 mL, 2.12 mmol, 1.1 equiv.) at 0 °C under N<sub>2</sub>. After 5 min, the solution was warmed to room temperature and stirred for 1.5 h. Volatiles were removed under reduced pressure and the crude material was purified by via flash silica gel chromatography, eluting with 33% EtOAc/hexanes to afford a pale yellow oil (304 mg, 93%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.40 (dd, *J* = 17.3, 1.5 Hz, 1H), 6.12 (dd, *J* = 17.3, 10.4 Hz, 1H), 5.82 (dd, *J* = 10.4, 1.5 Hz, 1H), 5.68 (ddd, *J* = 5.8, 4.0, 2.2 Hz, 2H), 4.16 (t, *J* = 6.6 Hz, 2H), 4.09 (dd, *J* = 3.4, 1.9 Hz, 2H), 2.15 (tdd, *J* = 6.5, 4.9, 2.1 Hz, 2H), 1.83–1.72 (m, 2H), 1.48 (s, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 166.41, 131.66, 130.81, 130.11, 128.62, 63.99, 63.74, 28.67, 28.13.



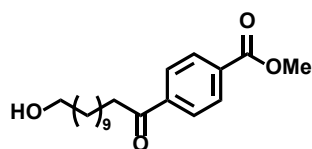
**Synthesis of methyl 4-(12-((*tert*-butyldimethylsilyl)oxy)-1-hydroxydodecyl)benzoate (2.44).** In a flame-dried flask, ((11-bromoundecyl)oxy)(*tert*-butyl)dimethylsilane (2.67 g, 7.31 mmol, 1.2 equiv.) was added dropwise to freshly etched Mg turnings (355 mg, 14.60 mmol, 2.4 equiv.) in dry THF (20 mL). The mixture was heated to reflux for 2.5 h under N<sub>2</sub> and then cooled to room temperature. To the resultant solution of (11-((*tert*-butyldimethylsilyl)oxy)undecyl)magnesium bromide was swiftly added a solution of methyl 4-formylbenzoate (1.0 g, 6.09 mmol, 1 equiv.) in dry THF (10 mL) at 0 °C. After stirring for 2 h, the reaction was warmed to room temperature and stirred for an additional 1 h. The reaction mixture was cooled to 0 °C and quenched with saturated aqueous NH<sub>4</sub>Cl (25 mL) and diluted with Et<sub>2</sub>O (50 mL). After stirring vigorously for 1 h, the layers were separated and the aqueous layer was extracted into Et<sub>2</sub>O three times. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 10% EtOAc/hexanes to afford a colorless oil (1.64 g, 60%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.99 (d, *J* = 8.3 Hz, 2H), 7.40 (d, *J* = 8.3 Hz, 2H), 4.72 (t, *J* = 6.7 Hz, 1H), 3.90 (s, 3H), 3.58 (t, *J* = 6.6 Hz, 2H), 2.13 (s, 1H), 1.85–1.64 (m, 2H), 1.57–1.43 (m, 2H), 1.41–1.14 (m, 16H), 0.88 (s, 9H), 0.03 (s, 6H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 162.45, 150.27, 129.86, 125.91, 74.27, 63.48, 52.21, 39.33, 33.00, 29.73, 29.65, 26.12, 25.78, 18.52, 14.32, –5.12.



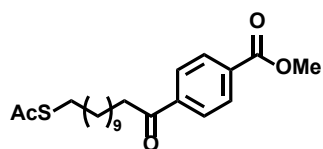
**Synthesis of methyl 4-(12-((*tert*-butyldimethylsilyl)oxy)dodecanoyl)benzoate (2.45).** To a mixture of MnO<sub>2</sub> (2.53 g, 29.11 mmol, 8 equiv.) in CH<sub>2</sub>Cl<sub>2</sub> (25 mL) was added methyl 4-(12-((*tert*-butyldimethylsilyl)oxy)-1-hydroxydodecyl)benzoate (1.64 g, 3.64 mmol, 1 equiv.) and the mixture was stirred at 40 °C for 12 h. The reaction mixture was cooled to room temperature and filtered through Celite, and concentrated under reduced pressure. The crude material was purified via flash silica gel chromatography, eluting with 2–20% EtOAc/hexanes to afford a white solid (1.31 g, 80%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.12 (dd, *J* = 8.5, 2.0 Hz, 2H), 8.05–7.97 (m, 2H), 3.95 (s, 3H), 3.59 (t, *J* = 6.6 Hz, 2H), 2.98 (t, *J* = 7.4 Hz, 2H), 1.73 (p, *J* = 7.3 Hz, 2H), 1.57–1.44 (m, 2H), 1.26 (m, 14H), 0.99–0.83 (m, 9H), 0.04 (s, 6H).



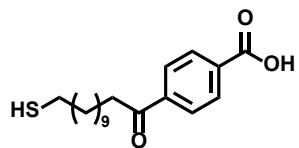
**Synthesis of 4-(12-hydroxydodecanoyl)benzoic acid (2.43).** To a solution of 4-(12-((*tert*-butyldimethylsilyl)oxy)dodecanoyl)benzoate (1.09 g, 2.43 mmol) in 4:3 MeOH/THF (35 mL) was added 40% aqueous NaOH (3 mL) and the reaction was stirred at 70 °C for 3 h. After cooling to room temperature, the mixture was acidified with 20% aqueous H<sub>2</sub>SO<sub>4</sub> and extracted into Et<sub>2</sub>O. The Et<sub>2</sub>O extract was concentrated and taken up in THF (10 mL). To this solution was added 50% TFA in H<sub>2</sub>O at room temperature. After 30 min, THF was removed under reduced pressure and the product precipitated as a white solid, which was collected by filtration (600 mg, 57%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.02 (s, 4H), 3.34 (t, *J* = 6.6 Hz, 2H), 3.00 (t, *J* = 7.2 Hz, 2H), 1.57 (p, *J* = 7.2 Hz, 2H), 1.36 (q, *J* = 6.8 Hz, 2H), 1.32–1.15 (m, 14H). LRMS (ESI) calculated for C<sub>19</sub>H<sub>27</sub>O<sub>4</sub><sup>-</sup> ([M-H]<sup>-</sup>) 320.2, found 320.3.



**Synthesis of methyl 4-(12-hydroxydodecanoyl)benzoate (2.47).** To a solution of methyl 4-(12-((*tert*-butyldimethylsilyl)oxy)dodecanoyl)benzoate (1.2 g, 2.67 mmol) in 1:2 MeOH/THF (30 mL) was added 15 drops of 1 M aqueous HCl at room temperature. After 1.5 h, the resulting precipitate was collected by filtration and washed with cold MeOH to afford the desired product as a white solid (822 mg, 92%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.12 (d, *J* = 8.2 Hz, 2H), 8.00 (d, *J* = 8.2 Hz, 2H), 3.95 (s, 3H), 3.64 (s, 2H), 2.98 (t, *J* = 7.4 Hz, 2H), 1.81–1.68 (m, 2H), 1.29 (d, *J* = 21.0 Hz, 16H).

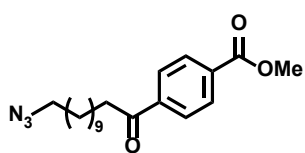


**Synthesis of methyl 4-(12-(acetylthio)dodecanoyl)benzoate (2.48).** To a suspension of methyl 4-(12-hydroxydodecanoyl)benzoate (400 mg, 1.20 mmol, 1 equiv.) and Et<sub>3</sub>N (0.21 mL, 1.50 mmol, 1.25 equiv.) in dry CH<sub>2</sub>Cl<sub>2</sub> (15 mL) was added methanesulfonyl chloride (0.11 mL, 1.32 mmol, 1.10 equiv.) under N<sub>2</sub> at 0 °C. The reaction was warmed to room temperature and stirred for 30 min. The reaction was quenched with saturated aqueous NH<sub>4</sub>Cl. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure to afford an off-white solid. To a solution of the crude mesylate in dry DMF (15 mL) was added KSAc (206 mg, 1.80 mmol, 1.80 equiv.) and the mixture was stirred at room temperature under N<sub>2</sub> for 3 h. The mixture was concentrated, taken up in CH<sub>2</sub>Cl<sub>2</sub>, washed with saturated aqueous NaHCO<sub>3</sub>, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude residue was purified by flash silica gel chromatography, eluting with 10% EtOAc/hexanes to afford a tan solid (320 mg, 68%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.15–8.09 (m, 2H), 8.03–7.98 (m, 2H), 3.95 (s, 3H), 2.98 (t, *J* = 7.4 Hz, 2H), 2.86 (t, *J* = 7.4 Hz, 2H), 2.32 (s, 3H), 1.73 (p, *J* = 7.3 Hz, 2H), 1.55 (p, *J* = 7.6, 7.2 Hz, 2H), 1.43–1.21 (m, 14H).

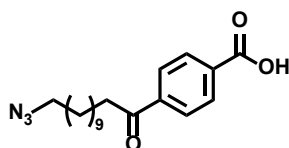


**Synthesis of 4-(12-mercaptododecanoyl)benzoic acid (2.46).** To a solution of methyl 4-(12-((methylsulfonyl)oxy)dodecanoyl)benzoate (50 mg, 0.13 mmol) in MeOH (2 mL) was added 10% aqueous NaOH (2 mL) followed by THF (2 mL) at room temperature. After 1 h, volatiles were removed under reduced pressure and the aqueous solution was acidified to pH 1 with 1 M aqueous HCl. The resultant precipitate was collected by centrifugation and lyophilized to afford an off-white solid (8 mg, 18%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.00 (d, *J* = 3.8 Hz, 4H), 3.07–2.95 (m, 2H), 2.69–2.60 (m, 2H), 1.56 (d, *J* = 8.6 Hz, 2H), 1.51–1.45 (m, 2H), 1.21 (s, 14H). LRMS (ESI) calculat-

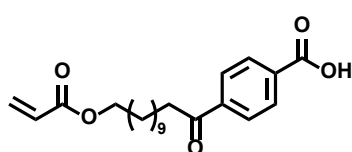
ed for  $C_{19}H_{27}O_3S^-$  ( $[M-H]^-$ ) 335.2, found 335.2.



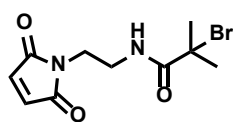
**Synthesis of methyl 4-(12-azidododecanoyl)benzoate (2.50).** To a solution of methyl 4-(12-hydroxydodecanoyl)benzoate (200 mg, 0.60 mmol, 1 equiv.) and diphenylphosphoryl azide (0.15 mL, 0.72 mmol, 1.2 equiv.) in dry DMF (1 mL) was added DBU (0.11 mL, 0.72 mmol, 1.2 equiv.) at room temperature under  $N_2$ . After 30 min, the reaction was heated to 65 °C for 30 min and then 90 °C for 30 min. The reaction was cooled to room temperature and diluted with EtOAc. The solution was washed with half saturated brine. The organic layer was dried over  $Na_2SO_4$  and concentrated under reduced pressure. The crude residue was purified by flash silica gel chromatography, eluting with 10–100% EtOAc/hexanes to afford a white solid (110 mg, 51%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  8.11 (dt,  $J = 8.5, 1.9$  Hz, 2H), 8.00 (dt,  $J = 8.5, 1.9$  Hz, 2H), 3.95 (s, 3H), 3.25 (t,  $J = 7.0$  Hz, 2H), 2.98 (t,  $J = 7.4$  Hz, 2H), 1.73 (p,  $J = 7.4$  Hz, 2H), 1.65–1.52 (m, 2H), 1.44–1.20 (m, 14H).



**Synthesis of 4-(12-azidododecanoyl)benzoic acid (2.49).** To a solution of methyl 4-(12-azidododecanoyl)benzoate (110 mg, 0.31 mmol) in 50% MeOH/THF (4 mL) was added 10% aqueous NaOH (4 mL) at room temperature. After 3 h, volatiles were removed under reduced pressure and the remaining aqueous solution was acidified to pH 1 with 1 M aqueous HCl. The solution was extracted with  $Et_2O$  twice and the combined organic layers were dried over  $Na_2SO_4$  and concentrated under reduced pressure to afford the desired product as a white solid (72 mg, 73%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  10.50 (br s, 1H), 8.18 (d,  $J = 7.9$  Hz, 2H), 8.02 (d,  $J = 8.1$  Hz, 2H), 3.24 (t,  $J = 6.9$  Hz, 2H), 2.99 (t,  $J = 7.3$  Hz, 2H), 1.73 (p,  $J = 7.3$  Hz, 2H), 1.58 (p,  $J = 7.1$  Hz, 2H), 1.49–1.12 (m, 14H).  $^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta$  200.23, 167.59, 140.80, 133.52, 130.46, 128.09, 51.56, 39.11, 29.53, 29.36, 29.22, 28.91, 26.79, 24.24. LRMS (ESI) calculated for  $C_{19}H_{26}N_3O_3^-$  ( $[M-H]^-$ ) 344.2, found 344.1.



