DEVELOPING FLARE: A CHEMI-GENETIC MALEIMIDE-BASED APPROACH TO FLUOROGENIC PROTEIN LABELLING

KELVIN KA-WAI TSAO

A thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Doctorate of Philosophy degree in Chemical Biology

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Abstract

Studying protein expression, localization, and behaviour within cells is invaluable to understanding biological processes. Fluorescent protein labelling can be used to easily visualize these biomolecules. However, many strategies use large fluorescent proteins or enzyme tags that disrupt the native localization and conformation of proteins of interest.

FlARe is a minimalistic, chemi-genetic fluorogenic protein labelling method with two components, the FlARe probe and the dC10 tag. The probe consists of a fluorophore linked to two maleimides, which quench the fluorophore's emission. They also react with the complementary dC10 peptide tag that is genetically fused to a specific protein and contains two cysteines positioned 10 Å apart. Reaction with dC10 anchors the FlARe probe to the tagged protein and abolishes the maleimides' quenching mechanisms. This releases the probe's intrinsic fluorescence, allowing the protein to be visualized by fluorescence.

This thesis sets out to complete three objectives: 1) To improve the FlARe probe chemical structure; 2) To implement novel developments of the FlARe protein labelling technique into live mammalian cells; 3) To expand the scope of the FlARe method to include bioconjugation.

To make the current FlARe probe design more stable in the aqueous milieu of the cell, a new FlARe probe, **KT12**, was synthesized. This new probe was more stable than the current **YC20** FlARe probe but not more selective. Despite this, valuable insight was made into the mechanisms of maleimide hydrolysis and thiol addition. After establishing that the current methoxymaleimide FlARe probe offered the most selective design, we invested our efforts into implementing the FlARe labelling technique in live mammalian cells. To this end, a

recombinant protein carrying different variants of the dC10 tag was successfully labelled with FlARe probes emitting at different wavelengths. Finally, we blended our expertise in probe design and biological implementation to develop a novel fluorogenic linker to form bioconjugates. As a result, a FlARe linker functionalized with both the dimaleimide moiety and an alkyne linker was designed. The successful synthesis and proof-of-principle demonstration of this novel FlARe linker were performed, allowing for the modular formation of fluorogenic bioconjugates.

These developments to FlARe advance this minimalistic and fluorogenic protein imaging method. This improves the current capability to fluorescently label proteins in cells in a non-disruptive manner. Thus, complex questions surrounding cell biology can be answered to better understand biological processes.

Acknowledgements

To Jeff, without whom none of this would have been possible. Thank you for taking me when I was just a second year McMaster Undergraduate. Thank you for being such a kind, caring and supportive mentor. For so long, tough love was the only mentorship I had known. Thank you for showing me a kinder mentorship, for being a friend, and for accepting me for who I am, sass, dark humour and all. I'll forever be grateful for all the lessons you have taught me, and all the soapbox talks you have given. You are the kind of scientist and teacher I strive to be…even if it means losing a little hair.

To the Keillor group members that I had the opportunity to share my time with, thank you for all the conversations, laughs, and drinks we shared. Yingche, you set my foundation as a chemist and have continued to help me even after you left the Keillors. Nicole, I'm so glad you crossed that tape boundary I set up when you first came. You have helped me so much with research as a post-doc and so much in life as a friend (with a shoulder). Thank you Mirka for helping me get started in everything biology. Thank you, Graham, for always coming to my rescue during synthesis (and for that Swern Oxidation protocol I never used), and for being my explosion buddy. Thank you to the Boddy group for always helping me with everything biochemistry. Thank you to the Pezacki group for helping me and pointing me in the right direction regarding the bioconjugation work. Kim, Mirka, Samuel, Amina, Abdullah, Karine, Jess, Nicole, Yingche, Eric, and Élise – thank you for adding so much colour and comedy during my graduate journey. Eric, thank you for keeping me company at the end of my studies, I'm a bit sad we didn't overlap more, but that's on you.

Karine, Jordan, Yen and Annie, thank you for all the times we went out for food and of course bubble *tea*. Karine (my lobster), walking around campus with you always brightened

my day and it always warmed my heart knowing how much you valued my advice. Alexander, thanks for letting me pop into your office, wherever it was, whenever I needed. Annie, I'm glad we keep each other on the right track and thank you for always sticking up for me. Jordan, thank you for always giving me advice whether it was about everything biochemistry, or living life a bit *crooked*.

To Alan Awez. Your story still breaks my heart, but I am here because of it. I hope you have found peace wherever you are. Rest in peace.

To my family – Mom, Dad, and Natalie – I am sincerely grateful for your unconditional support even when I didn't deserve it. Shirlene, you've helped me through every chapter of my life – thank you for being there for this one as well. Meaghan, thank you for always tackling the hardest problems with me. Peter, Jerry, Meaghan, and Shirlene – you all made up an invaluable support group who battled my pessimistic outlook by never failing to make me laugh or see the bright side when I needed to most. Thank you for reminding me who I truly am and for accepting me for all my flaws and virtues.

Dedication

This thesis is dedicated to Mom and Dad, who were there to celebrate my highs, but also supported me through my lows.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Symbols, Abbreviations, and Nomenclature

Symbol	Definition
% v/v	Percent by volume
% w/w	Percent by weight
%mol	Percent mole by mole
δ	Chemical Shift or Steric reaction constant
ρ	Electronic reaction constant
Φ	Quantum yield
τ	Fluorescence lifetime
Å	Angström
anh.	Anhydrous
BME	β-mercaptoethanol
Boc	*Tert*-butyloxycarbonyl
br	Broad
calc.	Calculated
Cys	Cysteine
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
d	Double
Da	Dalton
dC10	Dicysteine with 10 Å distance between cysteines
DCM	Dichloromethane
dd	Doublet of Doublets
dec.	Decomposition
DFT	Density Functional Theory
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-(dimethylamino)pyridine
DMF	*N*,*N*-dimethylformamide
DMP	Dess-Martin periodinane
DMSO	Dimethylsulfoxide
DPBS	Dulbecco's phosphate-buffered saline
E. coli	*Escherichia coli*
EI	Electron impact
eq.	Molar equivalents
ESI	Electrospray ionization
FlARe	Fluorogenic Addition Reaction
FlAsH	Fluorogenic Arsenic Hairpin
FRET	Förster resonance energy transfer
FLIM	Fluorescence lifetime imaging microscopy
GFP	Green fluorescent protein
GSH	Glutathione
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	Heteronuclear multiple bond correlation
HMDS	Hexamethyldisilazane
HRMS	High-resolution mass spectrometry
Hz	Hertz

IPTG	Isopropyl-β-D-thiogalactoside
k_{obs}	Observed rate constant
k_2	Second-order rate constant
LB	Lysogeny Broth
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
m	Multiplet
MBP	Maltose binding protein
m.p.	Melting point
min	Minutes
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Mass spectrometry
MWCO	Molecular Weight Cut Off
NEM	*N*-ethylmaleimide
NMR	Nuclear magnetic Resonance
COSY	Correlated Spectroscopy
o.n.	Overnight
OD_{600}	Optical density at 600 nm
P	Penicillin
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein Data Bank
PeT	Photoinduced electron transfer
pK_a	Negative logarithm of the acidity constant K_a
POI	Protein of interest
PBS	Phosphate Buffered Saline
q	Quadruplet
rpm	Rotations per minute
r.t.	Room temperature
s	Singlet
S	Streptomycin
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
SPAAC	Strain-promoted azide-alkyne cycloaddition
t	Triplet
TBET	Through-bond energy transfer
TBTA	Tris(1-benzyl-1H-1,2,3-triazol-4-yl)
TCEP	Tris(2-carboxyethyl)phosphine
TD-DFT	Time-dependent density functional theory
THF	Tetrahydrofuran
TMS	Tetramethylsilane or trimethylsilyl
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
v_{init}	Initial rate

Epigraph

Life is unpredictable. Not everything's in our control.

But as long as you're with the right people, you can handle anything.

- Amy Santiago, Brooklyn Nine-Nine, Season 5, Episode 22.

Chapter 1: Introduction

1.1. Lighting Up Biology

The word protein is derived from the Greek word πρωτειος (*proteios*), meaning "of the first rank," which alludes to their integral involvement within many cellular functions. Therefore, protein labelling is vital, allowing researchers to study the localization, function, expression, and behaviour of these biomolecules in their native environment. In cell imaging, fluorescent labels can be used to easily visualize specific proteins and are advantageous over label-free cellular imaging techniques such as quantitative phase-contrast microscopy and holotomography.[1,2] These methods only capture optical changes to image cell morphology, while fluorescent protein labelling can both image protein behaviour and intracellular events. Fluorescent labels are also preferable to histochemical stains, which can only visualize specific cellular compartments and fail to capture dynamic biological processes.

Fluorescence, along with phosphorescence, is a photoluminescent process, which occurs in some molecules that emit photons after absorbing energy from light. Photoluminescence occurs in three stages, which can be visualized using a Jablonski Diagram (Figure 1.1). This process starts when an electron from the ground state (S_0) in the molecule absorbs a photon with energy $h\upsilon_1$ and is promoted to a higher energy singlet excited state (S_1') (Figure 1.1A). The energy of the electron in the S_1' state is rapidly but only partially dissipated through non-radiative transitions until it is at the lowest excited state, S_1. These non-radiative transitions can be through vibrational relaxation or internal conversion (Figure 1.1B). During vibrational relaxation, the molecule transfers vibrational energy to the surroundings as it relaxes from an excited-vibrational state to a less energetic vibrational mode and occurs within the same electronic state. On the other hand, internal conversion involves the molecule

entering a lower electronic state through the coupling of two energy-degenerate vibrational modes. In the third step, the electron in the S_1 excited state returns to the S_0 ground state in a process called relaxation. During relaxation, the electron releases a photon with energy $h\upsilon_{2(F)}$, resulting in fluorescence (Figure 1.1C).

It is possible for two photons to excite a fluorophore, which leads to an emitted photon with higher energy and shorter wavelength than each individual exciting photon. This phenomenon is integral in two-photon excitation (TPE) microscopy. However, to ensure that two photons reach the fluorophores in the sample near simultaneously, a femtosecond pulse laser is typically used. TPE microscopy is beneficial since it uses longer wavelengths to achieve deeper tissue penetration with less scattering, resulting in better sample visualization, especially for opaque tissues.[3] Longer wavelengths also reduce the likelihood of altering a fluorophore such that it is unable to fluoresce, a process called photobleaching.

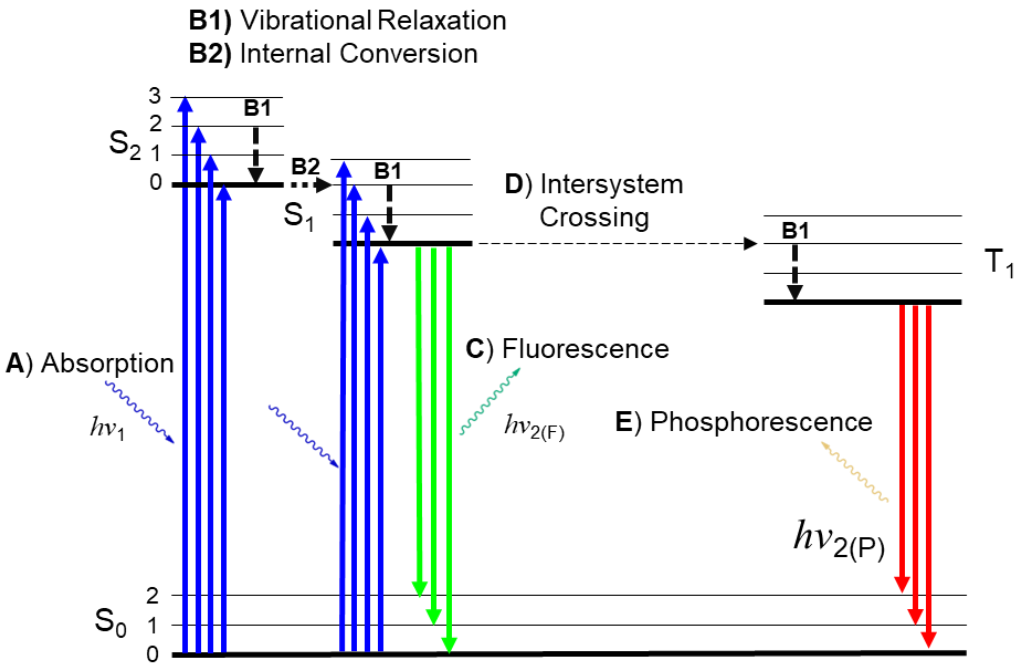


Figure 1.1. One form of a Jablonski Diagram. $h\nu_1$ = energy of the excitation photon. $h\nu_2$(F) = energy of the emitted photon as a result of fluorescence. $h\nu_2$(P) = energy of the emitted photon as a result of phosphorescence. The figure is adapted from Lakowicz, J. R., Principles of Fluorescence Spectroscopy, 3rd ed.[4]

Occasionally, the molecule in the S_1' excited state during the second step of photoluminescence can undergo intersystem crossing into a triplet state (T_1) instead of the S_1 singlet state (Figure 1.1D). The excited molecule is trapped in T_1 since it has the same spin as the ground state electron and can only relax back to S_0 through kinetically unfavourable spin "forbidden" transitions. As a result, the emission of photons with energy $h\nu_{2(P)}$, a process called phosphorescence, occurs on the order of milliseconds or longer, which is significantly slower than the nanosecond timescales on which fluorescence occurs (Figure 1.1E). These slower rates of emission make phosphorescence much less favourable than fluorescence as an imaging method.

The non-radiative transitions that occur in the second step take place extremely rapidly on the picosecond scale (10^{-12} s).[4] In particular, due to fast internal conversion from higher

electronic states to S_1, a phenomenon called Kasha's Rule, the energy of the emitted photons always occur from the lowest vibrational state of S_1. As a result of Kasha's Rule, the emission spectrum is generally independent of the excitation spectrum (Figure 1.2). Furthermore, the energy lost to these non-radiative transitions typically causes the energy of $h\upsilon_2$ to be lower than that of $h\upsilon_1$, and so fluorescence typically occurs at longer wavelengths. This Stokes shift is quantified by the difference between the wavelengths at which the excitation and emission are at maximal intensity (Figure 1.2). Kasha's Rule and fluorophores exhibiting different Stokes shifts allow for fluorescence multiplexing, in which a single photon can be used to excite multiple fluorophores, and thus multiple proteins, simultaneously. Furthermore, large Stokes shifts prevent fluorophores from self-quenching, which occurs when fluorescent molecules in close proximity absorb the emission from other molecules.[4]

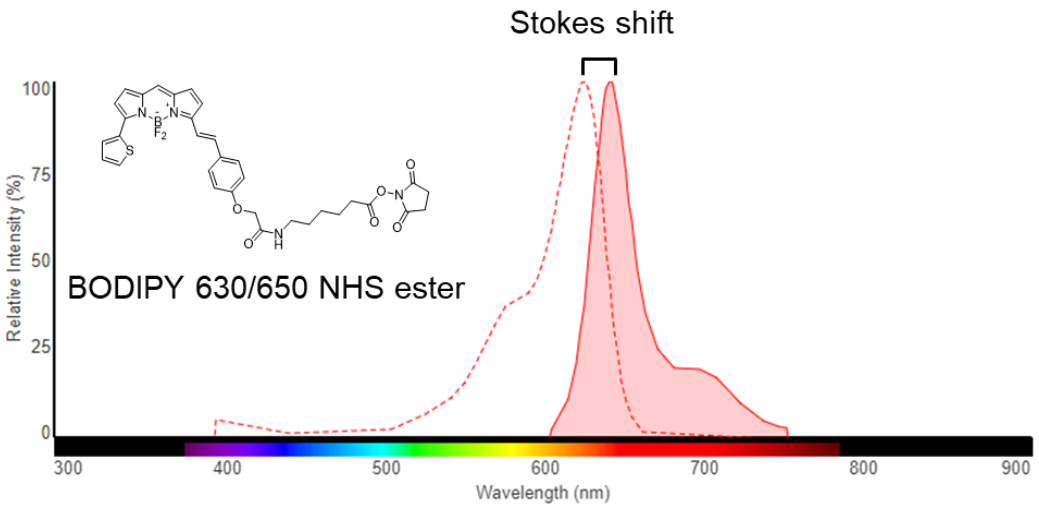


Figure 1.2. Excitation spectrum (dotted lines) and emission spectrum (solid line) of BODIPY 630/650 NHS ester. The Stokes shift is the difference in wavelength between the excitation maximum and emission maximum as indicated in the figure. The fluorophore's structure is shown in the top left corner. Spectra created using the ThermoFisher Scientific's Fluorescence SpectraViewer https://www.thermofisher.com/ca/en/home/life-science/cell-analysis/labeling-chemistry/fluorescence-spectraviewer.html.

A fluorophore is also characterized by its quantum yield (Φ), which is the number of emitted photons relative to the number of absorbed photons (Equation 1.1).[4] High quantum

yields indicate brighter fluorophores, with an upper limit of $\Phi = 1$. In addition to quantum yield, fluorescence lifetimes (τ) are also used to characterize fluorophores and describe the average time a fluorophore spends in the excited state before relaxation. This characteristic is used in Fluorescence-lifetime imaging microscopy (FLIM) and offers an imaging method that minimizes the effects of photon scattering, which occurs in thicker samples.[5]

$$\text{Equation } \mathbf{1.1}\text{: } \Phi = \frac{\#\ of\ photons\ emitted}{\#\ of\ photons\ absorbed}$$

Fluorescence detection methods can commonly detect analyte concentrations between one part per billion to one part per trillion.[6] In general, fluorescence detection can be 1 000 to 500 000 times more sensitive than absorbance spectroscopy, at least partly because it is easier to detect emitted photons as opposed to transmittance. Fluorescence is also quickly and easily visualized, requiring primarily an excitation light source, which is typically a monochromator, and a photodiode detector positioned perpendicular to the light source. Typically, mirrors and filters are used to manipulate the light that is involved. These components are essential to fluorescence microscopy, which is further explained in Chapter 3, section 3.2.2.