**Synthesis of 4-(12-(acryloyloxy)dodecanoyl)benzoic acid (2.51).** To a suspension of 4-(12-hydroxydodecanoyl)benzoic acid (25 mg, 0.078 mmol, 1 equiv.) in dry THF (2 mL) was added dimethylaniline (0.011 mL, 0.086 mmol, 1.1 equiv.) under  $N_2$  and the mixture was heated to 60 °C. Acryloyl chloride (0.038 mL, 0.468 mmol, 6 equiv.) was added to the hot mixture and the reaction was stirred for 3 h. Volatiles were removed under reduced pressure and the crude material was recrystallized from MeOH to afford the product as a white solid (11 mg, 38%).  $^1H$  NMR (300 MHz,  $DMSO-d_6$ )  $\delta$  8.04 (s, 4H), 6.31 (dd,  $J = 17.2, 1.9$  Hz, 1H), 6.15 (dd,  $J = 17.2, 10.1$  Hz, 1H), 5.92 (dd,  $J = 10.2, 1.8$  Hz, 1H), 4.08 (t,  $J = 6.6$  Hz, 2H), 3.04 (t,  $J = 7.2$  Hz, 2H), 1.71–1.48 (m, 4H), 1.40–1.16 (m, 14H). LRMS (ESI) calculated for  $C_{22}H_{29}O_5^-$  ( $[M-H]^-$ ) 373.2, found 373.4.



**Synthesis of 2-bromo-N-(2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)-2-methylpropanamide (2.52).** To a solution of 2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethan-1-aminium trifluoroacetate (150 mg, 0.59 mmol, 1 equiv.) and  $Et_3N$  (0.27 mL, 1.92 mmol, 3.25 equiv.) in  $CH_2Cl_2$  (5 mL) was added 2-bromo-2-methylpropionyl bromide (0.12 mL, 0.94 mmol, 1.6 equiv.) dropwise at 0 °C under  $N_2$ . After 2 h, the reaction was warmed to room temperature and stirred for 30 min. The reaction mixture was washed three times with water (5 mL) and the organic layer was dried over  $Na_2SO_4$  and

concentrated under pressure. The crude mixture was purified by flash silica gel chromatography, eluting with 25–100% EtOAc/hexanes to afford a white solid (144 mg, 84%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.00 (s, 1H), 6.73 (s, 2H), 3.77–3.71 (m, 2H), 3.52–3.45 (m, 2H), 1.91 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 134.62, 62.48, 39.85, 37.36, 32.49.

## 2.7 References

- (1) Steinmetz, N. F. Viral Nanoparticles as Platforms for Next-Generation Therapeutics and Imaging Devices. *Nanomed. Nanotech. Biol. Med.* **2010**, *6* (5), 634–641.
- (2) Ma, Y.; Nolte, R. J. M.; Cornelissen, J. J. L. M. Virus-Based Nanocarriers for Drug Delivery. *Adv. Drug Deliv. Rev.* **2012**, *64* (9), 811–825.
- (3) Yildiz, I.; Shukla, S.; Steinmetz, N. F. Applications of Viral Nanoparticles in Medicine. *Curr. Opin. Biotechnol.* **2011**, *22* (6), 901–908.
- (4) Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery. *Angew. Chem. Int. Ed.* **2009**, *48* (50), 9493–9497.
- (5) Zhao, Q.; Chen, W.; Chen, Y.; Zhang, L.; Zhang, J.; Zhang, Z. Self-Assembled Virus-Like Particles From Rotavirus Structural Protein VP6 for Targeted Drug Delivery. *Bioconjugate Chem.* **2011**, *22* (3), 346–352.
- (6) Flenniken, M. L.; Liepold, L. O.; Crowley, B. E.; Willits, D. A.; Young, M. J.; Douglas, T. Selective Attachment and Release of a Chemotherapeutic Agent From the Interior of a Protein Cage Architecture. *Chem. Commun.* **2005**, *4*, 447–3.
- (7) Lucon, J.; Qazi, S.; Uchida, M.; Bedwell, G. J.; LaFrance, B.; Prevelige, P. E.; Douglas, T. Use of the Interior Cavity of the P22 Capsid for Site-Specific Initiation of Atom-Transfer Radical Polymerization with High-Density Cargo Loading. *Nat. Chem.* **2012**, *4* (10), 781–788.
- (8) Zeng, Q.; Wen, H.; Wen, Q.; Chen, X.; Wang, Y.; Xuan, W.; Liang, J.; Wan, S. Cucumber Mosaic Virus as Drug Delivery Vehicle for Doxorubicin. *Biomaterials* **2013**, *34* (19), 4632–4642.
- (9) 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.
- (10) Yildiz, I.; Lee, K. L.; Chen, K.; Shukla, S.; Steinmetz, N. F. Infusion of Imaging and Therapeutic Molecules Into the Plant Virus-Based Carrier Cowpea Mosaic Virus: Cargo-Loading and Delivery. *J. Controlled Release* **2013**, *172* (2), 568–578.
- (11) Glasgow, J. E.; Capehart, S. L.; Francis, M. B.; Tullman-Ercek, D. Osmolyte-Mediated Encapsulation of Proteins Inside MS2 Viral Capsids. *ACS Nano* **2012**, *6* (10), 8658–8664.
- (12) Ren, Y.; Wong, S. M.; Lim, L.-Y. Folic Acid-Conjugated Protein Cages of a Plant Virus: a Novel Delivery Platform for Doxorubicin. *Bioconjugate Chem.* **2007**, *18* (3), 836–843.



- (13) Kwak, M.; Minten, I. J.; Anaya, D.-M.; Musser, A. J.; Brasch, M.; Nolte, R. J. M.; Müllen, K.; Cornelissen, J. J. L. M.; Herrmann, A. Virus-Like Particles Templated by DNA Micelles: a General Method for Loading Virus Nanocarriers. *J. Am. Chem. Soc.* **2010**, *132* (23), 7834–7835.
- (14) Hermanson, G. T. *Bioconjugate Techniques*; Elsevier: Oxford, 2013.
- (15) Li, J.; Chen, P. R. Development and Application of Bond Cleavage Reactions in Bioorthogonal Chemistry. *Nat. Chem. Biol.* **2016**, *12* (3), 129–137.
- (16) Pawlak, J. B.; Gentil, G. P. P.; Ruckwardt, T. J.; Bremmers, J. S.; Meeuwenoord, N. J.; Ossendorp, F. A.; Overkleeft, H. S.; Filippov, D. V.; van Kasteren, S. I. Bioorthogonal Deprotection on the Dendritic Cell Surface for Chemical Control of Antigen Cross-Presentation. *Angew. Chem. Int. Ed.* **2015**, *127* (19), 5720–5723.
- (17) Verhelst, S. H. L.; Fonović, M.; Bogyo, M. A Mild Chemically Cleavable Linker System for Functional Proteomic Applications. *Angew. Chem. Int. Ed.* **2007**, *46* (8), 1284–1286.
- (18) Ulbrich, K.; Etrych, T.; Chytil, P.; Jelínková, M.; Říhová, B. HPMA Copolymers with pH-Controlled Release of Doxorubicin. *J. Control. Release* **2003**, *87* (1-3), 33–47.
- (19) Castañeda, L.; Maruani, A.; Schumacher, F. F.; Miranda, E.; Chudasama, V.; Chester, K. A.; Baker, J. R.; Smith, M. E. B.; Caddick, S. Acid-Cleavable Thiomaleamic Acid Linker for Homogeneous Antibody-Drug Conjugation. *Chem. Commun.* **2013**, *49* (74), 8187–8189.
- (20) Klán, P.; Šolomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J. Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. *Chem. Rev.* **2013**, *113* (1), 119–191.
- (21) Streu, C.; Meggers, E. Ruthenium-Induced Allylcarbamate Cleavage in Living Cells. *Angew. Chem. Int. Ed.* **2006**, *45* (34), 5645–5648.
- (22) 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.
- (23) Kislukhin, A. A.; Hong, V. P.; Breitenkamp, K. E.; Finn, M. G. Relative Performance of Alkynes in Copper-Catalyzed Azide–Alkyne Cycloaddition. *Bioconjugate Chem.* **2013**, *24* (4), 684–689.
- (24) Versteegen, R. M.; Rossin, R.; Hoeve, ten, W.; Janssen, H. M.; Robillard, M. S. Click to Release: Instantaneous Doxorubicin Elimination Upon Tetrazine Ligation. *Angew. Chem. Int. Ed.* **2013**, *52* (52), 14112–14116.
- (25) Matikonda, S. S.; Orsi, D. L.; Staudacher, V.; Jenkins, I. A.; Fiedler, F.; Chen, J.; Gamble, A. B. Bioorthogonal Prodrug Activation Driven by a Strain-Promoted 1,3-Dipolar Cycloaddition. *Chem. Sci.* **2015**, *6*, 1212–1218.
- (26) Tilley, S. D.; Francis, M. B. Tyrosine-Selective Protein Alkylation Using  $\pi$ -Allylpalladium Complexes. *J. Am. Chem. Soc.* **2006**, *128* (4), 1080–1081.

- (27) Takacs, J. M.; Lawson, E. C.; Clement, F. On the Nature of the Catalytic Palladium-Mediated Elimination of Allylic Carbonates and Acetates to Form 1,3-Dienes. *J. Am. Chem. Soc.* **1997**, *119* (25), 5956–5957.
- (28) Stephanopoulos, N.; Carrico, Z. M.; Francis, M. B. Nanoscale Integration of Sensitizing Chromophores and Porphyrins with Bacteriophage MS2. *Angew. Chem. Int. Ed.* **2009**, *48* (50), 9498–9502.
- (29) Spicer, C. D.; Davis, B. G. Palladium-Mediated Site-Selective Suzuki–Miyaura Protein Modification at Genetically Encoded Aryl Halides. *Chem. Commun.* **2011**, *47* (6), 1698–1700.
- (30) Baldwin, A. D.; Kiick, K. L. Tunable Degradation of Maleimide–Thiol Adducts in Reducing Environments. *Bioconjugate Chem.* **2011**, *22* (10), 1946–1953.
- (31) Chalker, J. M.; Wood, C. S. C.; Davis, B. G. A Convenient Catalyst for Aqueous and Protein Suzuki–Miyaura Cross-Coupling. *J. Am. Chem. Soc.* **2009**, *131* (45), 16346–16347.
- (32) Norrish, R.; Bamford, C. H. Photodecomposition of Aldehydes and Ketones. *Nature* **1936**, *138*, 1016.
- (33) Yang, N. C.; Yang, D. Photochemical Reactions of Ketones in Solution. *J. Am. Chem. Soc.* **1958**, *80* (11), 2913–2914.
- (34) Lewis, F. D.; Hilliard, T. A. Photochemistry of Methyl-Substituted Butyrophenones. Nature of the 1,4-Biradical Intermediates. *J. Am. Chem. Soc.* **1970**, *92* (22), 6672–6674.
- (35) Epstein, W. W.; Jones, D. S.; Bruenger, E.; Rilling, H. C. The Synthesis of a Photolabile Detergent and Its Use in the Isolation and Characterization of Protein. *Anal. Biochem.* **1982**, *119* (2), 304–312.
- (36) Levrand, B.; Herrmann, A. Light Induced Controlled Release of Fragrances by Norrish Type II Photofragmentation of Alkyl Phenyl Ketones. *Photochem. Photobiol. Sci.* **2002**, *1* (11), 907–919.
- (37) Node, M.; Nishide, K.; Fuji, K.; Fujita, E. Hard Acid and Soft Nucleophile System. 2. Demethylation of Methyl Ethers of Alcohol and Phenol with an Aluminum Halide–Thiol System. *J. Org. Chem.* **1980**, *45* (22), 4275–4277.
- (38) Gomez, C.; Chen, J.; Wang, S. Efficient Synthesis of Phosphotyrosine Building Blocks Using Imidazolium Trifluoroacetate. *Tetrahedron Lett.* **2009**, *50* (48), 6691–6692.
- (39) Walker, M. A. A High Yielding Synthesis of *N*-Alkyl Maleimides Using a Novel Modification of the Mitsunobu Reaction. *J. Org. Chem.* **1995**, *60* (16), 5352–5355.
- (40) Kato, A.; Nakai, S. Hydrophobicity Determined by a Fluorescence Probe Method and Its Correlation with Surface Properties of Proteins. *Biochim. Biophys. Acta* **1980**, *624* (1), 13–20.
- (41) Haskard, C. A.; Li-Chan, E. C. Y. Hydrophobicity of Bovine Serum Albumin and Ovalbumin Determined Using Uncharged (PRODAN) and Anionic (ANS) Fluorescent Probes.

- J. Agric. Food Chem.* **1998**, 46 (7), 2671–2677.
- (42) Chao, C. C.; Ma, Y. S.; Stadtman, E. R. Modification of Protein Surface Hydrophobicity and Methionine Oxidation by Oxidative Systems. *Proc. Natl. Acad. Sci. USA* **1997**, 94 (7), 2969–2974.
- (43) Matulis, D.; Lovrien, R. 1-Anilino-8-Naphthalene Sulfonate Anion-Protein Binding Depends Primarily on Ion Pair Formation. *Biophys. J.* **1998**, 74 (1), 422–429.
- (44) Sackett, D. L.; Wolff, J. Nile Red as a Polarity-Sensitive Fluorescent Probe of Hydrophobic Protein Surfaces. *Anal. Biochem.* **1987**, 167 (2), 228–234.
- (45) Pattison, D. I.; Rahmanto, A. S.; Davies, M. J. Photo-Oxidation of Proteins. *Photochem. Photobiol. Sci.* **2012**, 11 (1), 38–53.
- (46) Kerwin, B. A.; Remmele, R. L. Protect From Light: Photodegradation and Protein Biology. *J. Pharm. Sci.* **2007**, 96 (6), 1468–1479.
- (47) Lam, X. M.; Yang, J. Y.; Cleland, J. L. Antioxidants for Prevention of Methionine Oxidation in Recombinant Monoclonal Antibody HER2. *J. Pharm. Sci.* **1997**, 86 (11), 1250–1255.
- (48) Bancirova, M. Sodium Azide as a Specific Quencher of Singlet Oxygen During Chemiluminescent Detection by Luminol and *Cypridina* Luciferin Analogues. *Luminescence* **2011**, 26 (6), 685–688.
- (49) Bleeke, T. Catecholamine-Induced Vascular Wall Growth Is Dependent on Generation of Reactive Oxygen Species. *Circ. Res.* **2004**, 94 (1), 37–45.
- (50) Ji, J. A.; Zhang, B.; Cheng, W.; Wang, Y. J. Methionine, Tryptophan, and Histidine Oxidation in a Model Protein, PTH: Mechanisms and Stabilization. *J. Pharm. Sci.* **2009**, 98 (12), 4485–4500.
- (51) Lam, X. M.; Lai, W. G.; Chan, E. K.; Ling, V.; Hsu, C. C. Site-Specific Tryptophan Oxidation Induced by Autocatalytic Reaction of Polysorbate 20 in Protein Formulation. *Pharm. Res.* **2011**, 28 (10), 2543–2555.
- (52) Creed, D. The Photophysics and Photochemistry of the Near-UV Absorbing Amino Acids—I. Tryptophan and Its Simple Derivatives. *Photochem. Photobiol.* **1984**, 39 (4), 537–562.
- (53) Grewal, P.; Mallaney, M.; Lau, K.; Sreedhara, A. Screening Methods to Identify Indole Derivatives That Protect Against Reactive Oxygen Species Induced Tryptophan Oxidation in Proteins. *Mol. Pharmaceutics* **2014**, 11 (4), 1259–1272.
- (54) Singh, A. K.; Khade, P. K. 3-Nitro-2-Naphthalenemethanol: a Photocleavable Protecting Group for Carboxylic Acids. *Tetrahedron* **2005**, 61 (42), 10007–10012.
- (55) Singh, A. K.; Khade, P. K. 7-Methoxy-3-Nitro-2-Naphthalenemethanol—a New Phototrigger for Caging Applications. *Tetrahedron Lett.* **2011**, 52 (38), 4899–4902.
- (56) Khade, P. K.; Singh, A. K. A New Caging Phototrigger Based on a 2-Acetonaphthyl Chro-

- mophore. *Tetrahedron Lett.* **2007**, 48 (39), 6920–6923.
- (57) Furuta, T.; Torigai, H.; Sugimoto, M. Photochemical Properties of New Photolabile cAMP Derivatives in a Physiological Saline Solution. *J. Org. Chem.* **1995**, 60 (13), 3953–3956.
- (58) Kulikov, A.; Arumugam, S.; Popik, V. V. Photolabile Protection of Alcohols, Phenols, and Carboxylic Acids with 3-Hydroxy-2-Naphthalenemethanol. *J. Org. Chem.* **2008**, 73 (19), 7611–7615.
- (59) Arumugam, S.; Popik, V. V. Attach, Remove, or Replace: Reversible Surface Functionalization Using Thiol–Quinone Methide Photoclick Chemistry. *J. Am. Chem. Soc.* **2012**, 134 (20), 8408–8411.
- (60) Thevenet, D.; Neier, R. An Efficient Photoinduced Deprotection of Aromatic Acetals and Ketals. *Helv. Chim. Acta.* **2011**, 94 (2), 331–346.
- (61) Chapman, R.; Gormley, A. J.; Stenzel, M. H.; Stevens, M. M. Combinatorial Low-Volume Synthesis of Well-Defined Polymers by Enzyme Degassing. *Angew. Chem. Int. Ed.* **2016**, 55 (14), 4500–4503.
- (62) Twaik, M. A.; Tahan, M.; Zilkha, A. Grafting of Poly(Ethylene Oxide) on Poly(Methyl Methacrylate) by Transesterification. *J. Polym. Sci., Part A: Polym. Chem.* **1969**, 7 (9), 2469–2480.
- (63) Otera, J. Transesterification. *Chem. Rev.* **1993**, 93 (4), 1449–1470.
- (64) Benaglia, M.; Alberti, A.; Giorgini, L.; Magnoni, F.; Tozzi, S. Poly(Glycidyl Methacrylate): a Highly Versatile Polymeric Building Block for Post-Polymerization Modifications. *Polym. Chem.* **2012**, 4 (1), 124–132.
- (65) Gadwal, I.; Stuparu, M. C.; Khan, A. Homopolymer Bifunctionalization Through Sequential Thiol–Epoxy and Esterification Reactions: an Optimization, Quantification, and Structural Elucidation Study. *Polym. Chem.* **2015**, 6 (8), 1393–1404.
- (66) Li, Q.; Bao, Y.; Wang, H.; Du, F.; Li, Q.; Jin, B.; Bai, R. A Facile and Highly Efficient Strategy for Esterification of Poly(Meth)Acrylic Acid with Halogenated Compounds at Room Temperature Promoted by 1,1,3,3-Tetramethylguanidine. *Polym. Chem.* **2013**, 4 (9), 2891–2897.
- (67) Parrish, B.; Breitenkamp, R. B.; Emrick, T. PEG- and Peptide-Grafted Aliphatic Polyesters by Click Chemistry. *J. Am. Chem. Soc.* **2005**, 127 (20), 7404–7410.
- (68) Wallat, J. D.; Rose, K. A.; Pokorski, J. K. Proteins as Substrates for Controlled Radical Polymerization. *Polym. Chem.* **2014**, 5 (5), 1545–1558.
- (69) Lucon, J.; Edwards, E.; Qazi, S.; Uchida, M.; Douglas, T. Atom Transfer Radical Polymerization on the Interior of the P22 Capsid and Incorporation of Photocatalytic Monomer Crosslinks. *Eur. Polym. J.* **2013**, 49 (10), 2976–2985.
- (70) Hovlid, M. L.; Lau, J. L.; Breitenkamp, K.; Higginson, C. J.; Laufer, B.; Manchester, M.; Finn, M. G. Encapsidated Atom-Transfer Radical Polymerization in Q $\beta$  Virus-Like

Nanoparticles. *ACS Nano* **2014**, *8* (8), 8003–8014.

- (71) Averick, S.; Simakova, A.; Park, S.; Konkolewicz, D.; Magenau, A. J. D.; Mehl, R. A.; Matyjaszewski, K. ATRP Under Biologically Relevant Conditions: Grafting From a Protein. *ACS Macro Lett.* **2012**, *1* (1), 6–10.
- (72) Hooker, J. M.; Datta, A.; Botta, M.; Raymond, K. N.; Francis, M. B. Magnetic Resonance Contrast Agents From Viral Capsid Shells: a Comparison of Exterior and Interior Cargo Strategies. *Nano Lett.* **2007**, *7* (8), 2207–2210.
- (73) Tong, G. J.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. Viral Capsid DNA Aptamer Conjugates as Multivalent Cell-Targeting Vehicles. *J. Am. Chem. Soc.* **2009**, *131* (31), 11174–11178.
- (74) Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. Oxidative Coupling of Peptides to a Virus Capsid Containing Unnatural Amino Acids. *Chem. Commun.* **2008**, *10*, 1205–1207.
- (75) Seim, K. L.; Obermeyer, A. C.; Francis, M. B. Oxidative Modification of Native Protein Residues Using Cerium(IV) Ammonium Nitrate. *J. Am. Chem. Soc.* **2011**, *133* (42), 16970–16976.
- (76) Galvagnion, C.; Smith, M. T. J.; Broom, A.; Vassall, K. A.; Meglei, G.; Gaspar, J. A.; Stathopoulos, P. B.; Cheyne, B.; Meiering, E. M. Folding and Association of Thermophilic Dimeric and Trimeric DsrEFH Proteins: Tm0979 and Mth1491. *Biochemistry* **2009**, *48* (13), 2891–2906.

## Chapter 3

### Elucidating the nature of conformational changes in peptide-CRM<sub>197</sub> conjugate vaccines

#### Abstract

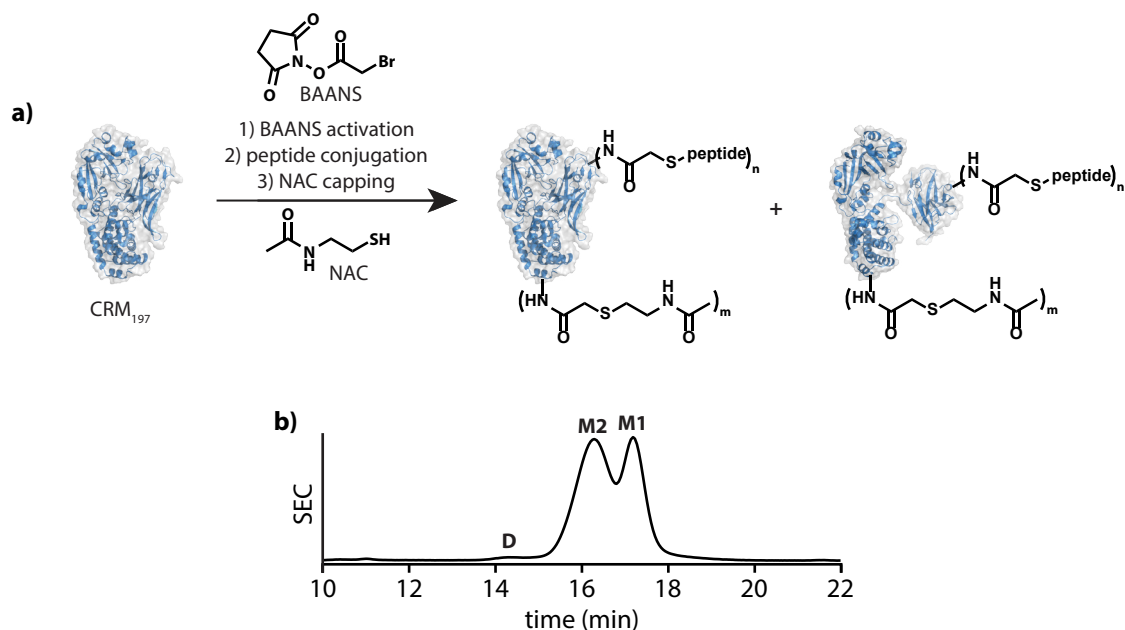
Conjugate vaccines prepared from the CRM<sub>197</sub> carrier protein have been successful in the clinic and are of great interest in the field of immunotherapy. Researchers at Pfizer are currently developing peptide-CRM<sub>197</sub> conjugate vaccines using an activation-conjugation strategy. This method has been found to result in two distinct populations of conjugates, believed to be the result of a conformational change during preparation. These two populations offer an interesting system for structure-immunogenicity relationship studies. In this chapter, the nature and causes of conformational change in peptide-CRM<sub>197</sub> conjugates will be explored using the bioconjugation toolbox. Efforts towards selective preparation of either of the two states will be discussed.