Cell imaging allows biologists to empirically visualize biological processes and their changes in response to external stimuli. Labelling proteins with fluorescent tags allow specific processes to be monitored in their environment and build towards a complete picture describing the biological mechanisms that govern cell biology. To achieve this goal, numerous methods have been developed (Figure 1.3). Of these protein labelling strategies, the most common involve 1) the genetic fusion to a fluorescent protein, 2) chemically labelling a

genetically fused pendant enzyme or protein, and 3) labelling a genetically encoded tag with a small molecule fluorophore.[7,8]

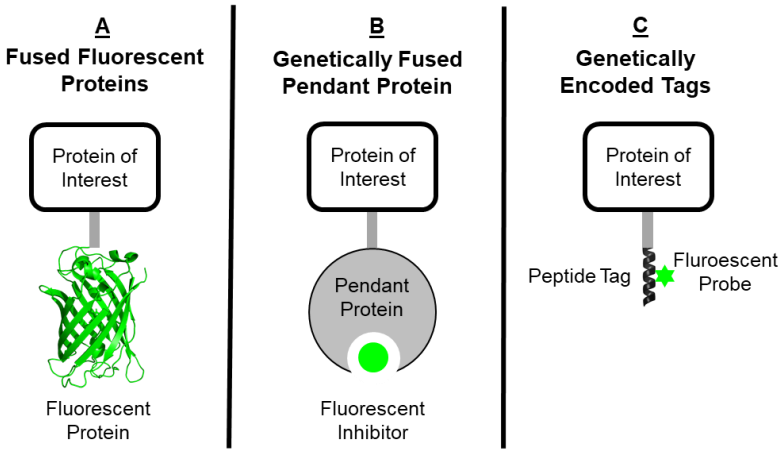


Figure 1.3. Three most common protein labelling strategies. A) Genetic fusion of the POI to a fluorescent protein. B) Chemically labelling a pendant protein on the POI with a fluorescent inhibitor. C) Labelling a genetically encoded tag with a small molecule fluorescent probe.

1.2. Protein Labelling Strategies

1.2.1. Genetic Fusion with a Fluorescent Protein

The pioneering method for protein labelling involved the genetic fusion of green fluorescent protein (GFP) to a protein of interest (POI) (Figure 1.3A). GFP was first isolated in 1962 by Shimomura from the jellyfish *Aequorea Victoria*. Chalfie then expressed the fluorescent protein in both *E. coli* and *C. elegans* in 1993 and 1994.[9–11] From Shimomura and Chalfie, Tsien contributed to the knowledge behind the mechanisms of GFP's fluorescence and developed novel GFP variants with improved brightness and folding as well as altered spectral characteristics.[12]

GFP is a 26.9-kDa large protein consisting of an eleven-stranded β-barrel with an α-helix threaded through the axis of the cylinder (Figure 1.4).[13] On the α-helix, at the centre of the can-like cylinder, resides the amino acids Ser65-Tyr66-Gly67, which spontaneously form

the fluorescent chromophore *p*-hydroxybenzylideneimidazolinone post-translationally (Figure 1.4).[12] In its fluorescent form, GFP is maximally excited at around 400 nm and has an emission maximum of 505 nm with a quantum yield $\Phi = 0.79$. It is typically expressed as a fusion partner with the POI such that the protein being studied retains its normal activity while its location, movement, and behaviour can be monitored by tracking GFP's fluorescence microscopically (Figure 1.3A).

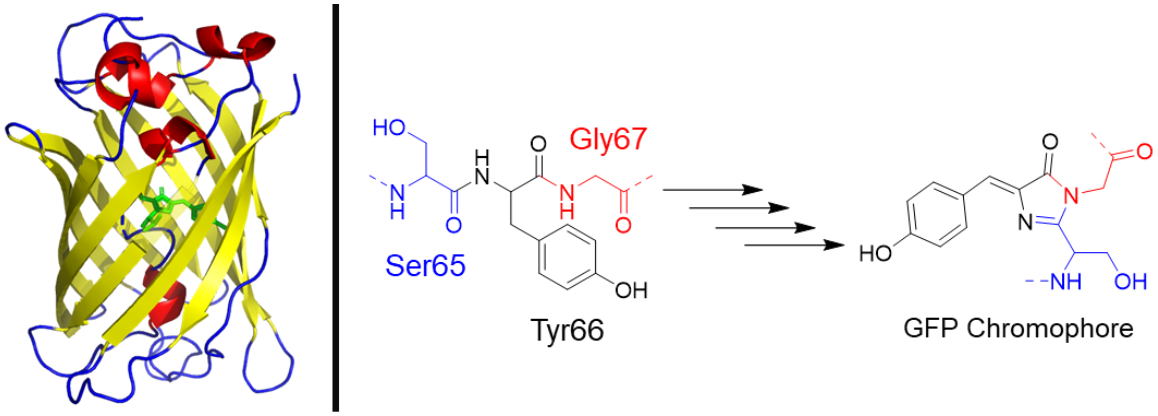


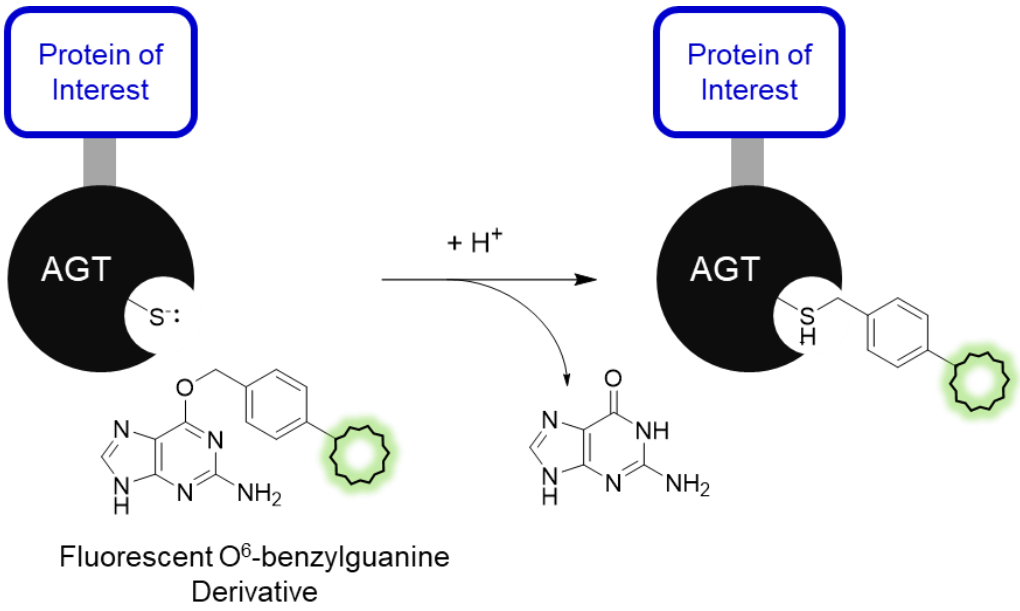
Figure 1.4. (Left) GFP's tertiary structure showing the α-helix (red) threading up through the barrel of β-sheets (yellow) with the chromophore (green) in the middle (PDB: 1EMB).[13] (Right) Formation of GFP's fluorescent chromophore, *p*-hydroxybenzylideneimidazolinone, from Ser65 (blue), Tyr66 (black), and Gly67 (red).

GFP and its augmented variants are widely used given that these fluorescent proteins fold and mature rapidly and efficiently within the cell, and their monomeric configurations facilitate their fusion to a POI. Additionally, their high brightness and photostability contribute to good signal-to-noise ratio during cell imaging. The numerous variants of GFP with different excitation and emission spectra throughout the entire visible wavelength also allow monitoring events with multiple fluorescent proteins using Förster resonance energy transfer (FRET) or to observe multiple proteins at once. Despite their strengths, fluorescent proteins still come with weaknesses. For example, GFP still takes time to fold, and maturation of the chromophore requires oxygen. Furthermore, GFP tends to form multimers and aggregates and

may disturb intercellular trafficking. The protein's large size can also disrupt the native localization and function of the POI.[14]

1.2.2. Chemically labelling a genetically fused pendant enzyme or protein

To address some of the limitations of GFP, new chemi-genetic protein labelling strategies were developed that combine the similar genetic modification of a POI with chemical labelling using a small molecule probe. One such strategy is to express the POI as a fusion partner with a pendant enzyme, which binds their natural ligands with high affinity. The pendant enzyme is subsequently labelled with a synthetic ligand that is fluorescent and shows a high affinity for the enzyme tag (Figure 1.3B). The SNAP method is a popular example of this strategy that involves appending the DNA repair protein O^6-alkylguanine-DNA alkyltransferase (AGT) to the POI through genetic modification.[15,16] O^6-alkylguanine or O^6-benzylguanine derivatized with a fluorophore will then bind and irreversibly label AGT (Scheme 1.1). Thus, the POI can then be monitored through the emission produced by this small molecule probe. Other examples of this strategy include the CLIP, Halo, and PYP labelling strategies.[7,17,18]



Scheme 1.1. Schematic of fluorescent protein labelling using the SNAP tag method, which involves the enzyme ATG and a fluorescent derivative of its ligand, O^6-benzylguanine.

This chemi-genetic approach to labelling proteins through a pendant enzyme using a fluorescent targeted inhibitor is both rapid and selective. Furthermore, synthetic ligands can be easily functionalized with various fluorophores that have different spectral properties and exhibit a range of excitation and emission spectra throughout the visible region. Unlike fluorescent proteins, the synthetic versatility of small molecule probes makes it easy to develop labels with comparable, and even superior, spectral range, brightness, and photostability. Additionally, these probes can be functionalized to cover a range of applications.[19] Regarding size, pendant enzyme tags offer a smaller alternative to fluorescent proteins, since some enzyme tags such as photoactive yellow protein are as small as 14 kDa in size.[17]

Although small molecule probes have many strengths in the way of functional flexibility and range, probes must be carefully designed for implementation into biological systems. These inhibitors must be cell-permeable and non-cytotoxic in addition to exhibiting

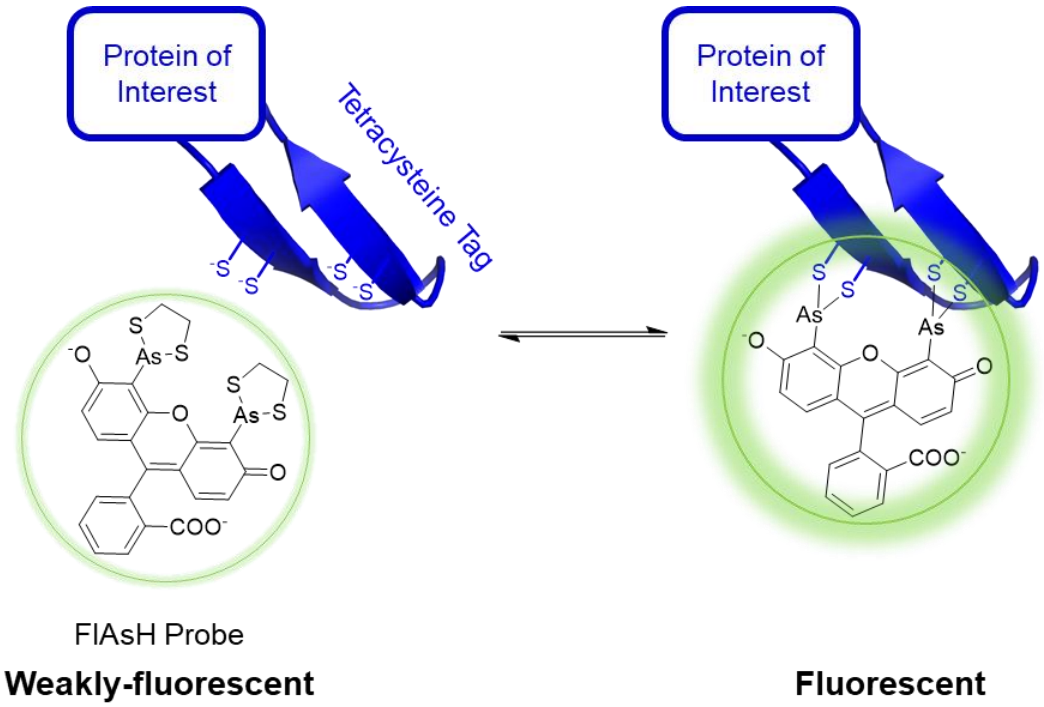
bright fluorescence, good photostability, and remain stable in the aqueous milieu of the cell. Furthermore, constantly fluorescent inhibitors are less convenient for live-cell imaging since washing steps are required to remove unreacted ligand, which can compromise cell survivability. To address this, many methods involving pendant enzyme tags have been modified or developed to incorporate **fluor**escence **gen**erating (fluorogenic) inhibitors.[7] Despite this, size remains an issue as these enzyme tags still range between 14-33 kDa in size.

1.2.3. Labelling a genetically encoded tag with a small molecule fluorophore

An even smaller alternative to using pendant enzymes as protein tags is to resort to rationally designed peptide tags. Like pendant enzymes, this chemi-genetic approach involves the genetic fusion of a small target peptide that is rationally designed to react complementary to a small molecule probe (Figure 1.3C). These probes are commonly fluorogenic and the peptide tags used in these methods generally range between 6 to 34 amino acids in size.[20–24] Although these peptide tags offer an even more minimalistic design than the pendant enzyme tags, this comes at the cost of reduced specificity and, in many cases, a more restrictive probe design that is limited to specific fluorophores.

The pioneering method to these small peptide chemi-genetic protein labelling strategies is the **Fl**uorescein **Ar**s**e**nical **H**airpin (FlAsH) method developed by Tsien (Scheme 1.2).[25] This fluorogenic protein labelling technique consists of two components: 1) The FlAsH probe, which comprises a small molecule probe composed of a fluorophore quenched by two arsenic functional groups; 2) A complementary Cys-Cys-Pro-Gly-Cys-Cys β-hairpin motif that is genetically fused to the POI.[23] When the tetra-cysteine motif reacts with the arsenic functionalities on the FlAsH probe, the quenching inherent to the arsenic groups is abolished.

This reaction restores the fluorophore's emission, allowing the POI to be monitored within the cell.



Scheme 1.2. FlAsH protein labelling strategy using a target peptide tag and a small molecule FlAsH probe.

FlAsH offers an extremely minimalistic design for a protein labelling method as it involves both a small molecule probe and a peptide tag that is six amino acids long. Tsien has also demonstrated its use in live-cells and has also developed other FlAsH-like probes with different excitation and emission spectra such as ReAsH.[23] However, FlAsH also suffers certain drawbacks due to the design of both the tetracysteine tag and the small molecule probe. For example, the arsenic-based FlAsH probe is cytotoxic and exhibits background fluorescence, despite its fluorogenic nature, and thus requires washing to remove any unreacted probe.[26,27] However, the non-covalent interaction between the tetracysteine motif and the arsenic functional groups means that extensive washing could lead to reduced labelling

efficiency. The probe design is also limited to fluorophores based on the xanthene scaffold, such as fluorescein and rhodamine, and expanding this colour palette requires conjugating multiple fluorophores together.[28] Furthermore, the tetracysteine motif itself is sensitive to oxidizing environments[29] and can form unreactive disulfide-linked aggregates.[30] Despite these weaknesses, Tsien's FlAsH protein labelling technique has inspired other similar strategies involving a small target peptide tag and a complementary small molecule probe.

The strategies mentioned above (see Figure 1.3) are not the only approaches to fluorescently labelling proteins. Other notable strategies involve the use of enzyme recognition peptide tags or unnatural amino acids, which come with their own strengths and weaknesses. Schneider reviews these strategies of fluorescent protein labelling in cells and shows that there is still much room for development regarding these chemi-genetic approaches to protein labelling.[7]

1.3. FlARe labelling method

Keillor has developed a chemi-genetic, small target peptide labelling approach to fluorescent protein tagging called FlARe, which addresses many of the weaknesses of FlAsH. The FlARe labelling method has two key components: 1) The FlARe probe, which is a fluorogenic small molecule probe; 2) The complementary target peptide called the dC10 tag (Figure 1.5). The FlARe probe comprises a fluorophore that is quenched by two maleimide functional groups connected through an aryl linker. These maleimides also serve to react with two cysteines positioned 10-10.8 Å apart on the complementary dC10 tag through Michael thiol additions.