### 3.1 Introduction

Conventionally, vaccines have been prepared from live attenuated pathogens or inactivated viruses.<sup>1</sup> While these pathogen-derived products have had a profound effect on global health,<sup>2</sup> they present challenges in terms of production and safety.<sup>3</sup> New technologies have allowed for the rational design of novel and safer vaccines.<sup>4</sup> Conjugate vaccines represent one class of this new wave of technologies. These synthetic vaccines are composed of hapten molecules (generally a B-cell epitope without T-cell epitopes) covalently attached to a protein carrier (source of T-cell epitopes).<sup>5</sup> As these materials are not derived from live pathogens, many issues are avoided, which would otherwise arise in the production of traditional vaccines. Additionally, hapten-carrier conjugates can be prepared for more diverse targets than existing vaccines. Beyond bacterial and viral targets, conjugate vaccines can also be developed against non-infectious diseases including cancer,<sup>6</sup> Alzheimer's disease,<sup>7</sup> and hypertension.<sup>8</sup>

A variety of carrier proteins, including diphtheria toxoid, diphtheria toxin (DT) cross-reactive material 197 (CRM<sub>197</sub>; DT Gly52Glu), tetanus toxoid (TT), keyhole limpet hemocyanin (KLH), and virus-like particles (VLPs), have been used for the production of conjugate vaccines.<sup>5</sup> Of these, CRM<sub>197</sub> is the most common carrier used in research studies and in clinically available conjugate vaccines, perhaps due to its inherent lack of toxicity.<sup>9</sup> Significant efforts have been made to prepare well-defined glycoconjugates of CRM<sub>197</sub>. Most commonly, the 39 lysine residues and the N-terminus of CRM<sub>197</sub> are targeted for conjugation. However, carboxylate containing residues, tyrosines, and disulfides have also been targeted for selective modification.<sup>10</sup> In addition to preparing conjugate materials, studies have demonstrated structure-immunogenicity relationships dependent on hapten loading, conjugation chemistry, and site(s) of conjugation.<sup>10</sup> However, similar rigorous studies are rare in the scientific literature for peptide-CRM<sub>197</sub> conjugates, even though some of these materials are promising vaccine candidates.<sup>11-13</sup>

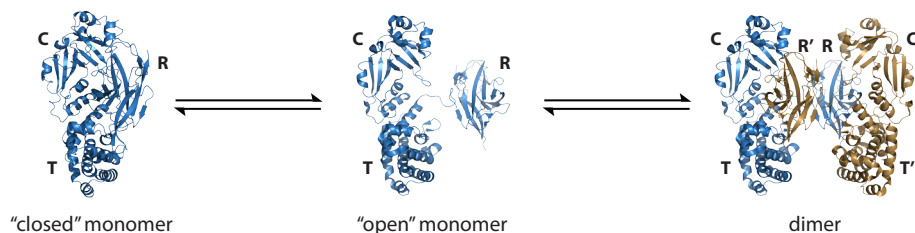
Pfizer is currently developing new well-defined peptide-CRM<sub>197</sub> conjugate vaccines using a patented activation-conjugation procedure (**Figure 3.1a**), in which solvent accessible lysines in CRM<sub>197</sub> are first activated with a bromoacetylating agent, BAANS (bromoacetic acid *N*-hydroxysuccinimide ester).<sup>14</sup> Cysteine-containing peptide epitopes of interest are then conjugated to the introduced bromoacetamides. Finally, unreacted bromoacetamide groups are capped with *N*-acetylcysteamine (NAC) to provide the final conjugate material. This chemistry has been optimized to provide the preferred number of a peptide epitopes (13 on average) for a given formulation. However, characterization by size-exclusion chromatography (SEC) has revealed these conjugates to be structurally inhomogeneous—in fact, three distinct populations have been observed (**Figure 3.1b**). By analytical ultracentrifugation (AUC), the two major species (M1 and M2) were determined to be monomeric, and the minor species (D) was determined to be a non-covalent dimer. It is important to note that M2 was not observed for unmodified CRM<sub>197</sub>, BAANS-activated CRM<sub>197</sub>, and NAC-capped BAANS-activated CRM<sub>197</sub>. An understanding of what these two monomeric populations comprise, chemically, and what their relative contributions are to the overall efficacy of the vaccine formulation is of great interest. In this chapter, work towards understanding the underlying causes of M1/M2 formation in CRM<sub>197</sub> conjugates will be discussed. Additionally, strategies for promoting or avoiding M2 formation will be presented.



**Figure 3.1.** Peptide-CRM<sub>197</sub> conjugates prepared via an activation-conjugation procedure were found to exist in two conformational states. a) Conjugates were prepared by first activating lysine residues with BAANS, and then conjugating a cysteine-containing peptide. Unreacted bromoacetamides were then capped with NAC. b) Characterization of the resulting conjugate material by SEC (monitored at 280 nm) revealed the presence of two populations of monomeric conjugates (M1 and M2) and one dimeric form (D) of the conjugate. M1 and M2 were believed to be conjugates of CRM<sub>197</sub> in a closed and open form, respectively.

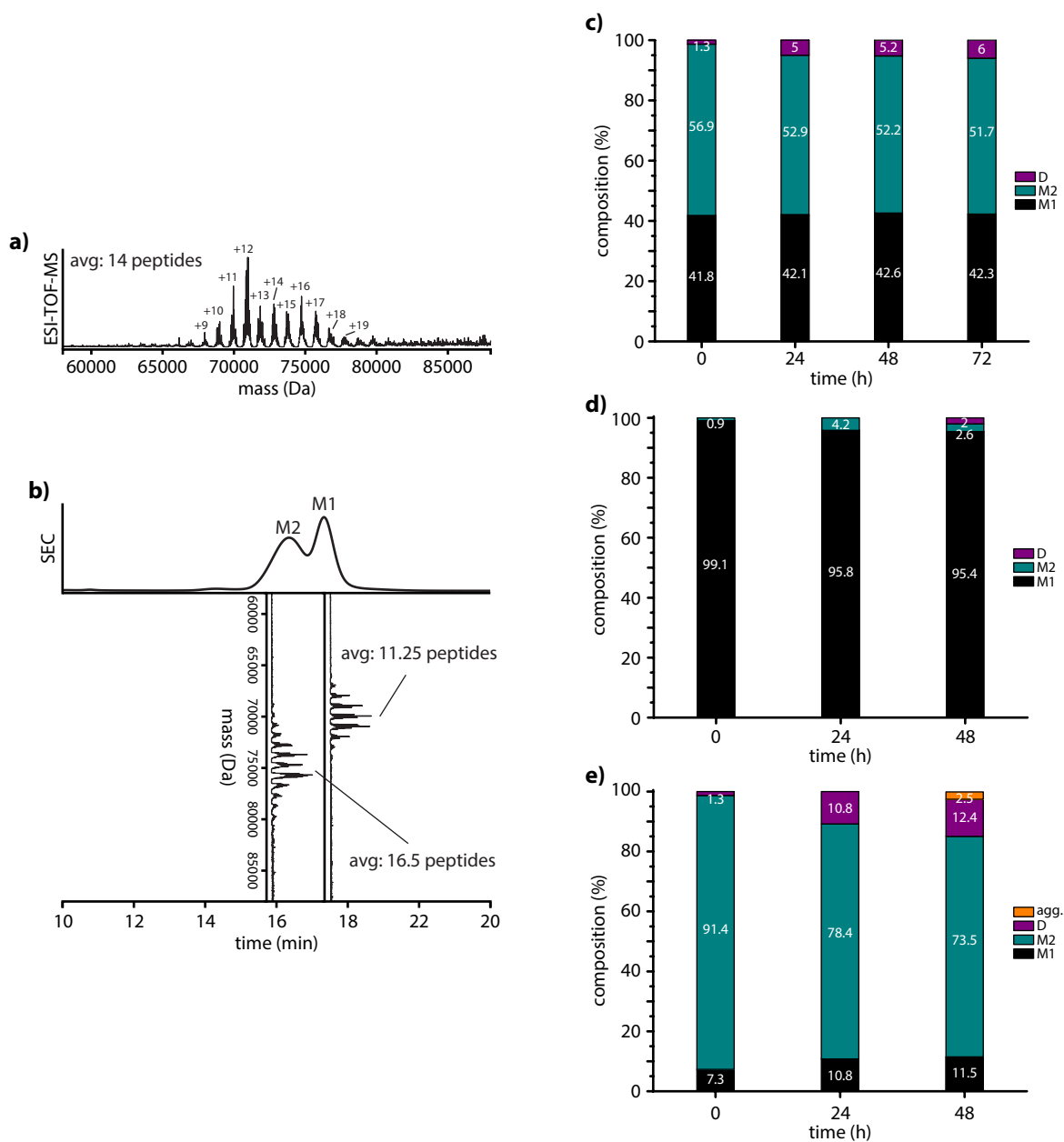
### 3.2 Conformational changes in CRM<sub>197</sub> conjugates

Diphtheria toxin (DT) is known to undergo significant conformational changes during cellular entry.<sup>15</sup> Additionally, DT and CRM<sub>197</sub> are known to exhibit 3D domain-swapping behavior.<sup>16</sup> In domain-swapping, the “closed” interface between the receptor-binding domain (R) and the translocation (T) and catalytic (C) domains of the DT/CRM<sub>197</sub> monomer is first disrupted to form an “open” monomer, wherein R is connected to T and C through a hinge-loop. Two open monomers can then interact to form a non-covalent dimer, wherein the R domain of each monomer forms a closed interface with the T and C domains of the other (**Figure 3.2**). The characteristic ability of CRM<sub>197</sub> to “open” suggests that the M1 and M2 populations observed in peptide-CRM<sub>197</sub> conjugates represent some form of closed and open monomers, respectively. CRM<sub>197</sub> glycoconjugates have been reported to take on more open conformations than unconjugated materi-



**Figure 3.2.** Domain swapping in CRM<sub>197</sub>. Disruption of the interface between the catalytic (C) and receptor-binding (R) domains of monomeric CRM<sub>197</sub> results in “open” form monomer. The open monomer can revert to the closed form or interact with another open monomer to form a noncovalent dimer, wherein the R domain of one monomer forms an interface with the catalytic domain (C’) of the other.





**Figure 3.3.** Stability and characterization of peptide-CRM<sub>197</sub> conjugates. A batch of peptide-CRM<sub>197</sub> was prepared via the standard activation-conjugation protocol and characterized by a) ESI-TOF-MS and b) SEC (monitored at 280 nm). Below the chromatogram are deconvoluted mass spectra for isolated M1 and M2. c) Unfractionated conjugate, d) isolated M1, and e) isolated M2 were stored at room temperature and the composition of each sample was monitored over time by SEC.

al.<sup>17-19</sup> However, distinct populations of open and closed monomers have not previously been observed through SEC characterization.<sup>18,20</sup>

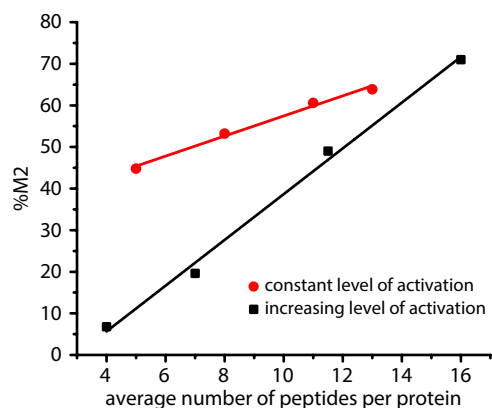
To gain a better understanding of the relationship between M1 and M2, stability studies were carried out on peptide-CRM<sub>197</sub> conjugates prepared via Pfizer's activation-conjugation protocol. The peptide (CTNEHFRG) used in this study, and for those throughout this chapter, was chosen for its similarity to a proprietary peptide epitope of interest to Pfizer. With this proprietary peptide epitope, Pfizer has determined that conjugates with an average of around 13 peptides per

CRM<sub>197</sub> monomer are best for their applications and, as such, conjugates with approximately this degree of modification were of greatest interest in this work.

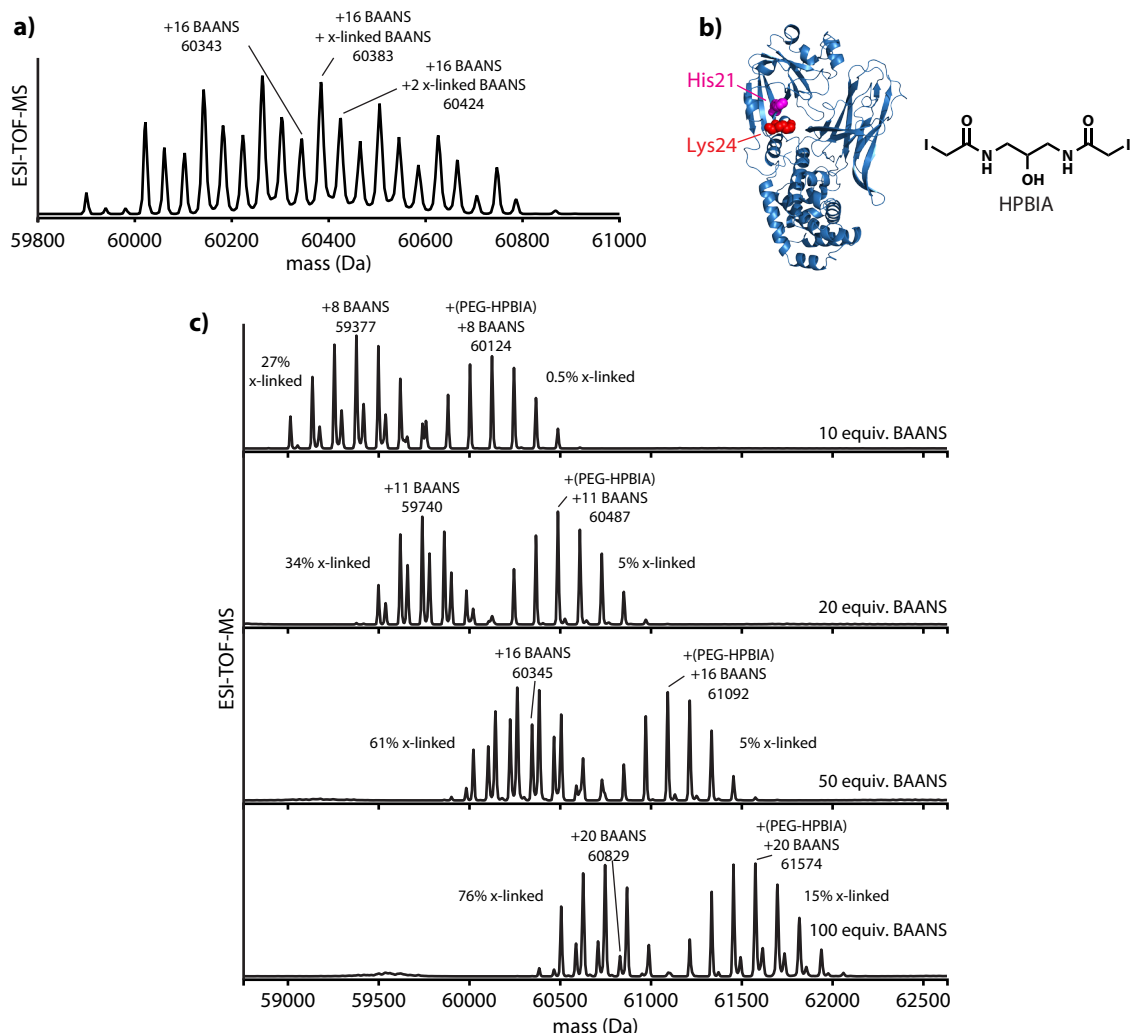
A batch of peptide-CRM<sub>197</sub> conjugate with an average of 14 peptides per protein was prepared and characterized by SEC and electrospray time-of-flight mass spectrometry (ESI-TOF-MS) (*Figure 3.3*). A portion of this batch was set aside and the remaining material was purified by SEC to obtain relatively pure samples of M1 and M2. The non-fractionated material, along with purified M1 and M2, were kept at room temperature and reanalyzed by SEC at 24 h intervals. While M1 was found to be quite stable under these conditions, M2 was found to convert slowly to dimer. However, little, if any, true interconversion between M1 and M2 was observed. Over time, the appearance of small amounts of M2 in isolated M1, and small amounts of M1 in isolated M2, were within the error of the peak fitting procedure used to determine composition. These findings suggested that M1 and M2 are not in equilibrium, but rather species trapped in an open or closed conformation. ESI-TOF-MS characterization supported this hypothesis, as isolated M2 was found to have increased peptide loading compared to M1 (*Figure 3.3b*).

### 3.2.1 Effects of the conjugation step on conformational change

The findings from the stability studies described above could suggest that after increasing the number of peptides per CRM<sub>197</sub> to a certain point, the protein undergoes a conformational change. However, multiple pieces of evidence suggest that this is not the case. First and foremost, the mass spectrum of isolated M1 and M2 partially overlap (*Figure 3.3b*). For example, the 14-peptide conjugate can be found as either M1 or M2. Second, decreasing peptide loading does not eliminate M2. For CRM<sub>197</sub> activated to a given level with BAANS, increasing peptide loading does lead to increased M2 formation (*Figure 3.4*, red curve). However, significant amounts of M2 is observed in batches with only a few peptides. Third, increasing peptide loading by increasing the level of BAANS-activation prior to conjugation has a greater effect on M2 formation than does increasing the degree of peptide loading on CRM<sub>197</sub> activated to a consistent degree (*Figure 3.4b*, black curve). Lastly, treatment of BAANS-activated CRM<sub>197</sub> with a thiol-containing discrete polyethylene glycol (m-dPEG<sub>12</sub>-SH) resulted in formation of both M1 and M2, much like the peptide conjugates. However, direct conjugation of a discrete PEG-NHS ester (m-dPEG<sub>12</sub>-NHS) to lysine residues in CRM<sub>197</sub> did not result in M2 formation. Taken together, this evidence suggests that peptide loading, and the conjugation step in general, is not the root cause of the irreversible conformational change. Rather, BAANS activation results in, or promotes, the conformational change. Clearly, a closer look at the activation chemistry was needed.



**Figure 3.4.** Effect of peptide loading on M2 formation in CRM<sub>197</sub> conjugates. Variable levels of peptide loading were achieved in two ways: 1) CRM<sub>197</sub> was activated with BAANS to an average of 16 activating agents per protein, then treated with increasing amounts of peptide (0.1875–1.5 g peptide per g protein) (red curve). 2) CRM<sub>197</sub> was activated with increasing amounts of BAANS (10–100 equiv.), then treated with a constant amount of peptide (1.5 g peptide per g protein) (black curve). Composition was determined by SEC.



**Figure 3.5.** BAANS-activated CRM<sub>197</sub> has intraprotein crosslinks. a) Deconvoluted mass spectrum demonstrating significant amounts of crosslinking in a sample of BAANS-activated CRM<sub>197</sub>. b) A significant portion of the observed crosslinking was believed to occur between His21 and nearby lysine residues, particularly Lys24. Using HPBIA, His21 was selectively capped to roughly 50% modification and further capped with m-dPEG<sub>12</sub>-SH. The resulting batch of 1:1 CRM<sub>197</sub>:(PEG-HPBIA)-CRM<sub>197</sub> was then treated with increasing amounts of BAANS. c) Mass spectra of the activated material clearly showed significantly lower amounts of crosslinking in material that had been treated with HPBIA, suggesting that His21 was the primary target of BAANS crosslinking.

### 3.2.2 Effects of the activation step on conformational change

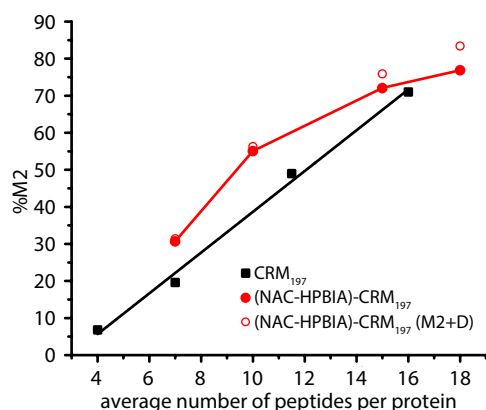
While the bromoacetyl functionality of the BAANS crosslinking agent is generally selective towards cysteines (or other thiol-containing molecules), cross-reactivity with other nucleophilic residues is not unprecedented.<sup>21</sup> In the case of BAANS activation of CRM<sub>197</sub>, off-target reactivity of the bromoacetyl group of BAANS is observed in the form of crosslinking. Mass spectra of BAANS-CRM<sub>197</sub> reveal the expected series of +121 Da peaks corresponding to lysine bromoacetylation, but also reveal +40 Da peaks, consistent with addition of BAANS followed by loss of bromide due to crosslinking (**Figure 3.5a**). The degree and number of crosslinks formed increases with increasing levels of BAANS activation. As CRM<sub>197</sub> does not contain any free cysteine residues, the crosslinking

must occur between lysines and other non-cysteine nucleophiles.

In early work towards the preparation of homogeneous CRM<sub>197</sub> conjugates, Chang *et al.* discovered that His21 displayed pronounced reactivity towards iodoacetamide-containing reagents.<sup>22</sup> Furthermore, Chang *et al.* demonstrated that specific crosslinking between His21 and Lys24 was possible using the bis(iodoacetamide) reagent HPBIA (**Figure 3.5b**). Based on this finding, BAANS crosslinking between His21 and Lys24 might be expected.

To examine whether BAANS activation results in a specific crosslink with His21 (versus other nucleophilic residues), His21 was first labeled to roughly 50% completion with HPBIA. The HPBIA-CRM<sub>197</sub> was then capped with m-dPEG12-SH and activated with varying amounts of BAANS. The resulting mixture contained two populations of activated CRM<sub>197</sub>, which were distinct by ESI-TOF-MS (**Figure 3.5c**). Comparison of the two populations revealed significantly greater crosslinking in the portion of the sample where His21 was not capped, relative to that in which it was capped with PEG-HPBIA. Moreover, this experiment suggested that 80–98% of crosslinking observed in samples of activated CRM<sub>197</sub> occurs at His21, presumably with Lys24.

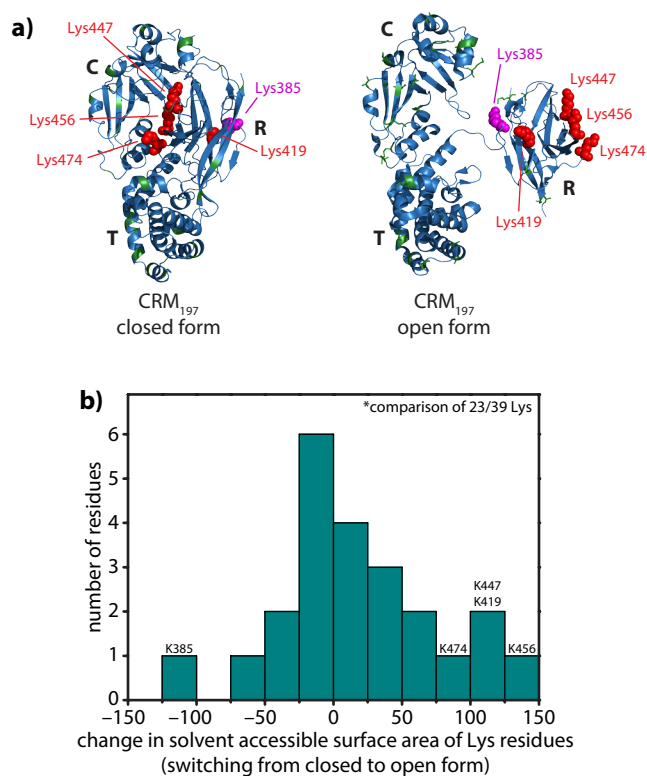
While this reactivity was an interesting finding in its own right, we thought that this specific crosslink might play a role in the formation of the open form conjugate (M2). His21 is located on the catalytic domain, close to the C-R domain interface. Perhaps modification of His21 through crosslinking could disrupt the C-R interface, partially opening the protein through a “bumping” effect. To probe this effect, Lys24-His21 crosslinking was mimicked through complete modification of His21 with HPBIA, followed by NAC capping. Peptide mapping of (NAC-HPBIA)-CRM<sub>197</sub> confirmed the high specificity of this reagent for His21 (not shown). (NAC-HPBIA)-CRM<sub>197</sub> was then activated with varying levels of BAANS, and subsequently conjugated with the peptide of interest (CTNEHFRG). Comparison by SEC of this series of conjugates prepared from (NAC-HPBIA)-CRM<sub>197</sub> to an analogous series prepared from standard CRM<sub>197</sub> revealed that modification of His21 has a distinct effect on M2 formation (**Figure 3.6**).



**Figure 3.6.** Effect of capping His21 on M2 formation. Samples of unmodified CRM<sub>197</sub> (black curve) and (NAC-HPBIA)-CRM<sub>197</sub> (red curve) were activated with BAANS (10–100 equiv.) and then capped with peptide. Analysis by SEC revealed roughly a 10% increase of M2 in His21-capped samples, controlling for peptide loading. The leveling seen in %M2 for His21-capped material was believed to be a result of M2 conversion into D (open red circles) rather than a true leveling effect.

Selective capping of His21 prior to activation resulted in nearly a 10% increase in M2 formation over standard material with equivalent levels of peptide loading. However, as complete crosslinking is not observed in standard BAANS-activated CRM<sub>197</sub>, the actual contribution of M2 formation due to His21 crosslinking is likely smaller than observed for the (NAC-HPBIA)-CRM<sub>197</sub> system. Efforts are currently underway to prepare His21 knockout CRM<sub>197</sub> to elucidate the role of His21 crosslinking on M2 further.

While His21 crosslinking appears to be M2-promoting, it is not independently sufficient to cause M2 formation. As mentioned above, PEGylated CRM<sub>197</sub> prepared by lysine modification with m-dPEG<sub>12</sub>-NHS does not result in observable M2 formation. If His21 crosslinking is the direct cause of M2 formation, (NAC-HPBIA)-CRM<sub>197</sub> should form M2 after direct PEGylation. However, PEGylation of (NAC-HPBIA)-CRM<sub>197</sub> did not result in M2 formation, as determined by SEC. Moreover, HPBIA capping after PEGylation of CRM<sub>197</sub>



**Figure 3.7.** Lysines at the interface of the catalytic (C) and receptor binding (R) domains of CRM<sub>197</sub> are unique. a) K419, K447, K456, and K474 become exposed upon transition from the closed form to the open form CRM<sub>197</sub>. b) Differences in solvent accessible surface area (SASA) of lysine residues between the two conformations of CRM<sub>197</sub> were quantified with BioLuminate. Lysine residues that were incomplete or missing from either one or both crystal structures were not included in this analysis.

C-R interface is partially disrupted, by His21 crosslinking or other factors, allowing for activation of the interfacial lysine residues. Peptide mapping experiments of isolated M1 and M2 are currently underway to determine if activation of these interfacial lysine residues is, indeed, the key difference between the two monomeric conjugate populations. It should be noted that previous studies probing CRM<sub>197</sub> lysine reactivity through peptide mapping have not shown these interfacial lysine residues to be reactive.<sup>19,23</sup> However, the conjugates examined in those studies were prepared by direct conjugation methods rather than through an activation-conjugation strategy like that used in the present work.

If the “bump and label” hypothesis is correct, activation of the lysine residues at the C-R interface, through any means possible, should result in M2 formation. A brute-force strategy was employed to activate these interfacial lysine residues. The most solvent accessible and reactive lysines on CRM<sub>197</sub> were acetylated to varying degrees with acetic acid NHS ester. These batches were then reacted with varying amounts of BAANS to produce activated Ac-CRM<sub>197</sub>, which were then conjugated to the peptide of interest. By pre-capping the most reactive lysine residues, the probability of activating a less accessible lysine, such as those in or near the interface, with BAANS increases. The

does not result in M2 formation.