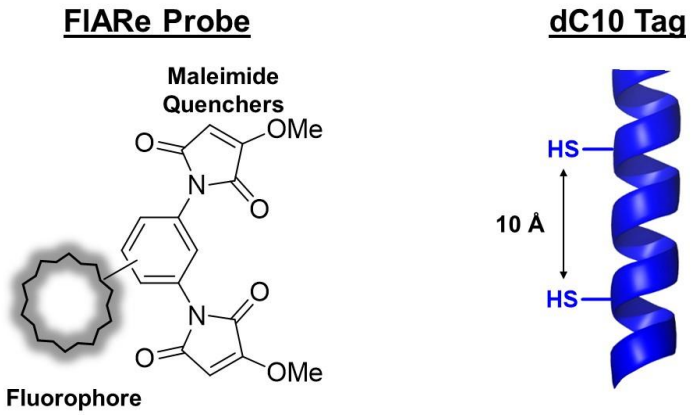
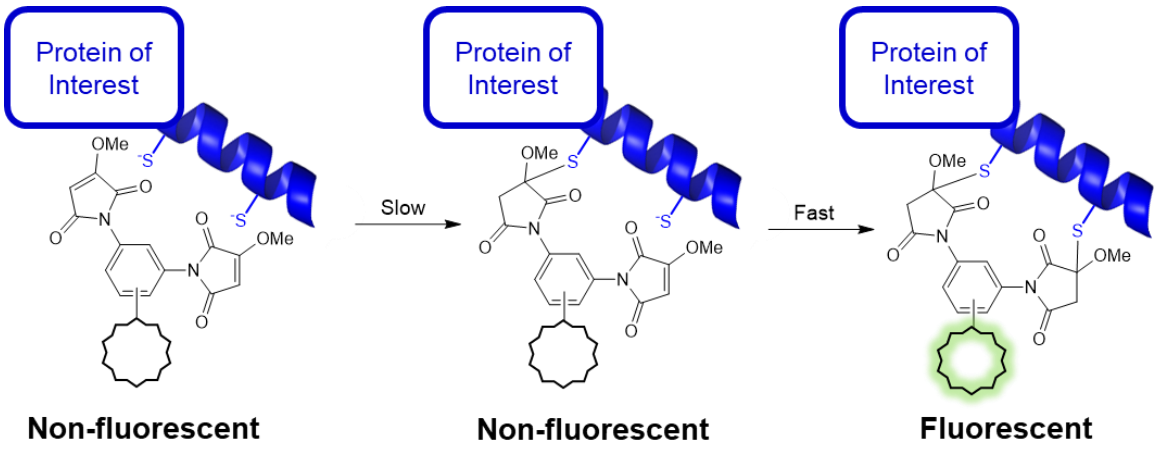


Figure 1.5. The two key components of the FlARe labelling method consisting of the FlARe probe and dC10 tag.

When the FlARe probe reacts with the dC10 tag, it does so through a **Fl**uorogenic **A**ddition **Re**action (FlARe).[31] FlARe occurs in two steps, and in the first step, one of the probe's maleimides undergoes a nucleophilic attack from one cysteine of the dC10 tag (Scheme 1.3). After this initial step, the probe remains quenched due to the presence of the second maleimide. However, the second reaction is now intramolecular, allowing the other cysteine to attack the second maleimide in a very rapid manner. After this second Michael addition reaction, both maleimides are converted to the respective succinimide, which disrupts the maleimides' quenching mechanism. This restores the fluorescence intrinsic to the FlARe probe, allowing the POI, which is genetically fused to the dC10 tag, to be monitored using microscopy.



Scheme 1.3. Schematic of the FlARe protein labelling method consisting of the FlARe probe reacting with both cysteine thiols on the dC10 tag to become fluorescent.

The strengths of the FlARe method as a fluorescent protein labelling strategy come from its synthetically versatile and non-cytotoxic probe. This probe also exhibits a very low background signal before the FlARe reaction, removing the necessity for a wash step during labelling. Even if washing was performed, the covalent nature of FlARe ensures that labelling is not compromised. Additionally, the maleimides' placement in the FlARe probe's design makes it facile to replace the fluorophore with another that exhibits different spectral properties. Regarding the tag itself, the distance between both cysteines on the dC10 tag ensures that a disulfide bridge will not form. As such, FlARe has been shown using microscopy to successfully label specific proteins carrying the dC10 tag without washing in living mammalian cells.[24] However, before being implemented into cells, which is the defining factor of a fluorescent protein labelling technique, much development had gone into the rational design of both the dC10 tag and the FlARe probe.

1.3.1. FlARe dC10 Target Peptide

The complementary dC10 peptide tag used in the FlARe labelling technique is an α-helix containing two cysteine residues.[32] The tag's helical nature ensures that the cysteine

thiols are exposed and maintain a distance of 10.8 Å between both thiols, which essentially separates both cysteines by two turns of an α-helix. Cysteine's under-representation in naturally occurring proteins and the low probability of finding two proximal cysteines contribute to FlARe's selectivity by decreasing the number of cellular interferents that could react with the FlARe probe.

Before the dC10 target peptide, a small α-helix from the Fos protein was used to show the proof-of-principle for the FlARe labelling reaction.[33] Later, a *de novo* di-cysteine peptide tag was proposed and designed based on the sequence of a synthetic peptide (Table 1.1A).[32] The salt bridges formed between the carboxylate and guanidinium groups of glutamate and arginine in this peptide sequence ensured that the peptide remained helical and soluble in aqueous media. To further stabilize this small helix, the terminal amine and acetyl groups were replaced with a C and N terminal cap known to help preserve the helix's dipoles (Table 1.1B). Additionally, two amino acids were also added to enhance helicity by ensuring an integer number of turns (Table 1.1C). Finally, two cysteine residues were introduced at i and i+7 to ensure 10.8 Å between the thiols to form the **d**i-**C**ysteine with **10** Å separation tag (dC10) (Table 1.1D). This is important as it complements the 10 Å distance between the maleimides on the FlARe probe. Ultimately, this dC10 target peptide was used for the characterization of most FlARe probes developed in the Keillor lab.[24]

Table 1.1. Peptide sequence progression of the dC10 tag.

	Peptide Sequence
A	Ac-E A A A R E A A A R E A A A R Q-NH_2
B	**L S A A** E A A A R E A A A R E A A A **K G G K**
C	L S A A E A A A R E A A A R E A A A **R A** K G G K
D (dC10)	L S A A E **C** A A R E A A **C** R E A A A R A K G G K
E (dC10*)	L S **K** A E C A A R E A A C R E **R K** A R A K G G K

Efforts have also been taken to improve the reactivity of the thiol groups of the cysteine residues in the dC10 tag. The first of these efforts introduced histidine residues proximal to the cysteines in order to help deprotonate the cysteine thiol and shift the equilibrium towards thiolate formation. However, this did not improve the reactivity of the dC10 tag as the histidine residues may have disrupted the tag's helicity and inhibit FlARe's distance dependant reactivity.[32] Another strategy was taken by introducing positively charged residues near the reactive cysteines (Table 1.1E).[34] This had the effect of decreasing the pK_a of the cysteine thiols, which shifted the thiol/thiolate equilibrium to favour the thiolate more at physiological pH. Even though doing so also decreases the nucleophilicity of the thiolate, the large increase in the proportion of thiolate compensates for this nucleophilic attenuation. This led to the dC10* sequence, which demonstrated improved reactivity with the FlARe probe.

1.3.2. FlARe Probes

The FlARe probe is the fluorogenic small molecule label that reacts with the dC10 tag to label the POI fluorescently. Its synthetic design employs a common strategy in which a reactive quencher is attached to a fluorophore to negate fluorescence through static quenching.

Typical quenching mechanisms involve **F**örster **R**esonance **E**nergy **T**ransfer (FRET), **T**hrough-**B**ond **E**nergy **T**ransfer (TBET), and **P**hotoinduced **e**lectron **T**ransfer (PeT).[35] Specifically, the FlARe probe consists of a fluorophore linked to two maleimide functional groups through an aryl linker, which quenches the probe's fluorescence through donor-excited PeT (d-PeT).[33,35] In the d-PeT mechanism, an electron promoted to the excited state in the fluorophore proximal to the maleimide is transferred to the maleimide's LUMO (Figure 1.6A). This occurs as the maleimide's LUMO is lower in energy than the fluorophore's excited state. The excited electron then returns to the fluorophore's ground state in a non-radiative manner. Thus, this electron transfer effectively quenches the fluorophore's emission if the maleimides remain present and unaltered.[31]

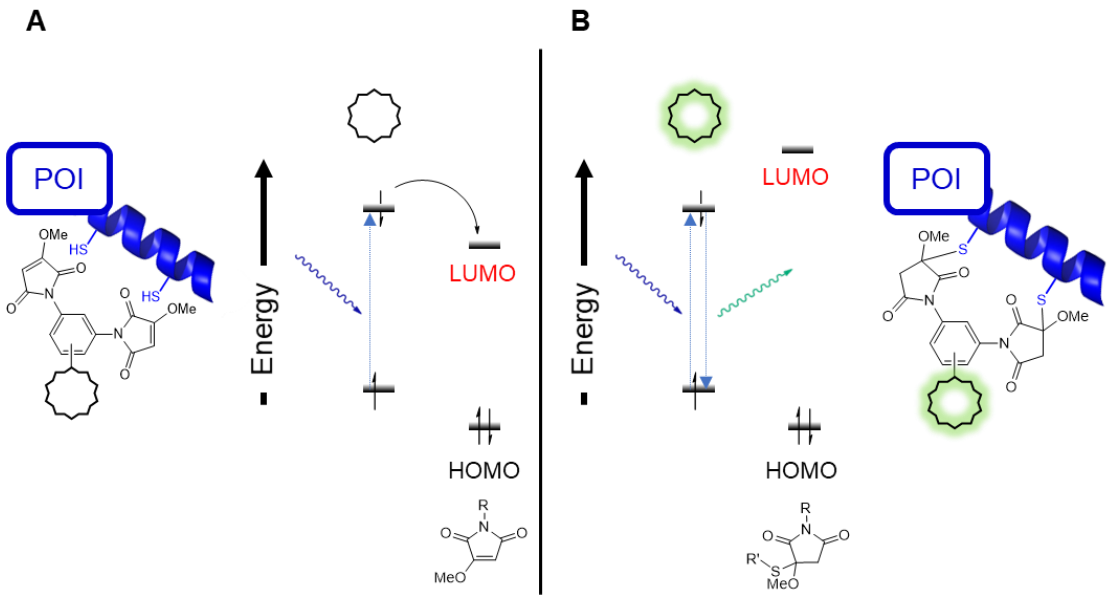


Figure 1.6. Donor-excited photoinduced electron transfer (d-PeT) mechanism that causes maleimides to quench the fluorescence of nearby fluorophores.

When both maleimides are converted to the succinimides after reacting with the dC10 tag through FlARe, the excited electron will no longer be transferred since the succinimides' LUMOs are higher in energy than fluorophore's excited state. Thus, in the absence of the

maleimides, the excited electron relaxes back to its ground state in a radiative manner and produces fluorescence (Figure 1.6B). This abrogation of the d-PeT mechanism inherent to the maleimides that are attached to the fluorophore confers the fluorogenic nature to the FlARe probes.[31] Although d-PeT offers an effective method to quenching fluorescence, it can only do so as long as the fluorophore's excited state remains higher in energy than the maleimide's LUMO. This restriction limits the fluorophores that can be implemented in the FlARe probe design as longer emitting fluorophores tend to have excited states that are lower in energy. To predict whether d-PeT is favoured and if the maleimide group can quench a fluorophore, time-dependent density functional theory (TD-DFT) calculations can be used to calculate a molecule's energy levels.[35] TD-DFT calculations have also shown that substituents can influence a molecule's energy level, allowing these molecules to be tuned for better PeT efficiency.[36,37]

Another key aspect of PeT quenching is its dependence on the distance between the electron donor and acceptor.[38] Early FlARe probe designs initially attached a di(methylmaleimide) aryl quenching moiety to the fluorophore through different diamine linkers.[39] Amongst these initial designs, the probe consisting of the 1,2-diaminocyclohexane linker demonstrated optimal quenching efficiency, which positioned the fluorophore and quenchers closest together in space. Building on this observation, a "spacerless" FlARe probe was designed, which demonstrated an immense increase in quenching efficiency. As such, new FlARe probes consist of a fluorophore directly attached to both maleimide moieties through an aryl linker.

An issue faced by the FlARe probe is its potential reactivity with adventitious thiols in the cell such as cysteine-rich proteins[27], cysteine-containing peptides like glutathione[40], or

small molecule thiols.[41] These thiols can react with the FlARe probe, leading to a fluorescent false-positive signal if both maleimides were to react, or at the very least consume the probe if only one maleimide were to react. Keillor discovered that adding methyl substituents to the maleimide ring impaired the probe's reactivity with GSH. Although the probe's reaction with the dC10 tag also slowed down, it was still faster than its reaction with GSH because of the entropic advantage offered by the constrained and dicysteine containing tag.[33] Keillor continued this strategy and explored using the increased electron-donating nature of a methoxy substituent to further attenuate the probe's reactivity with GSH. To this end, an assymetrical FlARe probe was developed containing one methylmaleimide and one methoxymaleimide under the hypothesis that although GSH may react with the methylmaleimide, it would not react with the methoxymaleimide. As such, GSH would not activate the FlARe probe. However, the intramolecular driving force that arises when the FlARe probe reacts with the dC10 tag would ensure the reaction of the tag's cysteine with the less reactive methoxymaleimide. This difference in reactivity between both maleimides was hypothesized to improve the FlARe probe's selectivity without significantly decreasing labelling speeds. Although this strategy worked to some degree, GSH was still found to react with the asymmetrical FlARe probe. In response, a symmetrical FlARe probe consisting of two methoxymaleimides was developed, which demonstrated negligible reactivity with GSH under physiological conditions while maintaining reactivity with the dC10 tag, albeit very slow reactivity. [24]

1.4. Thesis Objectives

A chemi-genetic approach to fluorescent protein labelling involves chemically labelling a target peptide genetically fused to the POI with a complementary small molecule

probe. This strategy offers a minimalistic labelling method that avoids the potential disruption caused by larger fluorescent proteins and pendant enzyme tags. The FlARe method, developed by Keillor, is an example of this strategy that involves reacting the genetically fused dC10 tag with a fluorogenic FlARe probe. Once complete, the fluorescently labelled, dC10 tagged POI can be visualized by fluorescence microscopy. The primary objective of this thesis is to develop the FlARe protein labelling method towards its future use in cellular systems and to explore additional applications for the FlARe method. This work was focussed on three main goals: 1) Tuning the selectivity and stability of the FlARe probe; 2) Implementing novel FlARe components in cells; 3) Expanding the FlARe labelling scope to include bioconjugation.

1.4.1. Tuning the selectivity and stability of the FlARe probe

In Chapter 2, the mechanisms governing maleimide reactivity were studied in order to improve the FlARe probe's selectivity for the dC10 tag and stability in the aqueous milieu of the cell. Model *N*-arylmaleimides substituted on the maleimide ring were studied to determine the regioselectivity of maleimide thiol addition and hydrolysis. Additionally, the rate constants of these reactions for a library of fluorogenic maleimide probes, also substituted on the ring alkene, were measured. Linear free energy relationships were studied to ascertain how the maleimide ring substituents influence these reactions. The insight gained from this mechanistic study was then applied to design a novel FlARe probe with improved selectivity and stability. The results of this chapter not only contributed to optimizing future FlARe probes as specific protein imaging agents but also to tuning novel maleimide-based conjugation systems towards their specific applications.

1.4.2. Implementing novel FlARe components in cells

A green fluorogenic FlARe probe (originally developed by Dr. Yingche Chen and Sydney Acton) and a new faster reacting dC10 peptide tag (originally developed by Miroslava Strmiskova) were previously developed in the Keillor group. In Chapter 3, these novel tools are implemented in mammalian cells, allowing cellular imaging using widefield fluorescence microscopy. The novel dC10 tag was assessed for faster labelling times in comparison to the old dC10 tag while the green fluorogenic probe was evaluated for its ability to label a dC10 tagged protein located in the nucleus. This chapter established the groundwork for demonstrating and evaluating the final cellular aspects of the FlARe protein imaging method.

1.4.3. Expanding the FlARe labelling scope to include bioconjugation

In Chapter 4, the scope of the FlARe method is expanded to include the bioconjugation of small molecules to proteins. The current FlARe probe design was adapted to include a terminal alkyne for the conjugation of azide-functionalized small molecule cargo onto a protein carrying the dC10 tag. A dC10 tagged protein was first labelled with a novel FlARe bioconjugating linker. Following this, an azide-functionalized fluorophore was conjugated to the protein through the FlARe linker at the dC10 tag. The results in Chapter 4 demonstrates the design, synthesis, and evaluation of a new fluorogenic and site-specific bioconjugating linker. This is significant as bioconjugation is an area of application that has gained traction in therapies combining the specificity of biomolecules and the functional range of small molecules.

1.5. Chapter 1 References

(1) Park, Y. K.; Depeursinge, C.; Popescu, G. (2018) Quantitative Phase Imaging in Biomedicine. *Nat. Photonics* 12, 578–589.

(2) Cotte, Y.; Toy, F.; Jourdain, P.; Pavillon, N.; Boss, D.; Magistretti, P.; Marquet, P.; Depeursinge, C. (2013) Marker-Free Phase Nanoscopy. *Nat. Photonics* 7, 113–117.

(3) Denk, W.; Strickler, J. H.; Webb, W. W. (2016) Two-Photon Laser Scanning Fluorescence Microscopy. *Science* 248, 73–76.

(4) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; (2006).

(5) Willem Borst, J.; Visser, A. J. W. G. (2010) Fluorescence Lifetime Imaging Microscopy in Life Sciences. *Meas. Sci. Technol.* 21, 1–21.

(6) Rye, H. S.; Dabora, J. M.; Quesada, M. A.; Mathies, R. A.; Glazer, A. N. (1993) Fluorometric Assay Using Dimeric Dyes for Double- and Single-Stranded DNA and RNA with Picogram Sensitivity. *Anal. Biochem.* 208, 144–150.

(7) Schneider, A. F. L.; Hackenberger, C. P. R. (2017) Fluorescent Labelling in Living Cells. *Curr. Opin. Biotechnol.* 48, 61–68.

(8) Crivat, G.; Taraska, J. W. (2012) Imaging Proteins inside Cells with Fluorescent Tags. *Trends Biotechnol.* 30, 8–16.

(9) Shimomura, O.; Johnson, F. H.; Saiga, Y. (1962) Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. *J. Cell. Comp. Physiol.* 59, 223–239.

(10) Chalfie, M.; Tu, Y.; Prasher, D. C. (1993) Glow Worms: A New Method of Looking

at C. Elegans Gene Expression. *Worm Breeder's Gaz.* 13, 19.

(11) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. (1994) Green Fluorescent Protein as a Marker for Gene Expression. *Science* 263, 802–805.

(12) Tsien, R. Y. (1998) The Green Fluorescent Protein. *Annu. Rev. Biochem* 67, 509–553.

(13) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. (1997) Structural Basis for Dual Excitation and Photoisomerization of the Aequorea Victoria Green Fluorescent Protein. *Proc. Natl. Acad. Sci. U. S. A.* 94, 2306–2311.

(14) Landgraf, D.; Okumus, B.; Chien, P.; Baker, T. A.; Paulsson, J. (2012) Segregation of Molecules at Cell Division Reveals Native Protein Localization. *Nat. Methods* 9, 480–482.

(15) Mollwitz, B.; Brunk, E.; Schmitt, S.; Pojer, F.; Bannwarth, M.; Schiltz, M.; Rothlisberger, U.; Johnsson, K. (2012) Directed Evolution of the Suicide Protein O 6-Alkylguanine-DNA Alkyltransferase for Increased Reactivity Results in an Alkylated Protein with Exceptional Stability. *Biochemistry* 51, 986–994.

(16) Gendreizig, S.; Kindermann, M.; Johnsson, K. (2003) Induced Protein Dimerization in Vivo through Covalent Labeling. *J. Am. Chem. Soc.* 125, 14970–14971.

(17) Hori, Y.; Ueno, H.; Mizukami, S.; Kikuchi, K. (2009) Photoactive Yellow Protein-Based Protein Labeling System with Turn-on Fluorescence Intensity. *J. Am. Chem. Soc.* 131, 16610–16611.

(18) Hoelzel, C. A.; Zhang, X. (2020) Visualizing and Manipulating Biological Processes

Using HaloTag and SNAP-Tag Technologies. *ChemBioChem* cbic.202000037.

(19) Terai, T.; Nagano, T. (2013) Small-Molecule Fluorophores and Fluorescent Probes for Bioimaging. *Pflugers Arch. Eur. J. Physiol.* 465, 347–359.

(20) Sunbul, M.; Nacheva, L.; Jäschke, A. (2015) Proximity-Induced Covalent Labeling of Proteins with a Reactive Fluorophore-Binding Peptide Tag. *Bioconjug. Chem.* 26, 1466–1469.

(21) Hu, G.; Jia, H.; Zhao, L.; Cho, D.; Fang, J. (2019) Small Molecule Fluorescent Probes of Protein Vicinal Dithiols. *Chinese Chem. Lett.* 30, 1704–1716.

(22) Lee, J. J.; Lee, S. C.; Zhai, D.; Ahn, Y. H.; Yeo, H. Y.; Tan, Y. L.; Chang, Y. T. (2011) Bodipy-Diacrylate Imaging Probes for Targeted Proteins inside Live Cells. *Chem. Commun.* 47, 4508–4510.

(23) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. (2002) New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *J. Am. Chem. Soc.* 124, 6063–6076.

(24) Chen, Y.; Clouthier, C. M.; Tsao, K.; Strmiskova, M.; Lachance, H.; Keillor, J. W. (2014) Coumarin-Based Fluorogenic Probes for No-Wash Protein Labeling. *Angew. Chem. Int. Ed. Engl.* 53, 13785–13788.

(25) Griffin, B. A., Adams, S. R., Tsien, R. Y. (1998) Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* 281, 269–272.

(26) Langhorst, M. F.; Genisyuerek, S.; Stuermer, C. A. O. (2006) Accumulation of

FlAsH/Lumio Green in Active Mitochondria Can Be Reversed by β-Mercaptoethanol for Specific Staining of Tetracysteine-Tagged Proteins. *Histochem. Cell Biol.* 125, 743–747.

(27) Stroffekova, K.; Proenza, C.; Beam, K. G. (2001) The Protein-Labeling Reagent FLASH-EDT2 Binds Not Only to CCXXCC Motifs but Also Non-Specifically to Endogenous Cysteine-Rich Proteins. *Pflugers Arch. Eur. J. Physiol.* 442, 859–866.

(28) Bhunia, A. K.; Miller, S. C. (2007) Labeling Tetracysteine-Tagged Proteins with a SpIAsH of Color: A Modular Approach to Bis-Arsenical Fluorophores. *ChemBioChem* 8, 1642–1645.

(29) Soh, N. (2008) Selective Chemical Labeling of Proteins with Small Fluorescent Molecules Based on Metal-Chelation Methodology. *Sensors* 8, 1004–1024.

(30) Chen, B.; Mayer, M. U.; Markillie, L. M.; Stenoien, D. L.; Squier, T. C. (2005) Dynamic Motion of Helix A in the Amino-Terminal Domain of Calmodulin Is Stabilized upon Calcium Activation. *Biochemistry* 44, 905–914.

(31) Guy, J.; Caron, K.; Dufresne, S.; Michnick, S. W.; Skene, W. G.; Keillor, J. W. (2007) Convergent Preparation and Photophysical Characterization of Dimaleimide Dansyl Fluorogens: Elucidation of the Maleimide Fluorescence Quenching Mechanism. *J. Am. Chem. Soc.* 129, 11969–11977.

(32) Guy, J.; Castonguay, R.; Campos-Reales Pineda, N. B.; Jacquier, V.; Caron, K.; Michnick, S. W.; Keillor, J. W. (2010) De Novo Helical Peptides as Target Sequences for a Specific, Fluorogenic Protein Labelling Strategy. *Mol. Biosyst.* 6, 976–987.

(33) Girouard, S.; Houle, M.-H.; Grandbois, A.; Keillor, J. W.; Michnick, S. W. (2005) Synthesis and Characterization of Dimaleimide Fluorogens Designed for Specific Labeling of Proteins. *J. Am. Chem. Soc.* 127, 559–566.

(34) Strmiskova, M.; Tsao, K.; Keillor, J. W. (2018) Rational Design of a Highly Reactive Dicysteine Peptide Tag for Fluorogenic Protein Labelling. *Org. Biomol. Chem.* 16, 6332–6340.

(35) Chen, Y.; Tsao, K.; Keillor, J. W. (2014) Fluorogenic Protein Labelling: A Review of Photophysical Quench Mechanisms and Principles of Fluorogen Design. *Can. J. Chem.* 398, 1–10.

(36) Greene, L. E.; Lincoln, R.; Cosa, G. (2019) Tuning Photoinduced Electron Transfer Efficiency of Fluorogenic BODIPY-α-Tocopherol Analogues. *Photochem. Photobiol.* 95, 192–201.

(37) Chen, Y.; Tsao, K.; De Francesco, É.; Keillor, J. W. (2015) Ring Substituent Effects on the Thiol Addition and Hydrolysis Reactions of N-Arylmaleimides. *J. Org. Chem.* 80, 12182–12192.

(38) Matsumoto, T.; Urano, Y.; Shoda, T.; Kojima, H.; Nagano, T. (2007) A Thiol-Reactive Fluorescence Probe Based on Donor-Excited Photoinduced Electron Transfer: Key Role of Ortho Substitution. *Org. Lett.* 9, 3375–3377.

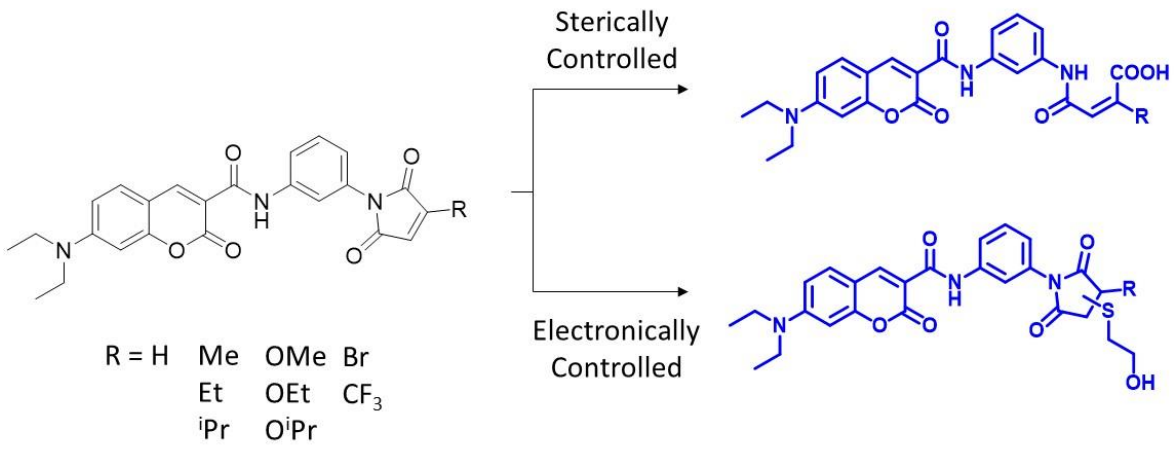
(39) Caron, K.; Lachapelle, V.; Keillor, J. W. (2011) Dramatic Increase of Quench Efficiency in “Spacerless” Dimaleimide Fluorogens. *Org. Biomol. Chem.* 9, 185–197.

(40) Jiang, X.; Yu, Y.; Chen, J.; Zhao, M.; Chen, H.; Song, X.; Matzuk, A. J.; Carroll, S. L.;

Tan, X.; Sizovs, A.; et al. (2015) Quantitative Imaging of Glutathione in Live Cells Using a Reversible Reaction-Based Ratiometric Fluorescent Probe. *ACS Chem. Biol.* 10, 864–874.

(41) Selhub, J. (1999) Homocysteine Metabolism. *Annu. Rev. Nutr.* 19, 217–246.

Chapter 2: Studying Maleimide Reactivity to Improve the FlARe Probe



2.1. Statement of Contribution

The results and conclusions reported in this chapter have been adapted with permission from Chen, Y.; Tsao, K.; De Francesco, É.; Keillor, J. W. (2015) Ring Substituent Effects on the Thiol Addition and Hydrolysis Reactions of N-Arylmaleimides. *J. Org. Chem.* 80, 12182–12192.[1] Copyright (2015) American Chemical Society. See Appendix for license. This chapter consists of the combined works from Dr. Yingche Chen, Élise De Francesco, and Kelvin Tsao, who worked together under Dr. Jeffrey Keillor as the principal investigator.

In addition to being the principal investigator, Dr. Keillor was also responsible for all the Density Functional Theory (DFT) calculations performed herein. He also guided the analysis of the linear free energy relationship analysis of the rate constants that were measured.

Dr. Yingche Chen designed the synthetic strategy concerning the coumarin maleimide substrates and the assay used to evaluate each substrate. The synthesis of compound **M2**, **M3**, **M5**, **2.10** and **2.11** was also performed by Dr. Chen, who also evaluated the kinetics of thiol

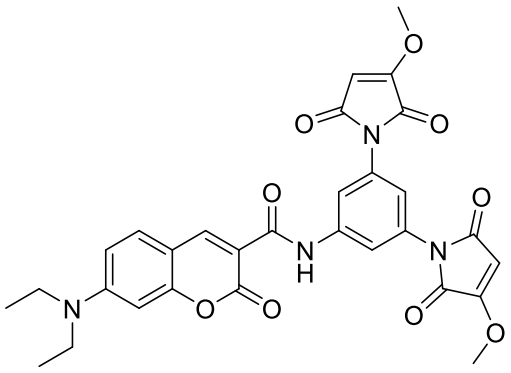
addition to **S1** and **S8**. Additionally, Dr. Chen performed the hydrolysis experiments for all the substrates.

Élise De Francesco first established the synthetic strategy designed by Dr. Chen and synthesized substrates **S1**, **S2**, **S5**, and **S8**. Élise also performed the kinetic evaluation of thiol addition to these substrates.

Kelvin Tsao followed the work established by Élise and Dr. Chen, leading to the synthesis and kinetic evaluation of substrates **S3**, **S4**, **S6**, and **S7**. He also developed the synthesis for **S9** and evaluated **S9** using the same assay designed by Dr. Chen. Additionally, **S1** and **S8** were also reconfirmed synthetically by Kelvin Tsao. Kelvin Tsao also performed the NMR studies necessary to establish the products of maleimide thiol addition and hydrolysis using **M2**, **M3**, and **M5**.

2.2. Background

To date, **YC20** is the most selective FlARe probe for the dC10 tag due to the two methoxymaleimides that are linked to a coumarin fluorophore core through an aryl linker (Figure 2.1).[2] Prior studies showed that this methoxymaleimide-FlARe probe demonstrated negligible reactivity with the cellular monothiol GSH and visible but slow reactivity with the complementary peptide tag. Likely, the entropic advantage offered by the distance-conforming, dicysteine dC10 tag made it more kinetically favourable for the FlARe probe to react with the target peptide but not monothiols.[3] This imparts a kinetically based selectivity on the FlARe probe against adventitious thiols that may activate the probe's fluorescence in the cell, which would produce background signal. In this chapter, we aim to better understand maleimide reactivity to improve the current FlARe probe structure through rational design.



YC20

Figure 2.1. Structure of YC20, the coumarin-based FlARe probe most selective for the dC10 target peptide.

2.2.1. Maleimide Reactivity

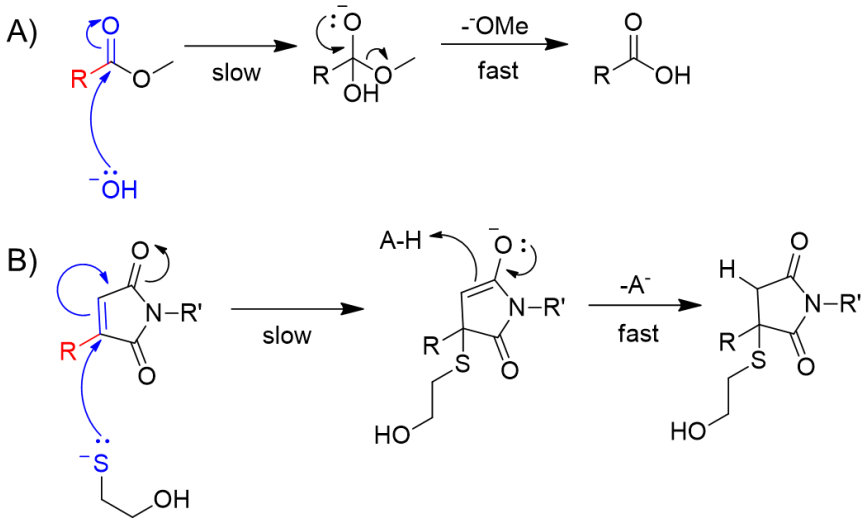
Maleimide groups are subject to Michael thiol addition during which a thiolate nucleophile attacks the C=C bond of the maleimide ring leading to the succinimide product. This reaction is integral to the FlARe labelling technique as it both disrupts the quenching of

the maleimide to activate the probe's fluorescence and anchors the FlARe probe to the target dC10 tag. Maleimides can also undergo hydrolysis if a hydroxide nucleophile attacks either imide carbonyl on the maleimide. This reaction is detrimental to FlARe labelling as maleimide hydrolysis regenerates some fluorescence, thus contributing to the slow appearance of background signal, and renders the maleimide unsusceptible to thiol addition, which prevents protein labelling.[4]

To understand the mechanisms of maleimide thiol addition and hydrolysis in-depth, the parameters that govern these reactions can be studied using linear free energy relationships (LFERs). In this process, the logarithms of a series of reaction rate constants or equilibrium constants for a given reaction are compared to a series of established substituent constants. In principle, many substituent constants are considered and those sets that demonstrate the best correlations inform on the mechanism. In this study, LFERs are used to study the maleimide substituent's electronic and steric effects on thiol addition and hydrolysis. The two series of substituent constants related to these effects are the electronic Hammett σ^+ constants and Taft steric constants, respectively.[5]

The Hammett σ constants were developed during studies that explored how the electronic characteristics of substituents directly attached to a conjugated π system of a benzene ring influenced the hydrolysis of substituted ethylbenzoates.[6] Both maleimide hydrolysis and thiol addition are similar as the maleimide acts as the electrophile in these reactions and the substituent is directly linked to the conjugated π system of the maleimide. On the other hand, the Taft constants were developed during studies that explored how the steric properties of substituents influenced the hydrolysis of substituted methylesters.[7] Similarly, thiol attack on the substituted C=C bond during maleimide thiol addition mirrors

the hydroxide attack on the C=O bond during ester hydrolysis (Scheme 2.1). As such, maleimide thiol addition can be modeled after the hydrolysis of a substituted methylester in addition to maleimide hydrolysis. Thus, the similarities shared between the reactions studied herein and the ester hydrolysis reactions allow the Hammett σ and Taft substituent constants to be used to study the maleimide substituent's electronic and steric effects on hydrolysis and thiol addition. However, in this study we implement the σ^+ constants, which were derived from the solvolysis of cumyl chlorides, [8] instead of Hammett's σ constants. This is because the maleimide substituent (R) is conjugated to a reaction centre that delocalizes the positive nature of the electrophilic carbon being attacked by the thiolate (Scheme 2.1B).[5]



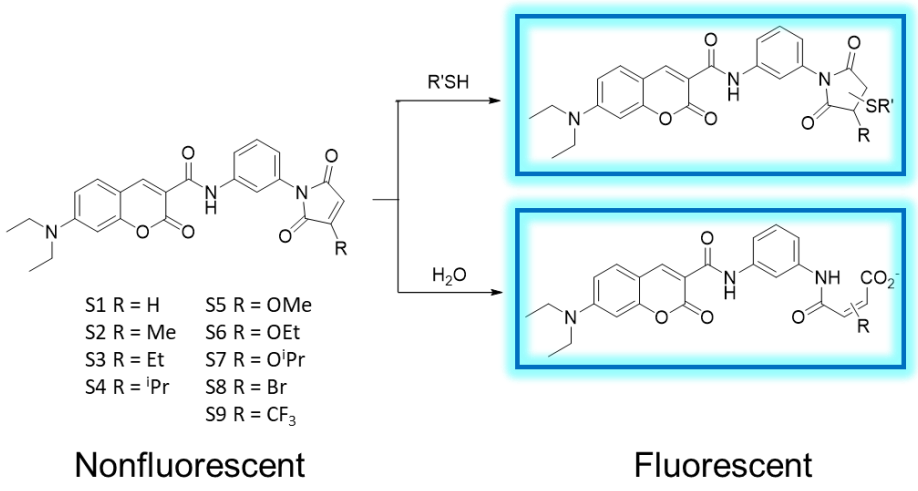
Scheme 2.1. A) Mechanism of the hydrolysis of substituted methyl ester used to establish the Taft substituent constants. B) Mechanism of maleimide thiol addition. Subsitutents are highlighted in red while the blue mechanism arrows show the similarities between both reaction mechanisms.