Given that His21 crosslinking promoted but did not directly result in M2 formation, the “C-R domain interface bumping” hypothesis was clearly missing an important factor. Close examination of the structures for “closed” (DT monomer PDB ID: 1MDT) and “open” (isolated monomer from CRM<sub>197</sub> dimer PDB ID: 4AE0) forms of CRM<sub>197</sub> led to a “bump and label” hypothesis (Figure 3.7a). A comparison of the differences between the solvent accessible surface area (SASA) of the lysine residues in the closed- and open-form monomers revealed that most lysines remained similarly accessible (Figure 3.7b). However, residues K419, K447, K456, and K474 become significantly more solvent accessible in the open conformation. Unsurprisingly, these four residues reside at the C-R interface in the closed monomer. This collection of lysines is roughly the same in number as the difference in average peptide loading between M1 and M2 in the isolation experiment discussed as the beginning of this section, perhaps not coincidentally. BAANS activation of this set of interfacial lysine residues may be the necessary factor for M2 formation after conjugation. If this is the case, M2 formation can be explained by a “bump and label” mechanism, in which the

interfacial lysines may become more accessible due to a bumping effect afforded by specific sites of acetylation or, due to the decreased labeling competition, BAANS is more readily able to modify these lysines during conformational fluctuations. Regardless of the mechanism, SEC analysis of the resulting conjugates revealed a strong correlation between the level of pre-acetylation and M2 formation for any given level of peptide loading (*Figure 3.8a*). Combining pre-acetylation with HPBIA capping of His21 results in near complete M2 formation (*Figure 3.8b*).

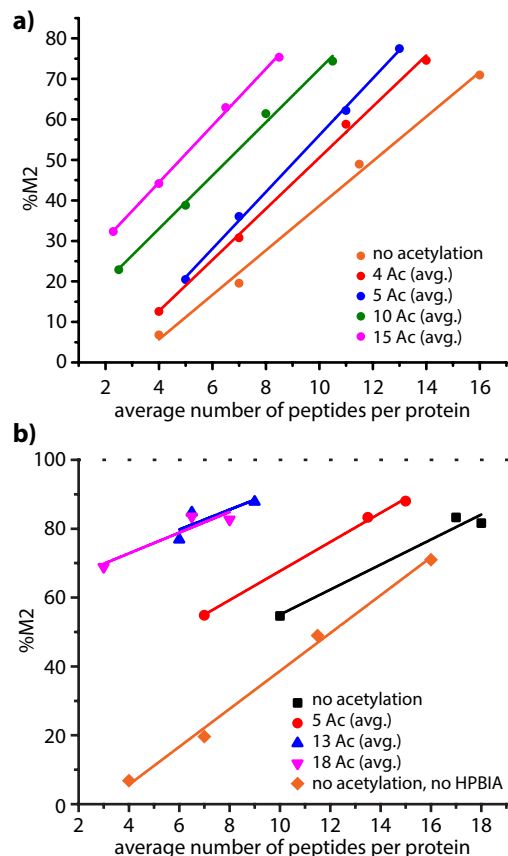
### 3.3 Alternative activation-conjugation strategies

As intraprotein crosslinking was found to be a contributing factor to conformational change in peptide-CRM<sub>197</sub> conjugates, alternative chemistries were explored as a means to create more homogeneous CRM<sub>197</sub> conjugates. In this section, four alternative activation-conjugation strategies will be discussed.

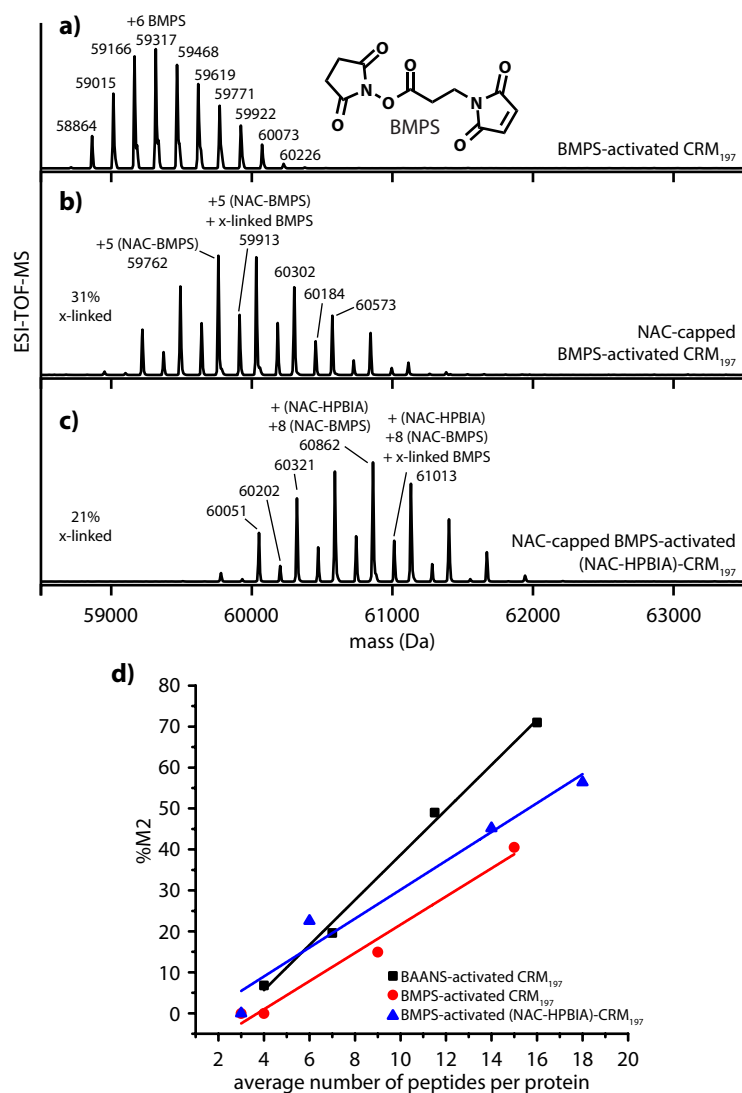
#### 3.3.1 Maleimide-NHS ester crosslinking

Like haloacetamides, maleimides are capable of labeling residues with nucleophilic side chains other than cysteine.<sup>21</sup> However, a maleimide-NHS ester crosslinking reagent, such as BMPS, may or may not form the Lys24-His21 M2-promoting crosslink found to occur upon BAANS activation. Unlike crosslinking with BAANS, BMPS crosslinking is undetectable by mass spectrometry of intact samples, as a Michael addition between a nucleophilic residue and the maleimide functionality results in no change of mass. However, capping the BMPS-activated material with excess thiol-containing small molecule, such as NAC, can reveal crosslinks by comparison of reactive versus unreactive activating agent. That is, any activated species observed by mass spectrometry after NAC-capping that do not show addition of NAC can be assumed to be crosslinked. NAC-capped BMPS-activated CRM<sub>197</sub> was found to be roughly 30% crosslinked when activated to an average of 9.5 activating agents per protein (*Figure 3.9b*). Comparable material prepared from (NAC-HPBIA)-CRM<sub>197</sub> was found to be only about 20% crosslinked, with an average of 7.5 activating agents per protein (*Figure 3.9c*). This difference suggests that, while His21 is involved in crosslinks in BMPS-activated CRM<sub>197</sub>, it is not the primary site of crosslinking in this system.

Peptide conjugates from BMPS-activated CRM<sub>197</sub> were prepared and analyzed by SEC. Much like conjugates prepared from BAANS-activated CRM<sub>197</sub>, increasing activation and pep-



**Figure 3.8.** Effect of lysine pre-capping on M2 formation. a) Peptide conjugates were prepared from BAANS-activated CRM<sub>197</sub> (orange curve) or BAANS-activated Ac<sub>n</sub>-CRM<sub>197</sub>. By SEC, increasing levels of lysine pre-acetylation resulted in higher levels of M2 formation, controlling for peptide loading. b) The effect was found to be considerably stronger if samples were capped with (NAC-HPBIA) prior to acetylation, and conjugation. Standard BAANS activation-conjugation protocols were used to prepare the non-capped control (orange curve).



**Figure 3.9.** Crosslinking in BMPS-activated CRM<sub>197</sub> and comparison of peptide conjugates prepared from BAANS-activated and BMPS-activated CRM<sub>197</sub>. a) Crosslinking in BMPS-activated CRM<sub>197</sub> cannot be observed directly as the Michael addition of a nucleophilic residue to the maleimide of BMPS results in no change of mass. b) Crosslinking can be observed indirectly by coupling a small molecule thiol to the activated material. Unreacted activating groups are assumed to have formed crosslinks. c) BMPS-activation of (NAC-HPBIA)-CRM<sub>197</sub> results in lower levels of crosslinking. However, crosslinking must occur at sites other than His21 in BMPS-activated CRM<sub>197</sub>. d) Conjugates prepared from BMPS-activated CRM<sub>197</sub> (red curve) resulted in lower levels of M2 formation than did comparable conjugates from BAANS-activated CRM<sub>197</sub> (black curve). BMPS-activated (NAC-HPBIA)-CRM<sub>197</sub> (blue curve) was found to contain roughly 10% more M2 than did comparable material without His21-capping.

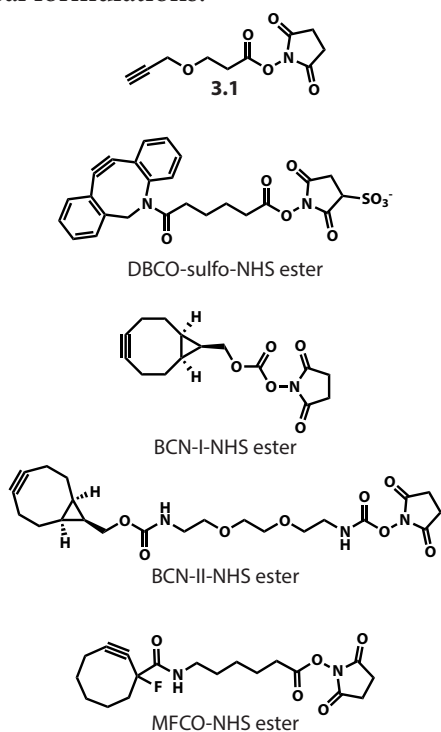
tide loading resulted in increased M2 formation (**Figure 3.9d**). While BMPS was demonstrated to form intraprotein crosslinks to a similar extent as BAANS, conjugates prepared from BMPS-activated CRM<sub>197</sub> did not result in as much M2 formation as those prepared from BAANS-activated CRM<sub>197</sub> on a per peptide basis. However, like conjugates prepared from BAANS-activated CRM<sub>197</sub> pre-capped with NAC-HPBIA, conjugates prepared from BMPS-activated CRM<sub>197</sub> pre-capped with NAC-HPBIA demonstrated roughly a 10% increase in M2 content relative to conjugates that were not capped with HPBIA.

The lower propensity of peptide-CRM<sub>197</sub> conjugates prepared through BMPS activation towards M2 formation further supports the bump and label hypothesis discussed above. This hypothesis requires that the activating agent or labeling agent is able to access the C-R interfacial lysine residues after bumping or during natural conformational fluctuations. While the small BAANS agent, with its zero-length linker, is able to access the interfacial lysine residues fairly readily, it was demonstrated that the relatively bulky m-dPEG<sub>12</sub>-NHS was unable to promote M2 formation. BMPS offers an intermediate case, in which the maleimide moiety minimizes reactivity towards interfacial lysine acylation through steric hindrance. Efforts are currently ongoing to demonstrate this effect, systematically, with a series of maleimide-NHS ester and bromoacetyl-NHS ester activating agents with varying linker lengths and steric bulkiness. It is ex-

pected that agents with bulky linkers (e.g., branched alkyl groups or a cyclohexyl linker) will result in minimal M2 formation after conjugation while still allowing for the desired levels of peptide loading.

### 3.3.2 CuAAC and SPAAC

To avoid intraprotein crosslinking at native residues completely, attempts were made to utilize copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) and strain-promoted alkyne-azide cycloaddition (SPAAC) strategies for conjugating peptides of interest to the CRM<sub>197</sub> scaffold (**Figure 3.10**). An azidolysine variant of the peptide of interest ( $\{K(N_3)\}$ TNEHFRG) was conjugated to CRM<sub>197</sub>, which was first activated with propargyl-NHS ester reagent **3.1**, using standard CuAAC conditions. While the resulting conjugates contained significantly less M2 than corresponding conjugates prepared from BAANS-CRM<sub>197</sub>, these conjugates were prone to forming dimers and higher order aggregates. Further studies using CuAAC to prepare CRM<sub>197</sub> conjugates were not undertaken, primarily due to concerns surrounding complete removal of toxic Cu(I) salts from pharmaceutical formulations.