2.2.2. Towards Studying N-arylmaleimides

Thus far, *N*-alkylmaleimide hydrolysis[9,10] and thiol addition[11,12] reactions have been studied extensively. However, there has been less focus on how the substituents on the maleimide C=C double bond influence the kinetics of maleimide hydrolysis[4] and thiol addition[13] of *N*-arylmaleimides. Given that the FlARe design uses two *N*-arylmaleimides, we

performed systematic kinetic studies on both the hydrolysis and thiol addition reactions involving *N*-arylmaleimides.

Although reactions with *N*-alkylmaleimides were monitored by spectrophotometrically following the decrease of absorbance at 302 nm, the aromatic group of *N*-arylmaleimides occludes any changes at this wavelength.[9–12,14,15] Therefore, we developed a fluorometric assay to study the reaction kinetics of these compounds.[13] As previously mentioned, maleimide groups can quench fluorophores through a photoinduced electron transfer (PeT) mechanism, which can be abrogated upon reaction to release fluorescence.[16–19] Thus, a strategy was adopted in which a fluorophore was attached to a fluorescence quenching *N*-arylmaleimide through an aryl linker.



Scheme 2.2. Fluorometric assay design to study the thiol addition and hydrolysis of coumarin maleimide substrates **S1-S9**.

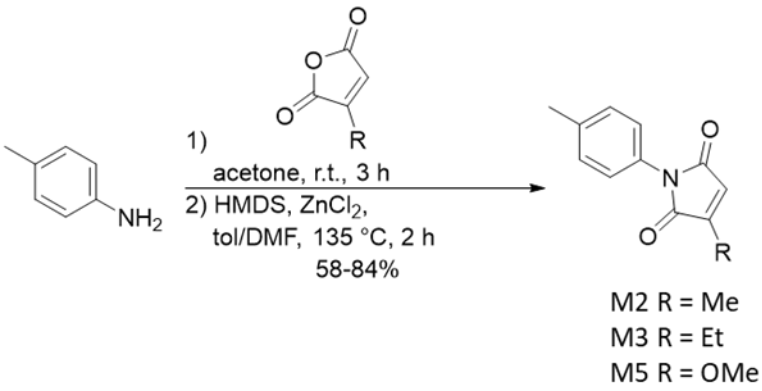
In order to remain consistent with our FlARe probe design, we decided to use 7-diethylamino-3-carboxylcoumarin as the quenched fluorophore.[2] The coumarin's synthetic accessibility, efficient quenching by maleimides, high quantum yield, photostability and hydrophilicity make this fluorophore suitable for kinetic studies under aqueous conditions.[20–

[26] As such, we designed a series of *N*-arylmaleimides, each bearing a different substituent on the maleimide C=C double bond (**S1-S9**) and a 7-diethylaminocoumarin-3-ylcarboxamide fluorophore at the *meta* position of the central phenyl ring. Upon reaction of the maleimide group by either hydrolysis or thiol addition, the cyan fluorescence of the coumarin should be restored as the PeT quenching mechanism inherent to the maleimide is abolished (Scheme 2.2). Unlike the FlARe probe design, our library of coumarin probes contained only one maleimide. This intended change simplifies the analysis involving the substituent's influence on the reactions studied herein.

2.3. Results and Discussion

2.3.1. Product Studies

Prior to kinetic studies, we performed product studies using structural analysis by NMR in order to confirm the predicted reactivity of maleimide hydrolysis and thiol addition. Therefore, we designed model compounds with simpler structures compared to the above coumarin maleimide substrates (Scheme 2.2). Model compounds **M2**,[27,28] **M3**,[29] and **M5** were prepared by reacting *p*-toluidine with the appropriate maleic anhydride as shown in Scheme 2.3. The synthesis of the maleic anhydrides is described later (*vida infra*).



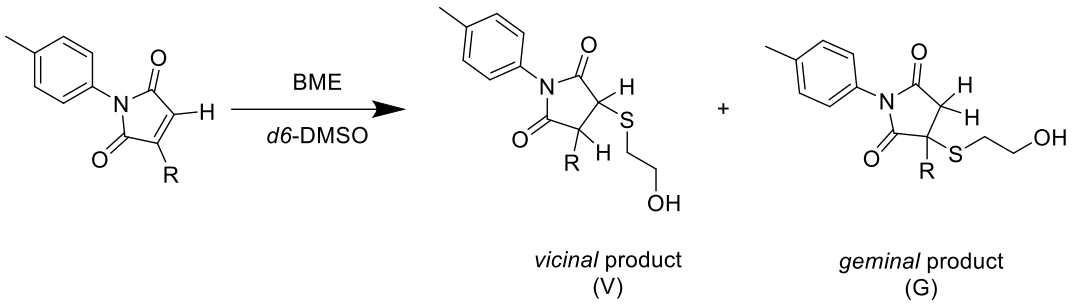
Scheme 2.3. Synthesis of model compounds M2, M3 and M5 for the product studies.[1]

2.3.1.1. Thiol Addition Product Studies

To study the products of maleimide thiol addition, model substrates **M2**, **M3** and **M5** were reacted with one equivalent of β-mercaptoethanol (BME) in *d*6-DMSO at room temperature overnight inside an NMR tube (see Experimental). Afterwards, the expected thiol addition products were confirmed by NMR analysis. The singlet at 2.34 ppm belonging to the toluidine's *p*-methyl substituent remained constant throughout the reaction and was thus chosen as an integration standard. In doing so, both the degree of conversion and regioselectivity of each thiol addition reaction was determined (Table 2.1). Thiol addition of BME to each model compound could occur either on the substituted carbon, resulting in the "*geminal*" product (G), or on the unsubstituted carbon to yield the "*vicinal*" product (V, Scheme 2.4).

Table 2.1. Product studies of thiol addition reaction with BME on model compounds **M2**, **M3**, and **M5** (150 mM model compound, 150 mM BME, *d6*-DMSO, 25 °C, overnight).[1]

Model compound	R	Conversion (%)	Regioselectivity (V:G)
M2	Me	93	58:42
M3	Et	77	84:16
M5	OMe	47	0:100



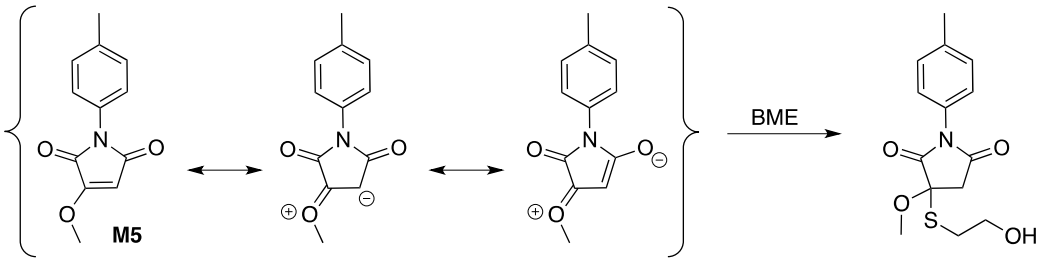
Scheme 2.4. Regioselectivity of BME addition to the model compounds.[1]

For **M2**, a peak at 6.77 ppm indicated residual **M2**, allowing for a percent conversion to product of 93% to be calculated (Figure 2.2). Additionally, two peaks at 1.64 ppm and 1.25 ppm had a combined integration of approximately 2.5, which is likely the combined integration of the methyl CH_3 peak of both regioisomers after accounting for unreacted **M2**. The singlet at 1.64 ppm likely belongs to the methyl substituent on the G product, as addition of BME adjacent to the substituent would deshield the methyl protons, resulting in a higher chemical shift (Figure 2.2). With this reasoning, the peak at 1.25 ppm belongs to the V product, which is expected to be a doublet due to the proton adjacent to the methyl substituent as opposed to BME as in the V product.

As with **M2**, residual **M3** was identified by the peak at 6.75 showing that 77% of **M3** was converted to the thiol addition product (Figure 2.3). The analysis of **M3** was more difficult with regards to product identification. However, a COSY spectrum distinguished the ethyl group's methylene in G (1.8-2.0 ppm) from that of V (1.6-1.7 ppm) (Figure 2.4). Furthermore, the protons between the succinimide ring of G, appearing as doublets at 3.25 and 2.66 ppm, demonstrated geminal coupling (J = 18.4 Hz), which further confirmed the integration of the assigned peaks. Overall, BME was found to add preferentially to the less substituted carbon when reacting with the methyl-substituted maleimide, and even more so for the ethyl-

substituted variant. This suggests that the observed regioselectivity may be influenced by steric hinderance.

In the case of the methoxy derivative **M5**, only 47% conversion was reached (Figure 2.5). A doublet at 2.93 ppm demonstrated geminal coupling (J = 20.0 Hz) and integrated to 0.5. This peak likely belongs to one of the succinimide protons (H^1) in the G product, while the integration suggests that only one regioisomer was synthesized. A COSY spectrum allowed the second succinimide proton in the G product (H^2) to be identified as the peaks at 3.51 and 3.56 ppm. Together, these peaks integrate to 0.5 and have a geminal coupling constant of J = 20.0 Hz, which is consistent with coupling to the peaks at 2.93 ppm. Exclusive formation of the G regiosimer may be because the strong electron donating character of the methoxy group through resonance with the maleimide double bond increases the electron density of the unsubstituted carbon, positioning the diffuse LUMO on the substituted carbon (Scheme 2.5). The increased shielding of the maleimide proton in **M5**, whose chemical shift is 6.01 ppm as opposed to 6.77 and 6.75 ppm for the methyl and ethyl derivatives respectively, also reflect this resonance delocalisation.



Scheme 2.5. Regioselectivity and resonance model of BME addition to **M5**.[1]

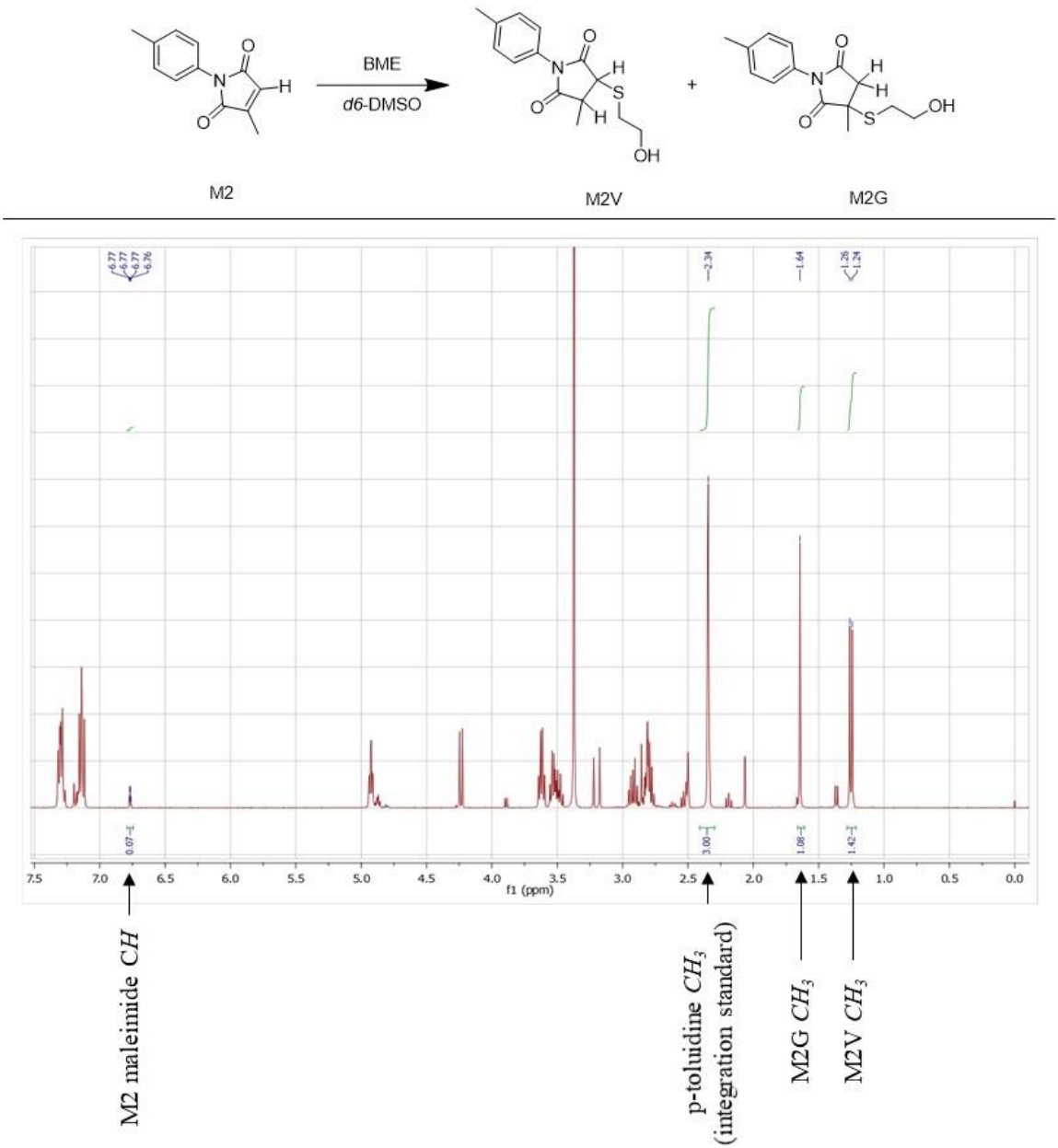


Figure 2.2. [1]H-NMR spectrum of the reaction between **M2** and BME (see Experimental). Protons used to determine % conversion and regioselectivity (V:G ratio) have been shown with arrows.

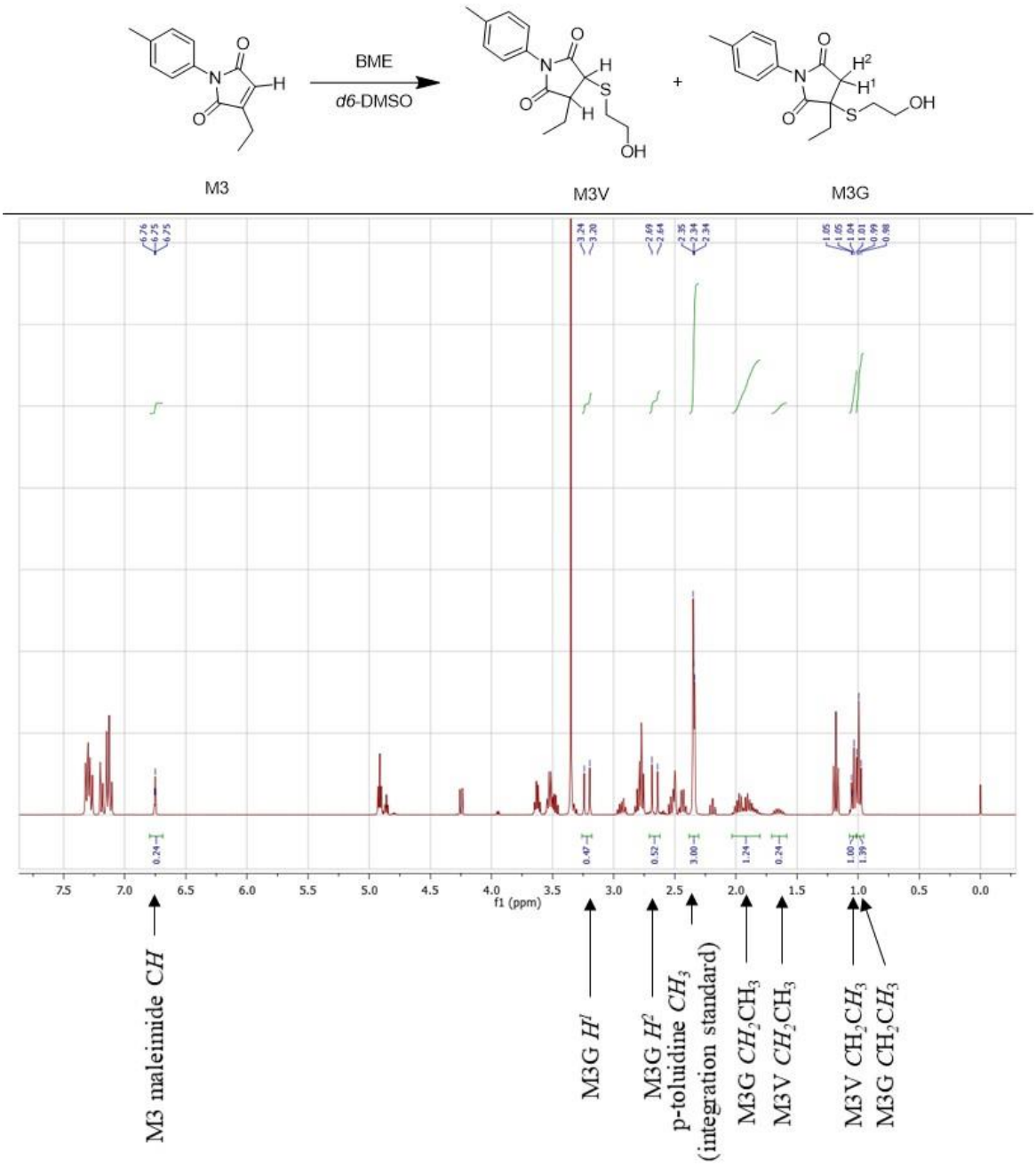


Figure 2.3. ^{1}H-NMR spectrum of the reaction between **M3** and BME (see Experimental). Protons used to determine % conversion and regioselectivity (V:G ratio) have been shown with arrows.

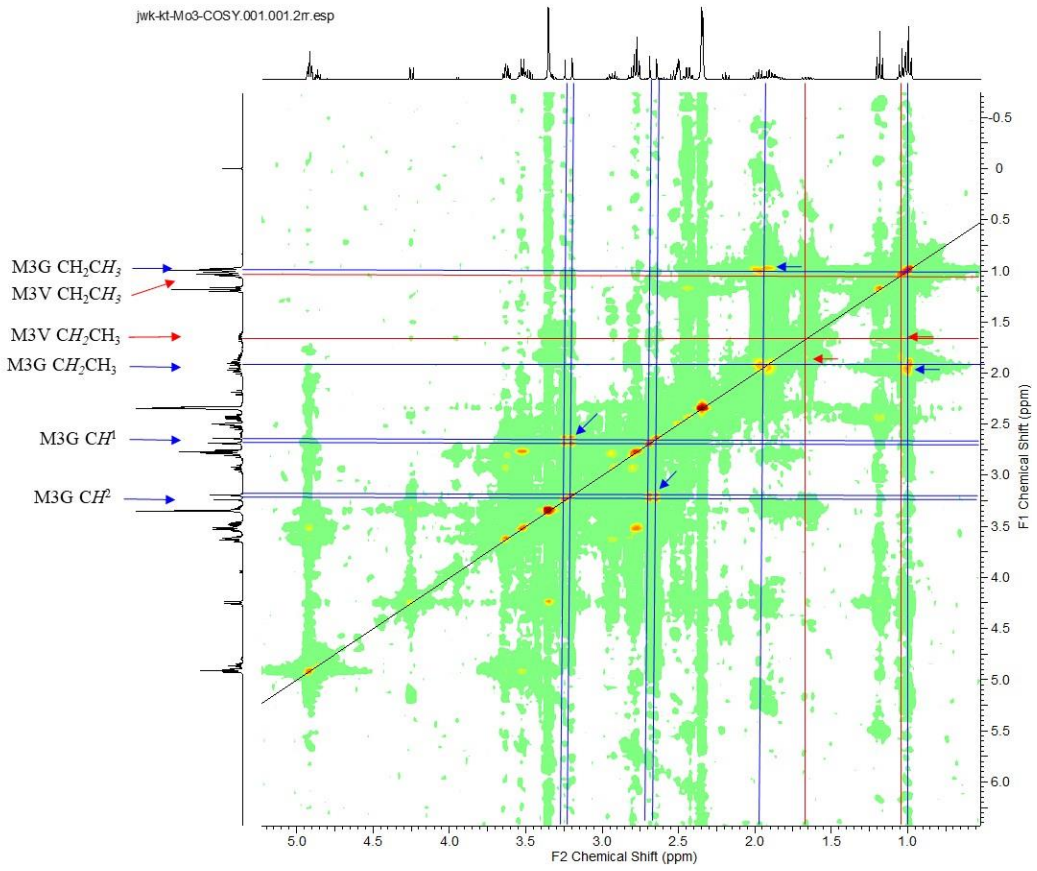


Figure 2.4. COSY NMR spectrum of the reaction between **M3** and BME. Used to identify protons belonging to the G and V products in order to determine % conversion and regioselectivity. Protons belonging to the G product are highlighted in blue and protons belonging to the V product are highlighted in red. Arrows highlight the correlations between protons.

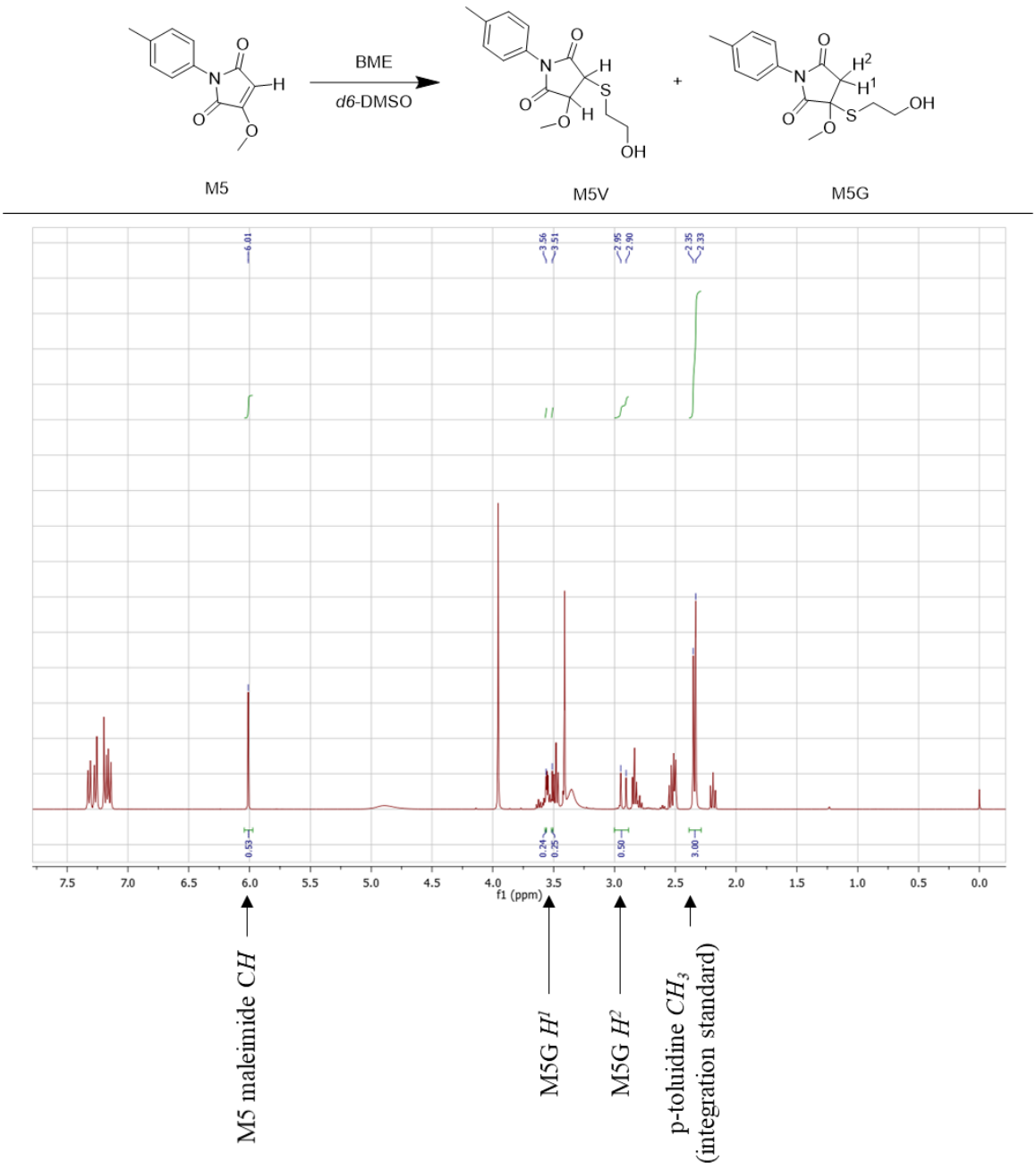


Figure 2.5. ^{1}H-NMR spectrum of the reaction between **M5** and BME (see Experimental). Protons used to determine % conversion and regioselectivity (V:G ratio) have been shown with arrows.

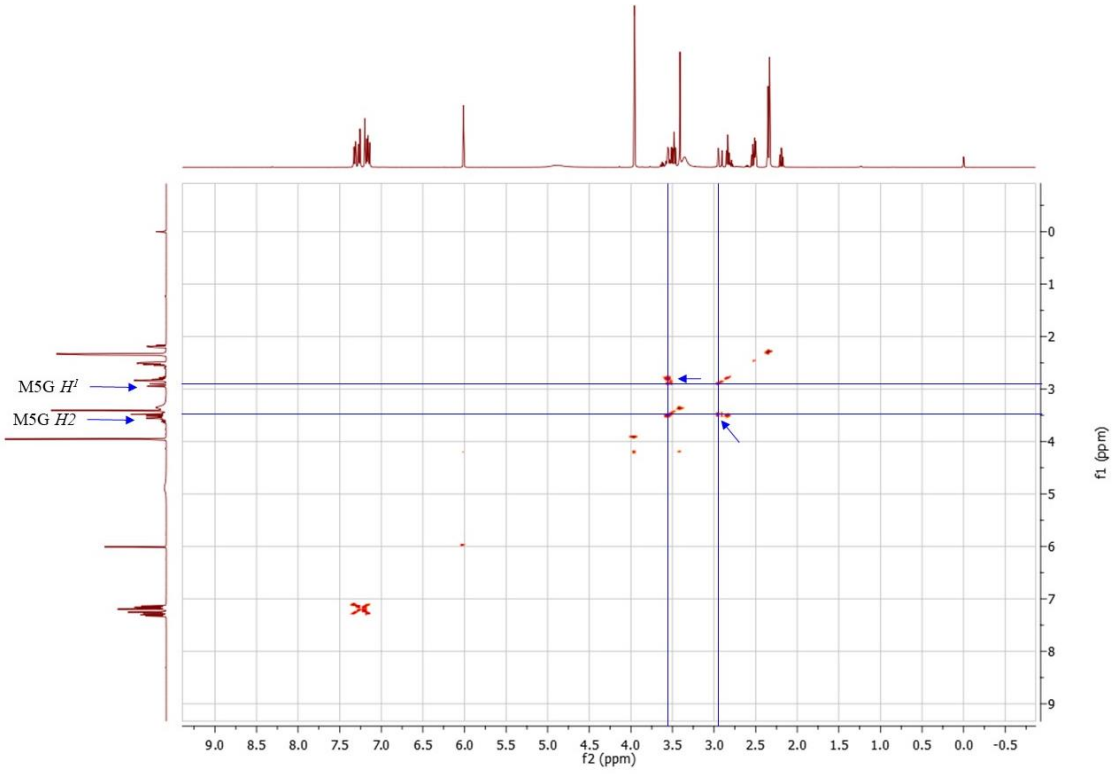


Figure 2.6. COSY NMR spectrum of the reaction between **M5** and BME. Used to identify protons belonging to the G and V products in order to determine % conversion and regioselectivity. Protons belonging to the G product are highlighted in blue. Arrows highlight the correlations between protons.

Lastly, we repeated the experiment with the methyl model compound **M2** in increasing amounts of D_2O in order to examine how solvent influences the regioselectivity of the thiol addition reaction. The V:G ratio was observed to shift towards increasing amounts of G product as the proportion of D_2O was raised (Table 2.2). This suggests that in our kinetic studies, which occur under conditions containing 95% water and 5% DMSO, the geminal addition product is probably formed in the majority. However, this could not be measured directly since the limited solubility of the model compound prevented us from attaining the necessary concentrations for NMR in conditions containing more than 20% D_2O in *d*6-DMSO.

Table 2.2. Solvent effect studies on the regioselectivity of BME addition to **M2** (150 mM model compound, 150 mM BME, *d6*-DMSO, 25 °C, overnight).[1]

D_2O %	Regioselectivity (V:G)
0	58:42
10	55:45
20	48:52

2.3.1.2. Hydrolysis

The model compounds shown above were also subjected to hydrolysis and similar product analysis. Briefly, the model compounds were hydrolyzed using 1% NaOH and the reaction mixture was then extracted prior to analyzing the product mixture by NMR. ^{1}H- and ^{13}C-NMR demonstrated that only one hydrolysis product was obtained from each reaction. To determine the regioselectivity of hydrolysis, products were analyzed by heteronuclear multiple bond correlation (HMBC), which gives the correlations between carbons and hydrogens that are two or three bonds apart. The more intense this correlation is, the more tightly they are coupled, due to the proximity and/or the rigidity of the intervening bonds. Thus, a strong correlation, as indicated by black arrows, was found between the methyl protons of **M2** (Figure 2.7) or the methylene protons of **M3** (Figure 2.8) and the carboxylate carbon of the maleamic acid product. These correlations demonstrated exclusive reactivity at the carbonyl adjacent to the substituted carbon. With regards to **M5**, the methoxy proton of the sole hydrolysis product correlated with the substituted alkene carbon (Figure 2.9), whose chemical shift of 160.3 ppm was consistent with an adjacent carboxylic acid rather than an adjacent anilide, which would have been expected to be at 170 ppm. Therefore, the hydrolysis of **M5** occurred with similar chemoselectivity.

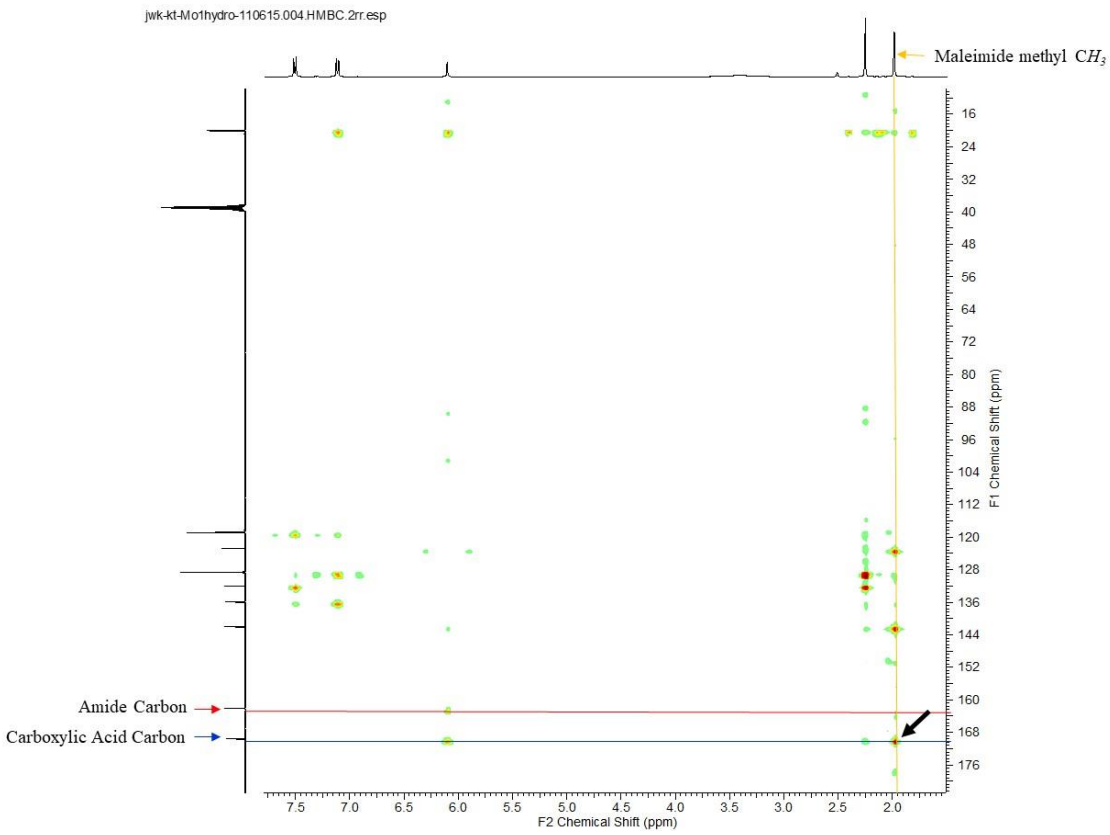


Figure 2.7. HMBC spectra of the hydrolysis product of **M2**.[1]

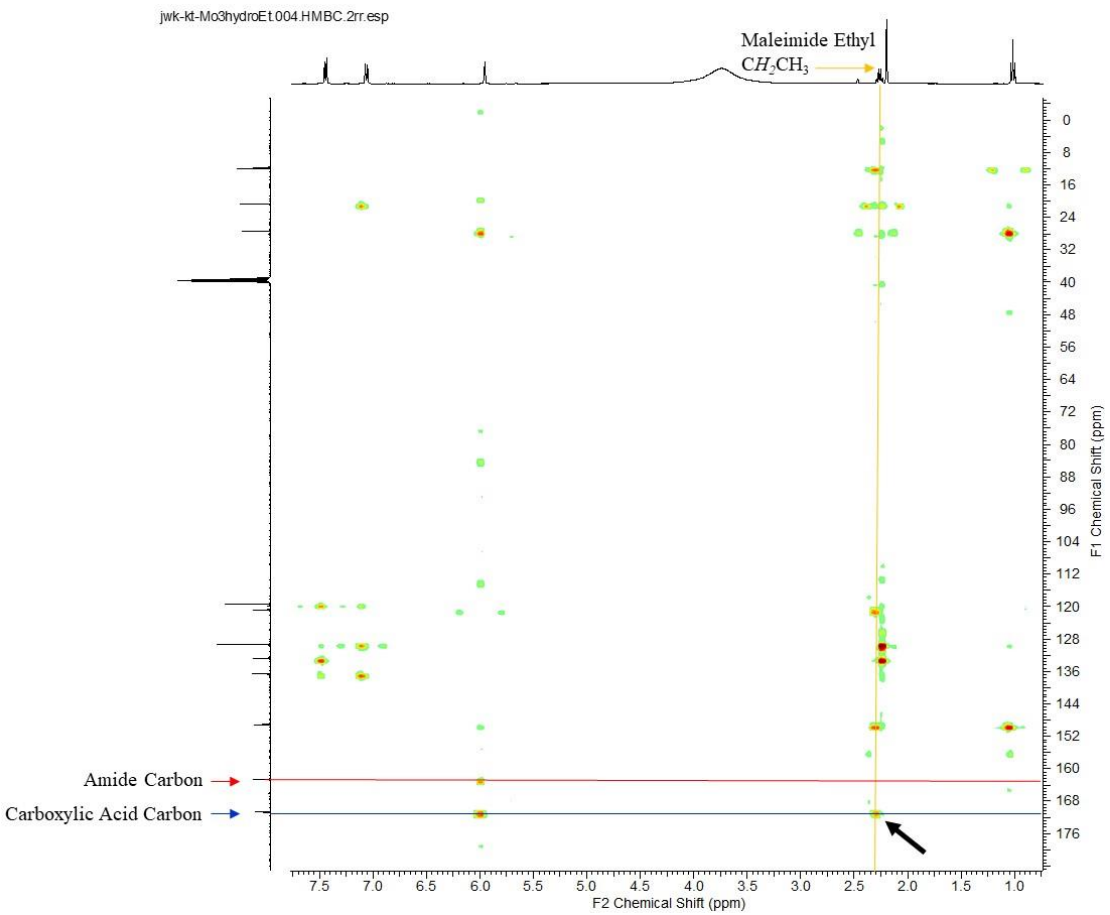


Figure 2.8. HMBC spectrum of the hydrolysis product of M3.[1]