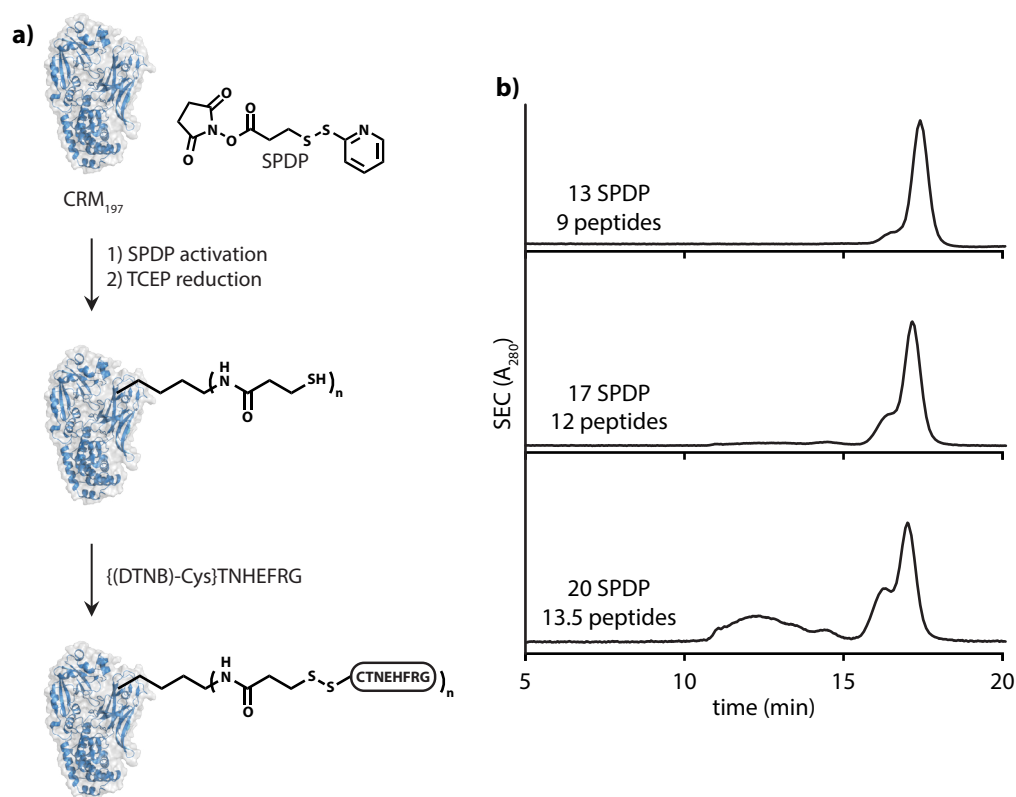
**Figure 3.10.** Reagents for activating CRM<sub>197</sub> prior to CuAAC or SPAAC conjugation with azidolysine-containing peptide  $\{K(N_3)\}$ TNEHFRG.

As an alternative to CuAAC for preparing crosslink-free CRM<sub>197</sub> peptide conjugates, activation of CRM<sub>197</sub> was attempted with four commercially available lysine reactive SPAAC reagents. Activation with a dibenzocyclooctyne reagent, DBCO-sulfo-NHS ester, resulted in significant aggregation (>50%) at levels of activation as low as four modifications, on average, per protein. Strained cyclooctyne derivatives BCN-I-NHS ester and BCN-II-NHS ester, along with the monofluoro cyclooctyne derivative MFCO-NHS ester, were then screened. Only low levels of activation ( $\leq 8$  activators per protein on average) were achieved, and even lower levels of peptide ( $\{K(N_3)\}$ TNEHFRG) incorporation were attained ( $\leq 4$  peptides per protein on average). At these low levels of modification, no M2 was observed. However, this would be the case for conjugates prepared through BAANS or BMPS activation as well. Improvement of the activation-conjugation strategy using SPAAC would certainly be possible through the use of bespoke activating agents optimized for this system. However, the resulting conjugate materials may not be appropriate for use as conjugate vaccines. It seems quite likely that the cycloaddition adduct would be strongly antigenic.<sup>12</sup>

### 3.3.3 Disulfide exchange

The final strategy investigated for crosslink-free preparation of peptide protein conjugates relied on the use of the mixed disulfide-NHS ester reagent SPDP (succinimidyl 3-(2-pyridyldithio)





**Figure 3.11.** Preparation of peptide-CRM<sub>197</sub> conjugates through disulfide exchange. a) A three-step process was used to prepare peptide-CRM<sub>197</sub> conjugates through disulfide exchange. CRM<sub>197</sub> was first acylated with SPDP and then activated by disulfide reduction with TCEP. Peptide, pre-labeled with DTNB, was then conjugated to the exposed thiols on the surface of CRM<sub>197</sub>. Conjugates were prepared from CRM<sub>197</sub> labeled with varying amounts of SPDP. b) Characterization by SEC showed that the resulting conjugates contained levels of M2 similar to those found for conjugates prepared from BMPS-activated CRM<sub>197</sub>. However, significant aggregation was observed in conjugates prepared from CRM<sub>197</sub> activated to high levels with SPDP.

propionate) for CRM<sub>197</sub> activation. Disulfide exchange, while not as bioorthogonal as CuAAC and SPAAC, is more chemoselective than maleimide or haloacetamide chemistry. Additionally, the protein-peptide linker would not be expected to be as antigenic as that resulting from CuAAC and SPAAC conjugation strategies.

The preparation of peptide-CRM<sub>197</sub> conjugates using disulfide exchange required the activation of both CRM<sub>197</sub> and the peptide (**Figure 3.11a**). CRM<sub>197</sub> was first reacted with SPDP. The introduced disulfides in the SPDP-CRM<sub>197</sub> material were then reduced by treatment with tris(2-carboxyethyl)phosphine hydrochloride (TCEP). It should be noted that the native disulfides in unmodified CRM<sub>197</sub> did not reduce under these conditions, as determined by Ellman's test.<sup>24</sup> The cleaved SPDP-CRM<sub>197</sub> was then treated with the peptide of interest, which was first activated with Ellman's reagent (DNTB) and purified. This conjugation procedure produced peptide-CRM<sub>197</sub> conjugates with levels of M2 similar to those observed for BMPS-activated CRM<sub>197</sub> conjugates. This was not unexpected based on the similarities in steric bulk between the two activating agents. However, at higher levels of activation these conjugates tended towards aggregation, potentially due to interprotein crosslink formation as a result of disulfide exchange (**Figure 3.11b**). Ultimately, while avoiding M2-promoting His21 crosslinking, disulfide exchange was found not to be a superior al-

ternative to haloacetamide or maleimide chemistry due to issues of aggregation.

### 3.4 Conclusions

These studies have laid the groundwork for understanding conformational change in peptide-CRM<sub>197</sub> conjugate vaccines. It was determined that crosslinking, or direct modification, of His21 in CRM<sub>197</sub> promotes conformational change when using an activation-conjugation strategy to prepare peptide-CRM<sub>197</sub> conjugates. Furthermore, it was hypothesized that activation of the lysine residues at the interface of the catalytic and receptor-binding domains is necessary and sufficient for M2 formation. However, confirmation of this theory awaits the results of ongoing peptide mapping studies.

In an effort to prepare homogeneous peptide-CRM<sub>197</sub> conjugates, alternative activation-conjugation strategies were investigated. While a strategy that resulted in material free of conformational inhomogeneity was not established, it was determined the steric bulk of the activating agent plays a significant role in M2 formation. Ongoing studies are further probing this effect. Additionally, we are currently engineering mutants of CRM<sub>197</sub> that can be site-selectively tethered to produce either stabilized monomers, through stapling of the C and R domains, or stabilized covalent dimers of CRM<sub>197</sub>, which could then be used to create conformationally-locked conjugate vaccines. Ultimately, these well-defined constructs can be used to gain insights into structure-immunogenicity relationships in peptide-CRM<sub>197</sub> conjugate vaccines.

### 3.5 Materials and methods

#### 3.5.1 General procedures and materials

Unless otherwise noted, all chemicals were purchased from commercial sources and used as received without further purification. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F<sub>254</sub> plates and visualized by ultraviolet (UV) irradiation at 254 nm and staining with potassium permanganate. Purifications by flash silica gel chromatography were performed using EM silica gel 60 (230–400 mesh). Chromatography solvents were used without distillation. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H<sub>2</sub>O) used in all procedures was deionized using a NANOpure™ purification system (Barnstead, USA). Centrifugations were performed with an Eppendorf 5424 R at 4 °C (Eppendorf, Hauppauge, NY). CRM<sub>197</sub> from *Corynebacterium diphtheriae* and standard BAANS-activated CRM<sub>197</sub> were obtained from Pfizer (St. Louis, MO). All samples of CRM<sub>197</sub>, activated CRM<sub>197</sub>, and CRM<sub>197</sub> conjugates were handled and stored at or below 4 °C. All containers (e.g., Eppendorf tubes, spin columns, spin filters, and LC-MS vials) were chilled on ice prior to addition of samples containing CRM<sub>197</sub> or its derivatives. Peptides were procured from GenScript (Piscataway, NJ). Discrete PEG reagents were purchased from Quanta Biodesign (Plain City, OH). Cyclooctyne derivatives were purchased from Berry & Associates (Dexter, MI).

#### 3.5.2 Instrumentation and sample analysis

NMR. <sup>1</sup>H and <sup>13</sup>C spectra were measured with a Bruker AVQ-400 (400 MHz) spectrometer. <sup>1</sup>H

NMR chemical shifts are reported as  $\delta$  in units of parts per million (ppm) relative to residual  $\text{CHCl}_3$  ( $\delta$  7.26, singlet) or DMSO ( $\delta$  2.50, pentet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (quintet), or br s (broad singlet). Coupling constants are reported as a  $J$  value in Hertz (Hz). The number of protons ( $n$ ) for a given resonance is indicated as  $n\text{H}$  and is based on spectral integration values.  $^{13}\text{C}$  NMR chemical shifts are reported as  $\delta$  in units of parts per million (ppm) relative to  $\text{CDCl}_3$  ( $\delta$  77.16, triplet) or DMSO- $d_6$  ( $\delta$  39.52, septet).

**Mass Spectrometry.** Proteins and protein conjugates were analyzed on an Agilent 6224 Time-of-Flight (TOF) mass spectrometer with a dual electrospray source connected in-line with an Agilent 1200 series HPLC (Agilent Technologies, USA). Chromatography was performed using a Proswift RP-4H (Thermo Scientific, USA) column with a  $\text{H}_2\text{O}/\text{MeCN}$  gradient mobile phase containing 0.1% formic acid. Mass spectra of proteins and protein conjugates were deconvoluted with MassHunter Qualitative Analysis Suite B.05 (Agilent Technologies, USA).

**High Performance Liquid Chromatography (HPLC).** HPLC was performed on Agilent 1100 series HPLC systems (Agilent Technologies, USA) equipped with in-line diode array detector (DAD) and fluorescence detector (FLD). Size exclusion chromatography (SEC) was accomplished on a TSKgel G3000SW<sub>XL</sub> column fitted with a TSKgel SW<sub>XL</sub> guard column (Tosoh Bioscience LLC, King of Prussia, PA) using an aqueous mobile phase (100 mM sodium phosphate, 200 mM NaCl, pH 7.0) at a flow rate of 0.55 mL/min. Column integrity was confirmed by analyzing bovine serum albumin (Sigma-Aldrich, St. Louis, MO) and CRM<sub>197</sub> (Pfizer, St. Louis, MO) analytical standards, and a 1,350–670,000 Da gel filtration standard mixture (Bio-Rad, Hercules, CA). To integrate partially overlapping SEC peaks accurately, multiple Gaussian fits were performed using OriginPro 9.0 (OriginLab Corp., Northampton, MA).

### 3.5.3 Experimental procedures

**General procedure for preparing CRM<sub>197</sub> conjugates from BAANS-CRM<sub>197</sub>.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu\text{M}$  final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of BAANS (generally 10, 20, 50, or 100 equiv. from a 200 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant BAANS-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

BAANS-CRM<sub>197</sub> in reaction buffer (33  $\mu\text{M}$ ) was treated with stock peptide solution (1.5 mg peptide/mg protein {15  $\mu\text{L}$  peptide stock solution per 100  $\mu\text{L}$  of 33  $\mu\text{M}$  BAANS-CRM<sub>197</sub>}, 20 mg/mL stock concentration, 0.6 M  $\text{NaHCO}_3$ , pH 9.2), or neat m-dPEG<sub>12</sub>-SH followed by addition of 0.6 M  $\text{NaHCO}_3$ , pH 9.2. The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 3 h. Unreacted bromoacetamide groups were capped by addition of neat *N*-acetylcysteamine (0.15  $\mu\text{L}$  per 100  $\mu\text{L}$  of 33  $\mu\text{M}$  BAANS-CRM<sub>197</sub>). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant BAANS-CRM<sub>197</sub>-peptide conjugate was analyzed by ESI-TOF-MS and SEC;

the sample was kept on ice until  $\leq 2$  min prior to injection.

**General procedure for direct PEGylation of CRM<sub>197</sub>.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a stock solution of m-dPEG<sub>12</sub>-NHS ester (25–100 equiv. from a 250 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant PEGylated CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

**General procedure for preparing CRM<sub>197</sub> conjugates from BMPS-CRM<sub>197</sub>.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of BMPS (generally 10, 20, 50, or 100 equiv. from a 200 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant BMPS-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

BMPS-CRM<sub>197</sub> in reaction buffer (33  $\mu$ M) was treated with stock peptide solution (1.5 mg peptide/mg protein {15  $\mu$ L peptide stock solution per 100  $\mu$ L of 33  $\mu$ M BMPS-CRM<sub>197</sub>}, 20 mg/mL stock concentration, 0.6 M NaHCO<sub>3</sub>, pH 9.2). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 3 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant BMPS-CRM<sub>197</sub>-peptide conjugate was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

**General procedure for the preparation of (NAC-HPBIA)-CRM<sub>197</sub>.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of HPBIA (200 mM in DMF, 137 equiv.). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. Additional HPBIA was added (137 equiv.), the solution was mixed by pipetting, and incubated for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. The resultant HPBIA-CRM<sub>197</sub> (33  $\mu$ M) was treated with neat *N*-acetylcysteamine (0.15  $\mu$ L per 100  $\mu$ L of 33  $\mu$ M HPBIA-CRM<sub>197</sub>) or m-dPEG<sub>12</sub>-SH. The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant (NAC-HPBIA)-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection. General procedures for activation and conjugation were used to further elaborate (NAC-HPBIA)-CRM<sub>197</sub>.