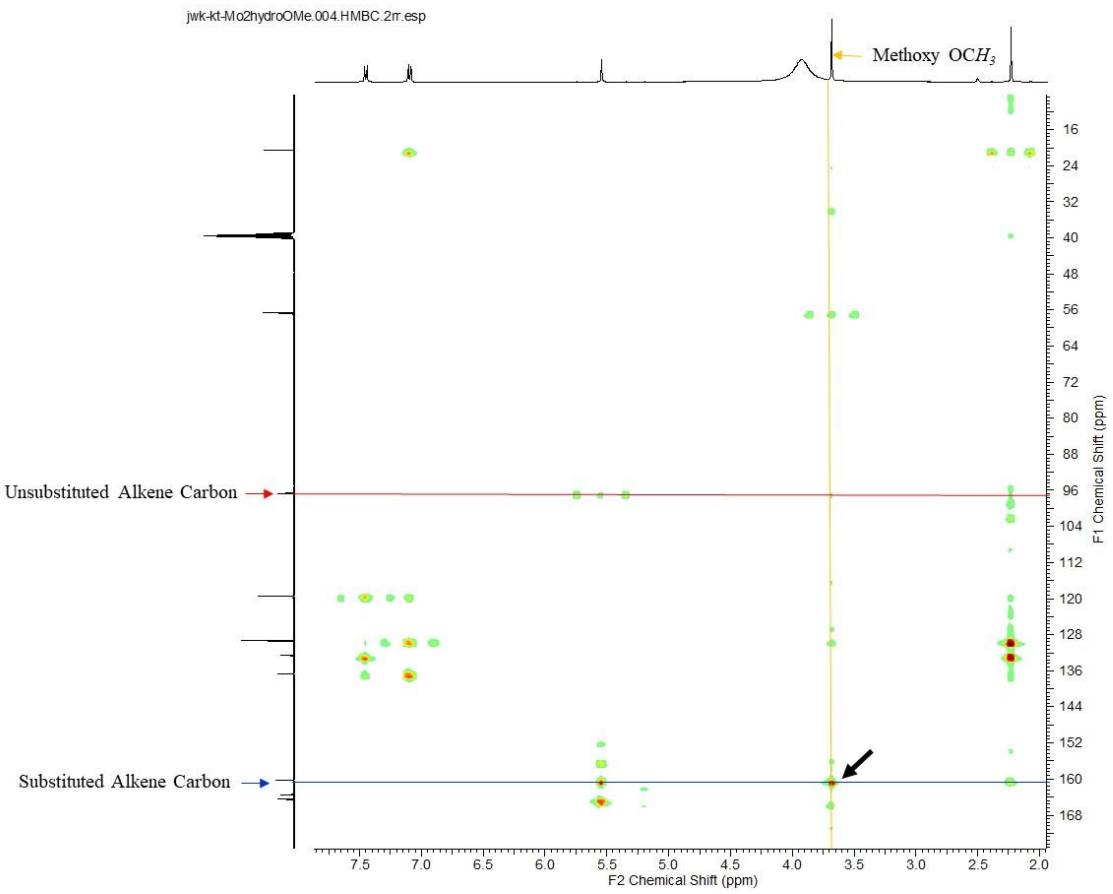


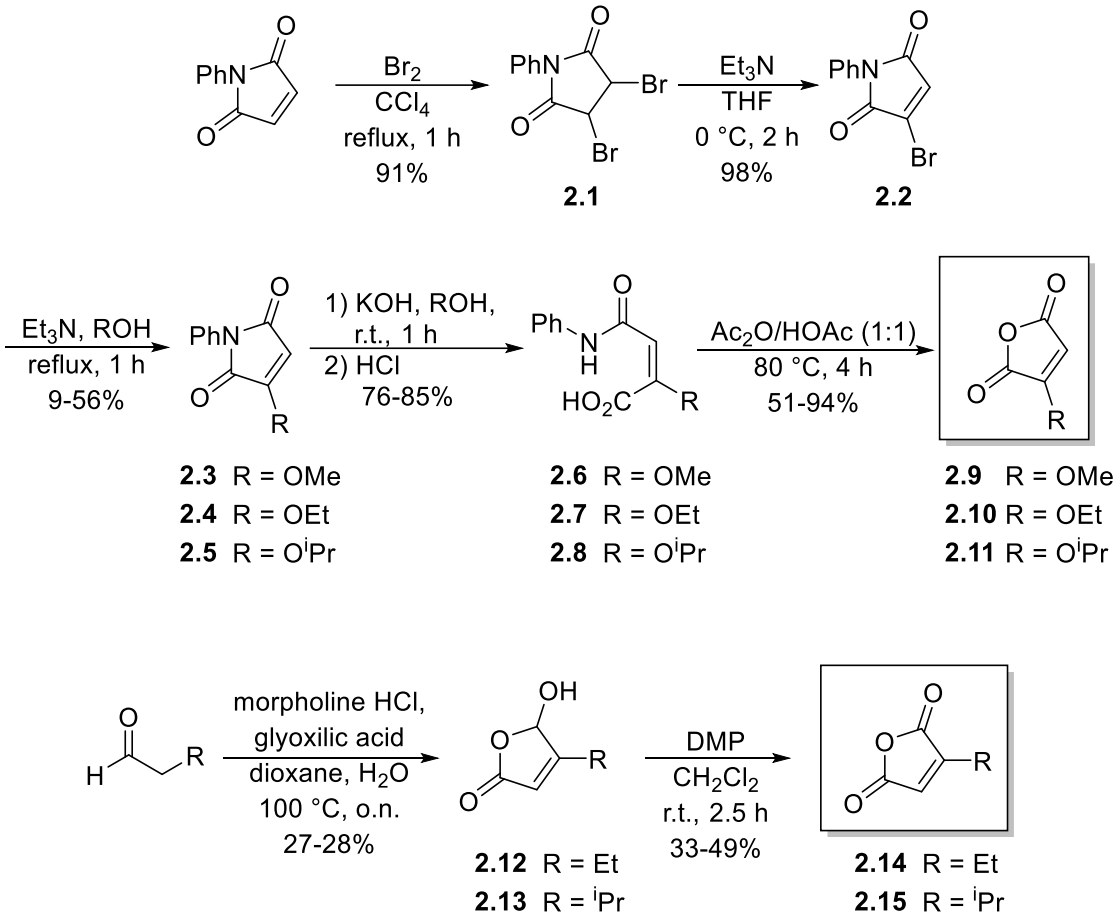
Figure 2.9. HMBC spectrum of the hydrolysis product of M5.[1]

2.3.2. Kinetic Studies

As mentioned above, a library of coumarin maleimide substrates (**S1-S9**) were synthesized in order to study the kinetics of maleimide thiol addition and hydrolysis. A convergent synthesis approach was taken to make each substrate, which consisted of three key pieces: 1) a 7-diethylaminocoumarin-3-carboxylate fluorophore; 2) a 1,3-diaminobenzene linker; 3) the appropriately functionalized maleic anhydride. The fluorophore fused to the aryl linker comprised a key amine intermediate to which a substituted maleic anhydride was added to form the final probe (Scheme 2.9, *vida infra*).

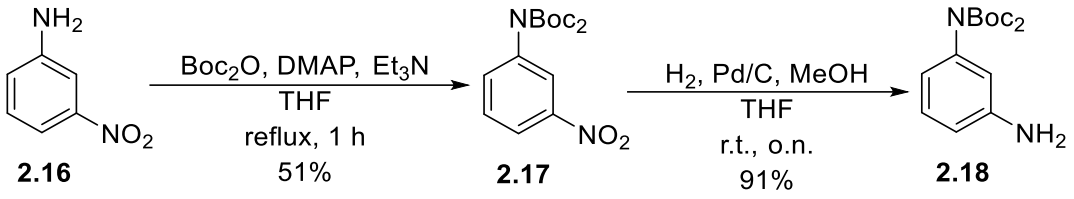
The substituted maleic anhydrides were either commercially available or prepared as shown in Scheme 2.6. Specifically, the alkoxy maleimides were synthesized by the direct

displacement of bromide from bromomaleimides, an approach that has been previously used in the synthesis of amino- and aryl substituted maleimides.[30,31] Thus, *N*-phenylmaleimide was dibrominated to obtain **2.1** from which HBr was eliminated to obtain monobromomaleimide **2.2** in excellent yield. RO^- was used to displace the bromide of **2.2** in low to moderate yield, giving substituted *N*-phenylmaleimides **2.3-2.5**, prior to obtaining *N*-phenylmaleamic acids **2.6-2.8** in very good yields through ring-opening hydrolysis. Maleic anhydrides **2.9-2.11** were obtained in moderate to excellent yields using a ring-closing reaction in which aniline was eliminated from the previous maleamic acids. To obtain the alkyl-substituted maleic anhydrides, an initial condensation of substituted aldehydes with glyoxylic acid gave lactones **2.12** and **2.13** in moderate yields. This was followed with oxidation by Dess-Martin periodinane (DMP) to obtain the final maleic anhydrides **2.14** and **2.15** in good yield.



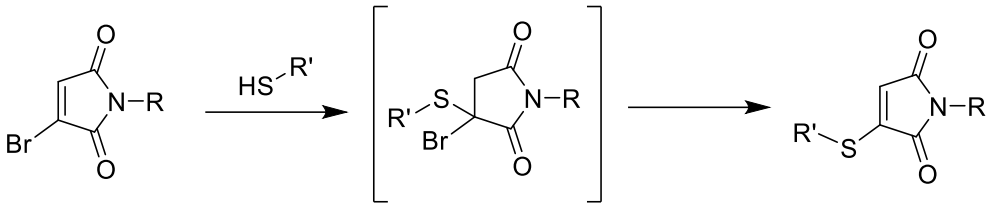
Scheme 2.6. Synthesis of substituted maleic anhydride intermediates.[1]

In order to prepare the key amine intermediate to which the maleic anhydrides would be added, the 1,3-diaminobenzene linker was first prepared asymmetrically and was coupled to the coumarin fluorophore in a controlled manner using protecting groups. Thus, *m*-nitroaniline **2.16** was first protected with two Boc groups in decent yield to give compound **2.17** and the nitro group was subsequently reduced using palladium-catalyzed hydrogenation to obtain the primary amine **2.18** in near quantitative yield (Scheme 2.7).

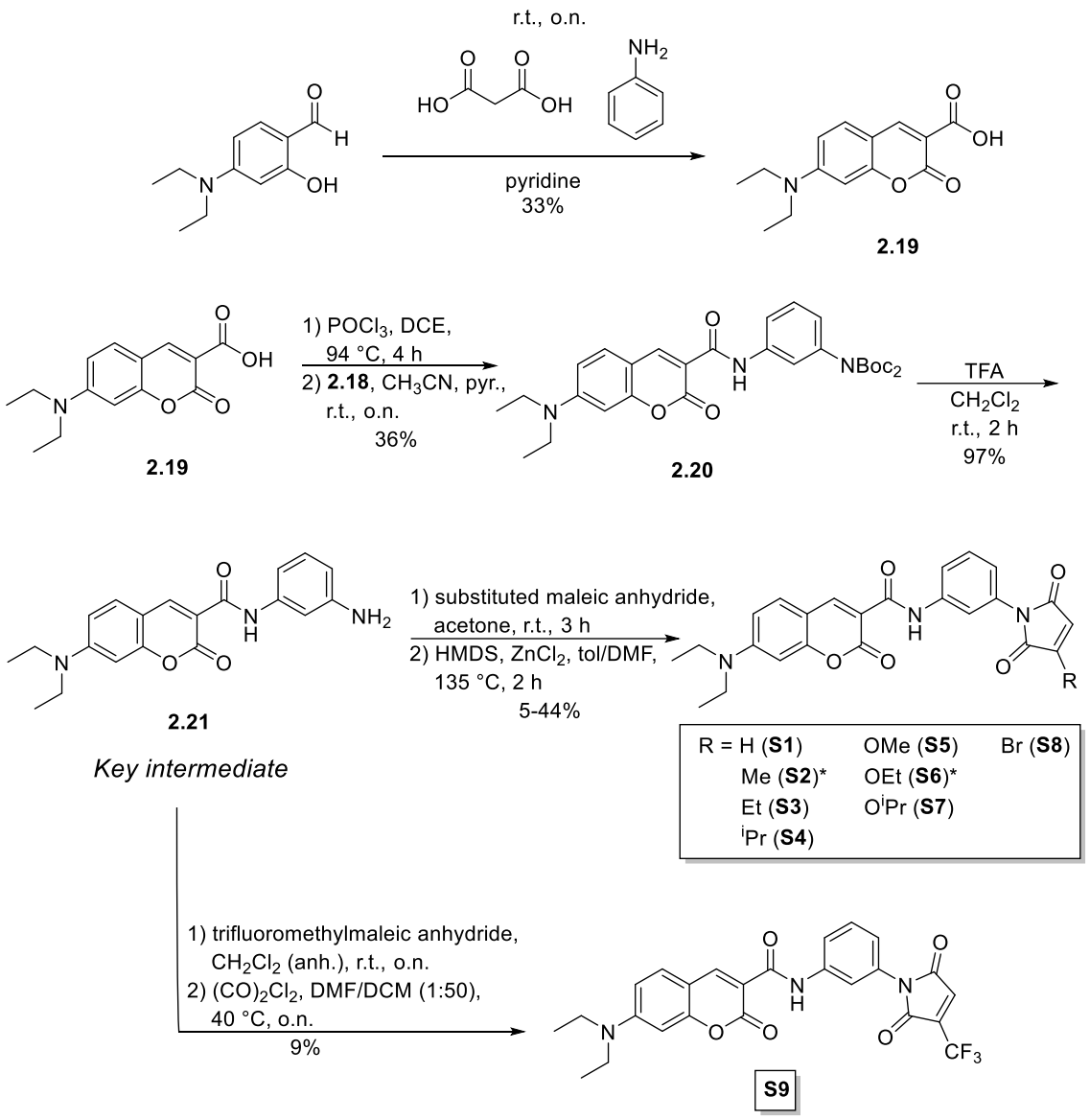


Scheme 2.7. Synthesis of the aryl linker intermediate.[1]

Coumarin **2.19** was synthesized from a previously established protocol using the Knoevenagel condensation and was then coupled with the aniline linker **2.18** by first converting **2.19**'s carboxylic acid to the acid chloride using $POCl_3$ (Scheme 2.9). The Boc groups on amide **2.20** were removed using TFA, giving key intermediate **2.21** in near quantitative yield. Coincidentally, an alternative strategy to synthesizing key intermediate **2.21** has also been reported in another study.[32] In our work, key intermediate **2.21** was reacted with the appropriate maleic anhydride to obtain fluorogenic substrates **S1**-**S9**. Substrate **S9** was prone to degradation, and so oxalyl chloride was used to convert the CF_3 substituted maleamic acid intermediate into the acyl chloride. Under anhydrous conditions, this was enough for ring closure to occur and form the resulting CF_3 substituted coumarin maleimide substrate, albeit in low yield due to the unstable nature of the compound. The synthesis of these substrates was not optimized as the absolute yields obtained were enough for the kinetic studies central to this study. **S8** was not studied further as previous studies have shown that thiol addition to bromomaleimides are followed by the subsequent elimination of the bromide, resulting in a thiol-substituted maleimide (Scheme 2.8).[33]



Scheme 2.8. Thiol addition-elimination reaction of bromomaleimides.[33]



Scheme 2.9. Synthesis of the final coumarin maleimide substrates.[1]

2.3.2.1. Hydrolysis

Given that FlARe probes are expected to react within cellular environments, it is important to examine the stability of maleimides in aqueous solutions since maleamic acids are generally unreactive towards thiol addition reactions.[34] *N*-phenylmaleimide was previously found to hydrolyze 25 times faster than *N*-ethylmaleimide at pH 7.0 and 30 °C.[4] Despite this, *N*-arylmaleimides are still stable enough for conjugation reactions at near neutral pH due to the much faster thiol addition that occurs.[13,16,35,36]

Our primary aim was to study the maleimide ring substituent's effect on the hydrolysis reaction by measuring the rate constants of hydrolysis for our fluorogenic substrate library between pH 5.0 and 10.5. The most electrophilic of these substrates, **S9**, hydrolyzed extremely rapidly under aqueous conditions, which is likely due to the electron-withdrawing nature of CF_3 group. The strong electron-withdrawing nature of CF_3 would also increase the rate of thiol addition to the maleimide, as demonstrated by McLeod's[37] report that an *N*-aryltrifluoromethylmaleimide reacts completely with *N*-Ac-Cys-OH in under 3 min in DMSO. Additionally, the CF_3 functional group would give rise to a particularly acidic thiol-addition product. As a result, the reverse of the Michael addition (ie. the E1cb elimination reaction) could occur to yield the starting trifluoromethylmaleimide. This effect is similarly observed with acrylamides functionalized with an electron-withdrawing cyano-group on the α-carbon.[38] Given this rapid degradation and potential reversibility, we decided that it would be highly impractical to examine the kinetics of **S9** in water. Thus, the substrate was not studied any further. Unlike **S9**, for substrates **S1-S7,** pseudo-first-order rate constants of hydrolysis, k_{obs}, were easily obtained. However, **S6** could not be studied below pH 7.7 due to its low solubility under assay conditions.