**General procedure for preparing Ac-CRM<sub>197</sub>.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of Ac-NHS (10, 20, 50, or 100 equiv. from a 200 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant Ac-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection. General procedures for activation and conjugation were used to further elaborate Ac-CRM<sub>197</sub>.

**General procedure for preparing CRM<sub>197</sub> conjugates through CuAAC.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of propargyl-NHS ester **3.1** (100 or 200 equiv. from a 100 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant **3.1**-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

**3.1**-CRM<sub>197</sub> in reaction buffer (33  $\mu$ M) was treated with stock azido-peptide solution (1.5 mg peptide/mg protein {15  $\mu$ L peptide stock solution per 100  $\mu$ L of 33  $\mu$ M **3.1**-CRM<sub>197</sub>}, 20 mg/mL stock concentration in reaction buffer), followed by addition of solutions of CuSO<sub>4</sub> (50  $\mu$ M final concentration), tris(3-hydroxypropyltriazolylmethyl)amine (250  $\mu$ M final concentration), and sodium ascorbate (2.5 mM final concentration). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 2 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant **3.1**-CRM<sub>197</sub>-peptide conjugate was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

**General procedure for preparing CRM<sub>197</sub> conjugates through SPAAC.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of cyclooctyne-NHS ester (10–100 equiv. from a 100 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant cyclooctyne-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

Cyclooctyne-CRM<sub>197</sub> in reaction buffer (33  $\mu$ M) was treated with stock azido-peptide solution (1.5 mg peptide/mg protein {15  $\mu$ L peptide stock solution per 100  $\mu$ L of 33  $\mu$ M cyclooctyne-CRM<sub>197</sub>}, 20 mg/mL stock concentration, in reaction buffer). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 6 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant cyclooctyne-CRM<sub>197</sub>-peptide conjugate was

analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

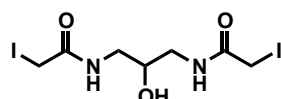
**General procedure for preparing CRM<sub>197</sub> conjugates through disulfide exchange.** An Eppendorf tube was pre-chilled on ice, charged with a stock solution of CRM<sub>197</sub> in storage buffer (33  $\mu$ M final concentration, 25 mM HEPES, 150 mM NaCl, 10% sucrose, pH 7.5), and diluted with ice-cold reaction buffer (DPBS, pH 8.0). To this solution was added a freshly prepared stock solution of SPDP (50–200 equiv. from a 100 mM stock in DMF). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 1.5 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant SPDP-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

SPDP-CRM<sub>197</sub> was reduced by addition of TCEP (100 equiv. relative to protein concentration, 0.5 M aqueous solution, pH 7.0). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 30 min. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the reduced SPDP-CRM<sub>197</sub> was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

Reduced SPDP-CRM<sub>197</sub> in reaction buffer (33  $\mu$ M) was treated with DTNP-labeled peptide (1.5 mg peptide/mg protein {15  $\mu$ L peptide stock solution per 100  $\mu$ L of 33  $\mu$ M reduced SPDP-CRM<sub>197</sub>}, 20 mg/mL stock concentration in reaction buffer, purified by HPLC). The resulting mixture was mixed thoroughly by gentle pipetting and incubated on ice for 3 h. The reaction mixture was then purified through five successive rounds of centrifugal filtration in 30 kDa MWCO filters with ice-cold reaction buffer. An aliquot of the resultant SPDP-CRM<sub>197</sub>-peptide conjugate was analyzed by ESI-TOF-MS and SEC; the sample was kept on ice until  $\leq 2$  min prior to injection.

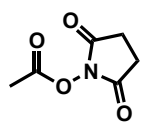
**Determination of solvent accessible surface area (SASA) of lysines in CRM<sub>197</sub>.** Total SASA of lysine residues was determined using BioLuminate (Schrödinger, LLC, USA). The crystal structure of monomeric DT (PDB ID: 1MDT) was analyzed as an analog of closed form CRM<sub>197</sub>. Total SASA of lysines in the open form of CRM<sub>197</sub> was determined by characterization of an isolated monomer from the crystal structure of CRM<sub>197</sub> dimer (PDB ID: 1AE0). Change in SASA was not calculated for lysine residues that were either incomplete or missing from one or both crystal structures.

### 3.5.4 Small molecule synthesis



#### Synthesis of *N,N'*-(2-hydroxypropane-1,3-diyl)bis(2-iodoacetamide) (HPBIA).

To a solution of 1,3-diamino-2-propanol (0.47 g, 5.22 mmol, 1 equiv.) in 50% H<sub>2</sub>O/MeOH (15 mL) was added 4-nitrophenyl iodoacetate (3.08 g, 10.03 mmol, 1.92 equiv.) portionwise at room temperature under N<sub>2</sub>. After 80 min, MeCN was added and the resulting yellow precipitate was collected by filtration and washed with MeCN to afford the desired product as a white solid (960 mg, 43%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.22 (t, *J* = 5.7 Hz, 2H), 5.05 (d, *J* = 5.0 Hz, 1H), 3.66 (s, 4H), 3.58–3.48 (m, 1H), 3.09 (dt, *J* = 13.6, 5.4 Hz, 2H), 2.99 (dt, *J* = 13.1, 6.1 Hz, 2H). <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  167.94, 68.02, 43.37, 1.03.



**Synthesis of 2,5-dioxopyrrolidin-1-yl acetate (Ac-NHS ester).** To a solution of *N*-hydroxy-succinimide (1.0 g, 8.69 mmol, 1 equiv.) in dry THF (8 mL) was added Ac<sub>2</sub>O (0.93 mL, 9.82 mmol, 1.13 equiv.) followed by Et<sub>3</sub>N (1.22 mL, 8.78 mmol, 1.01 equiv.) at 0 °C under N<sub>2</sub>. After 1 h, a white precipitate had formed and the reaction mixture was diluted with hexanes (50 mL). The product was isolated as a white solid by filtration (1.23 g, 90%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.80 (s, 4H), 2.29 (s, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 169.03, 165.44, 25.36, 17.21.

### 3.6 References

- (1) Plotkin, S. A. Vaccines: Past, Present and Future. *Nat. Med.* **2005**, *11* (4s), S5–S11.
- (2) Greenwood, B. The Contribution of Vaccination to Global Health: Past, Present and Future. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **2014**, *369* (1645), 20130433.
- (3) Jones, L. H. Recent Advances in the Molecular Design of Synthetic Vaccines. *Nat. Chem.* **2015**, *7* (12), 952–960.
- (4) De Gregorio, E.; Rappuoli, R. From Empiricism to Rational Design: a Personal Perspective of the Evolution of Vaccine Development. *Nat. Rev. Immunol.* **2014**, *14* (7), 505–514.
- (5) Costantino, P.; Rappuoli, R.; Berti, F. The Design of Semi-Synthetic and Synthetic Glycoconjugate Vaccines. *Expert Opin. Drug Discov.* **2011**, *6* (10), 1045–1066.
- (6) Zhu, J.; Warren, J. D.; Danishefsky, S. J. Synthetic Carbohydrate-Based Anticancer Vaccines: the Memorial Sloan-Kettering Experience. *Expert Rev. Vaccines* **2014**, *8* (10), 1399–1413.
- (7) Fettelschoss, A.; Zabel, F.; Bachmann, M. F. Vaccination Against Alzheimer Disease: an Update on Future Strategies. *Hum. Vaccin. Immunother.* **2014**, *10* (4), 847–851.
- (8) Do, T. H.; Chen, Y.; Van T Nguyen; Phisitkul, S. Vaccines in the Management of Hypertension. *Expert Opin. Biol. Ther.* **2010**, *10* (7), 1077–1087.
- (9) Malito, E.; Bursulaya, B.; Chen, C.; Surdo, P. L.; Picchianti, M.; Balducci, E.; Biancucci, M.; Brock, A.; Berti, F.; Bottomley, M. J.; et al. Structural Basis for Lack of Toxicity of the Diphtheria Toxin Mutant CRM197. *Proc. Natl. Acad. Sci. USA* **2012**, *109* (14), 5229–5234.
- (10) Hu, Q.-Y.; Berti, F.; Adamo, R. Towards the Next Generation of Biomedicines by Site-Selective Conjugation. *Chem. Soc. Rev.* **2016**, *45*, 1691–1719.
- (11) Ferro, V. A.; Khan, M. A. H.; Earl, E. R.; Harvey, M. J. A.; Colston, A.; Stimson, W. H. Influence of Carrier Protein Conjugation Site and Terminal Modification of a GnRH-I Peptide Sequence in the Development of a Highly Specific Anti-Fertility Vaccine. Part I. *Am. J. Reprod. Immunol.* **2003**, *48* (6), 361–371.
- (12) Peeters, J. M.; Hazendonk, T. G.; Beuvery, E. C.; Tesser, G. I. Comparison of Four Bifunctional Reagents for Coupling Peptides to Proteins and the Effect of the Three Moieties on the Immunogenicity of the Conjugates. *J. Immunol. Methods* **1989**, *120* (1), 133–143.

- (13) Briand, J. P.; Muller, S.; Van Regenmortel, M. H. Synthetic Peptides as Antigens: Pitfalls of Conjugation Methods. *J. Immunol. Methods* **1985**, *78* (1), 59–69.
- (14) Champion, B. R.; Stead, D. R.; Wright, P. A.; Pfizer Vaccines LLC IGE CH3 Peptide Vaccine. US Patent Office April 30, 2015.
- (15) Kent, M. S.; Yim, H.; Murton, J. K.; Satija, S.; Majewski, J.; Kuzmenko, I. Oligomerization of Membrane-Bound Diphtheria Toxin (CRM197) Facilitates a Transition to the Open Form and Deep Insertion. *Biophys. J.* **2008**, *94* (6), 2115–2127.
- (16) Steere, B.; Eisenberg, D. Characterization of High-Order Diphtheria Toxin Oligomers. *Biochemistry* **2000**, *39* (51), 15901–15909.
- (17) Crane, D. T.; Bolgiano, B.; Jones, C. Comparison of the Diphtheria Mutant Toxin, CRM197, with a Haemophilus Influenzae Type-B Polysaccharide-CRM197 Conjugate by Optical Spectroscopy. *Eur. J. Biochem.* **1997**, *246* (2), 320–327.
- (18) Pecetta, S.; Surdo, L.; Tontini, M.; Proietti, D.; Zambonelli, C.; Bottomley, M. J.; Biagini, M.; Berti, F.; Costantino, P.; Romano, M. R.; et al. Carrier Priming with CRM 197 or Diphtheria Toxoid Has a Different Impact on the Immunogenicity of the Respective Glycoconjugates: Biophysical and Immunochemical Interpretation. *Vaccine* **2015**, *33* (2), 314–320.
- (19) Crotti, S.; Zhai, H.; Zhou, J.; Allan, M.; Proietti, D.; Pansegrau, W.; Hu, Q.-Y.; Berti, F.; Adamo, R. Defined Conjugation of Glycans to the Lysines of CRM<sub>197</sub> Guided by Their Reactivity Mapping. *ChemBioChem* **2014**, *15* (6), 836–843.
- (20) Tontini, M.; Berti, F.; Romano, M. R.; Proietti, D.; Zambonelli, C.; Bottomley, M. J.; De Gregorio, E.; Del Giudice, G.; Rappuoli, R.; Costantino, P.; et al. Comparison of CRM197, Diphtheria Toxoid and Tetanus Toxoid as Protein Carriers for Meningococcal Glycoconjugate Vaccines. *Vaccine* **2013**, *31* (42), 4827–4833.
- (21) Hermanson, G. T. *Bioconjugate Techniques*; Elsevier: Oxford, 2013.
- (22) Chang, J. Y.; Ramseier, U.; Hawthorne, T.; O'Reilly, T.; van Oostrum, J. Unique Chemical Reactivity of His-21 of CRM-197, a Mutated Diphtheria Toxin. *FEBS Lett.* **1998**, *427* (3), 362–366.
- (23) Möglinger, U.; Resemann, A.; Martin, C. E.; Parameswarappa, S.; Govindan, S.; Wamhoff, E.-C.; Broecker, F.; Suckau, D.; Pereira, C. L.; Anish, C.; et al. Cross Reactive Material 197 Glycoconjugate Vaccines Contain Privileged Conjugation Sites. *Sci. Rep.* **2016**, *6*, 20488.
- (24) Ellman, G. L. Tissue Sulfhydryl Groups. *Arch. Biochem. Biophys.* **1959**, *82* (1), 70–77.