The resulting log k_{obs} vs pH plot makes three observations evident (Figure 2.10). First, each substrate's pH-rate profile appears to plateau between pH 5 and about 6. The pH independence within this range suggests a mechanism that involves the attack of neutral water on the maleimide during the hydrolysis reaction. Furthermore, varying buffer concentrations did not show significant changes in the measured rate constants, consistent with studies suggesting the absence of buffer catalysis.[39] Second, a slope of 1 was calculated for the basic portion of the pH-rate profile, indicating that 1 equivalent of hydroxide is involved in the reaction above pH 7. Lastly, each substrate's calculated second-order rate constant for hydroxide (k_{OH}) (see experimental) vary significantly from one another, demonstrating that the maleimide ring substituent likely influences the hydrolysis reaction.

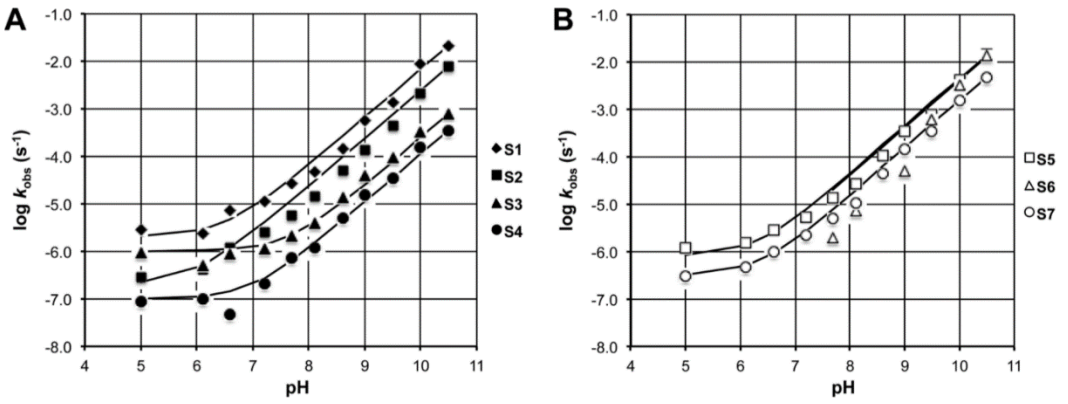


Figure 2.10. pH−rate profile for the hydrolysis of (A) substrates **S1−S4** and (B) substrates **S5−S7**. Reaction conditions: 25 μM coumarin *N*-arylmaleimide substrate, 50 mM aqueous buffer (I = 75 mM adjusted with KCl, pH 5-10.5), 5% DMSO, 22 °C. Solid lines are for fitted second-order rate constants, as described in the Experimental Section.[1]

The second-order rate constants calculated in this study are reported in Table 2.3. Regarding the unsubstituted maleimide, Salhany[13] has also shown a linear pH-rate profile from pH 7 to 10.2 at 25 °C for the hydrolysis of an unsubstituted maleimide linked to an eosin fluorophore through its ring nitrogen. Despite no reported second-order rate constant, half-lives deriving from the reported log k_{obs} values of -4.00 at pH 7.0 and -2.21 at pH 9.0 can be

roughly compared to the unsubstituted maleimide **S1** used in this study, which was found to be an order of magnitude less reactive (Figure 2.10). Upon examining substrates **S2-S7**, it is unsurprising that maleimide hydrolysis is sensitive to the maleimide ring's substituent. In fact, a study of six *para*-substituted *N*-phenylmaleimides by Kanaoka[4] clearly demonstrates this sensitivity. Furthermore, Kanaoka's reported half-lives at pH 8.0 translate to log k_{obs} values from -3.46 (*p*-NMe_2) to -2.77 (*p*-NO_2) at 30 °C, which once again show that **S1** reacts 1-2 orders of magnitude slower. Kanaoka *et al.* also calculated a Hammett ρ value of +0.35 for their library using the σ_{para} values for the substituents of their six substrates. This supports that nucleophilic attack on the maleimide carbonyl is rate limiting.[4]

Table 2.3. Second-order rate constants for maleimide hydrolysis (k_{OH}) and BME addition reactions (k_2) of *N*-arylmaleimides.[1] Reaction conditions for maleimide hydrolysis: 25 μM coumarin *N*-arylmaleimide substrate, 50 mM aqueous buffer (I = 75 mM adjusted with KCl, pH 5 – 10.5), 5% DMSO, 22 °C. Reaction conditions for BME addition reaction: 25 μM coumarin *N*-arylmaleimide substrate, 10-100X excess BME, 50 mM HEPES (I = 75 mM adjusted with KCl, pH 7.4), 5% DMSO. 37 °C,

Substrate	**R**	k_{OH} ($M^{-1}s^{-1}$)	k_2 ($M^{-1}s^{-1}$)	k_{OH} / k_2 Ratio
S1	H	67.2 ± 16.7	8.93 ± 0.73	7.55
S2	Me	24.4 ± 3.9	4.38 ± 0.16	5.55
S3	Et	2.47 ± 0.52	2.19 ± 0.18	1.07
S4	^{i}Pr	1.11 ± 0.28	1.32 ± 0.05	0.8
S5	OMe	44.5 ± 5.3	1.31 ± 0.05	40.5
S6	OEt	41.3 ± 7.2	N.D.[a]	-
S7	O^{i}Pr	15.4 ± 1.1	N.D.[b]	-

[a] Not determined due to **S6**'s insolubility in aqueous solution at pH 7.4, which were the conditions used for the thiol addition reaction. [b] Not determined because the hydrolysis of S7 dominates at pH 7.4, which were the conditions used for the thiol addition reaction (*vida infra*).

A study by Matsui[10] reported k_{OH} values for a series of maleimides bearing a H or alkyl group on the maleimide nitrogen. Notably, *N*-ethylmaleimide (NEM) had k_{OH} values of 34.7 $M^{-1}s^{-1}$ at 20 °C and 46.8 $M^{-1}s^{-1}$ at 30 °C, which are comparable to that of **S1** (Table 2.3). Additionally, the pH-rate plots reported by Matsui show a log k_{obs} value of -5 for NEM at pH

7 and 30 °C, like the log k_{obs} value measured for **S1** (Figure 2.10). Thus, these comparisons suggest that not all *N*-arylmaleimides hydrolyze faster than all *N*-alkylmaleimides. Lastly, Matsui plotted the log k_{OH} values against the Taft polar parameter σ* to obtain a ρ* value of 0.55 (30 °C) for their library. Like the studies by Kanaoka, this also supports the idea that the attack by hydroxide is the rate limiting step in the hydrolysis of their *N*-alkylmaleimides. Caution must be taken in making these comparisons as the hydrolysis rate constants reported herein were obtained at 22 °C. However, the values reported by Salhany, Kanaoka, and Matsui were obtained at higher temperatures, which will increase the rate constant of the reaction as suggested by the Arrhenius Equation. Therefore, given the differences in structure of the *N*-substituent of the studied maleimides and the differences in reaction temperature, one cannot draw concrete conclusions from the above comparisons.

2.3.2.2. Thiol Addition

To increase the relevance of these kinetic studies to cellular labelling applications, the thiol addition reactions involving substrates **S1-S7** were done at pH 7.4 and at 37 °C. Additionally, BME was used as a model thiol given its aqueous solubility and structural similarity to biological thiols. The reactions were performed under pseudo-first-order conditions in which 25 μM of substrate was reacted with 10-100 equivalents of BME. The pseudo-first-order rate constants were then calculated from the monoexponential increase in fluorescence (see Experimental). The second-order rate constants (k_2) were then measured from the plots of k_{obs} vs [BME]. However, the limited solubility of **S6** under assay conditions prevented the measurement of k_2, and **S7** reacted very slowly, making it indiscernible from background hydrolysis. Despite this, the thiol addition reactions for **S1-S5** at the concentrations studied, albeit not their rate constants, were typically >100-fold faster than

their background hydrolysis reactions, allowing for second-order rate constants to be easily measured. These second-order rate constants are shown in Table 2.3.

In a study by Bednar,[11] the reactivity of neutral BME with NEM was shown to be insignificant, but the thiolate of BME reacted with an intrinsic rate constant of 1.8×10^5 $M^{-1}s^{-1}$ at 25 °C. Given that the pK_a of BME is 9.61,[40] the proportion of BME thiolate at pH 7.4 can be calculated and in doing so, the predicted rate constant for the reaction of BME with NEM would be 1.2×10^3 $M^{-1}s^{-1}$. This value is lower than the value of 5.27×10^3 $M^{-1}s^{-1}$ reported by Keillor[39] for the reaction of BME with *N*-methylmaleimide at pH 7.5 (25 °C) and the value of 9.00×10^3 $M^{-1}s^{-1}$ for the reaction of cysteine with *N*-eosin-maleimide at pH 7.4 (25 °C).[13] However, it is more in line to compare the value of 1.1×10^3 $M^{-1}s^{-1}$ that Keillor previously reported for the reaction of GSH with a bis-*N*-arylmaleimide fluorogen at pH 7.5 and 20 °C.[35] Here, the unsubstituted maleimide **S1** reacted the fastest. Despite this, the rate constant for the reaction between **S1** and BME (8.9 $M^{-1}s^{-1}$, Table 2.3) is still 2 orders of magnitude lower than the rate constants reported in the above studies. However, a direct comparison is impossible given the structural differences between the thiols and the maleimide nitrogen substituents across these studies. This suggests that that thiol addition to maleimides are influenced by more than just the alkene substituent.

Upon examination of the rate constants from **S1** to **S5**, the k_2 rate constants differ by almost an order of magnitude, suggesting a significant substituent effect for the thiol addition reaction. However, this sensitivity is less pronounced than what was observed for hydrolysis given that the k_{OH} values differed by almost 2 orders of magnitude. Additionally, a different effect on hydrolysis and thiol addition is observed depending on the substituent, which is evident from the k_{OH}/k_2 ratio reported in Table 2.3. If the substituents had the same effects on

both reactions, the ratio of rate constants should be constant or vary smoothly. On the contrary, the methoxy substituent of **S5** is shown to decrease its k_2 value relative to the unsubstituted **S1** but demonstrates much less effect on the k_{OH} value relative to **S1**. Therefore, the dependence on the subsitutents may vary between the mechanisms for maleimide hydrolysis and thiol addition.

2.3.3. Linear Free Energy Relationships

The hydrolysis rate constants in Table 2.3 for substrates **S1-S4** reveals that the steric bulk of the alkyl substituents heavily influence hydrolytic reactivity. Furthermore, the k_{OH} values for **S5** to **S7** may indicate that the alkoxy substituents promote hydrolysis relative to the alkyl substituents. Despite this, it is difficult to discern the electronic from the steric effects of the substituents, which suggests that a linear free energy relationship (LFER) involving these rate data likely includes both steric and electronic parameters for the ring substituents. This is likely also the case for the substituent effects on the thiol addition rate constants even though these rate constants show a different trend than observed for the hydrolysis rate constants. Therefore, appropriate substituent parameters must first be established before in-depth mechanistic analysis.

As mentioned earlier, the similarities between maleimide thiol addition and hydrolysis and the ester hydrolysis reactions used to establish the Hammett σ^+ and Taft substituent constants allow these constants to be used to measure the electronic and steric effects of substituents for substrates **S1-S9**. Despite this, a poor correlation was observed between the measured rate constants and the isolated electronic substituent effect upon plotting the log k_{OH} and log k_2 values against σ^+ alone (Figure 2.11). Density maps and LUMO energy levels for substrates **S1-S9** obtained using DFT calculations (see Table 2.7 in Appendix) show a good

correlation with σ^+ (Figure 2.12), which supports that the σ^+ parameter is appropriate for measuring electronic effects like previous studies.[41] However, given the proximity of the maleimide ring's substituent to the site of nucleophilic attack, we reasoned that both steric and electronic effects influenced the reaction rates and had to be considered together. To do this, steric parameter values were either calculated and then normalized to H or taken directly from literature prior to the two-parameter fitting of log k_{OH} and log k_2 values (see Table 2.4 and experimental).[42]

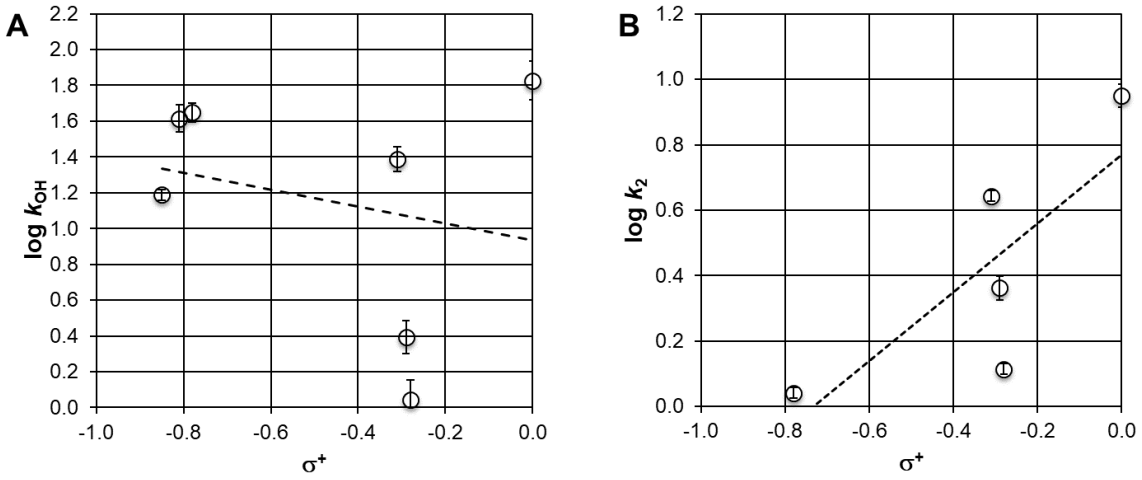


Figure 2.11. A) The observed poor correlation between log k_{OH} and Hammett σ^+ parameter for the maleimide ring substituent on substrates **S1-S7**. R^2 = 0.0539. B) The observed poor correlation between log k_2 and Hammett σ^+ parameter of maleimide ring substituent for substrates **S1-S5**. R^2 = 0.6113.

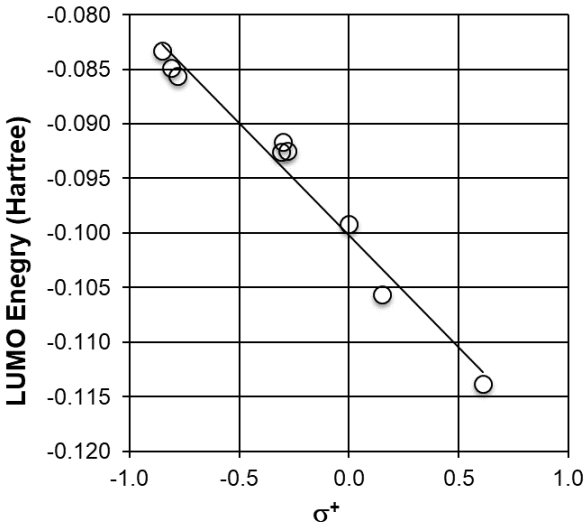


Figure 2.12. Correlation between LUMO energy level (as calculated by DFT) of substrates **S1-S9** and Hammett σ+ parameter for the corresponding maleimide ring substituent. $R^2 = 0.9724$. σ^+ (Br) = 0.15 and σ^+ (CF_3) = 0.61.[5] LUMO energy levels can be found in **Table 2.7** in the Appendix.

Table 2.4. Steric (E_s^{norm}) and electronic (σ^+) parameter values used for the substituents (R) of substrates **S1-S7**.

Substrate	R	σ^+ [a]	E_s^{norm} [b]
S1	H	0.00	0.00
S2	Me	-0.31	-1.24
S3	Et	-0.29	-1.31
S4	^{i}Pr	-0.28	-1.71
S5	OMe	-0.78	-0.25
S6	OEt	-0.81	-0.34
S7	O^{i}Pr	-0.85	-0.82 [c]

[a] Taken from reference [5]. [b] Taken from reference 42. [c] Calculated from extrapolated values (See Experiment, section 2.4.5).

Fitting the log k_{OH} of substrates **S1-S7** using two-parameter fitting yielded a steric reaction constant of $\delta = 0.9 \pm 0.2$ and an electronic reaction constant of $\rho = -0.1 \pm 0.5$, resulting in the plane shown in Figure 2.13A and the correlation of observed predicted values in Figure 2.13B. Compared to the correlation obtained for fitting only the electronic parameters (Figure 2.11), this correlation is much better. Despite only accounting for ~80% of the measured data

points, it is evident that the ring substituent effect on maleimide hydrolysis is primarily steric and minimally electronic. If hydroxide attack is rate limiting, as supported by other maleimide hydrolysis LFER studies,[4,10] the large δ value indicates that the -OH nucleophilic attack on at least one imide carbonyl is sterically hindered by the substituent. This may be due to the nucleophile's trajectory on the delocalized π* system. As shown by our product studies, it is the imide carbonyl adjacent to the substituted carbon that undergoes hydrolysis, which may be the most sensitive to the substituent's steric bulk. This preference may be because formation of a tetrahedral intermediate during nucleophilic attack relieves the steric strain between the carbonyl and the adjacent coplanar substituent. Contrary to this, nucleophilic attack on the carbonyl distal from the substituent to form the corresponding tetrahedral intermediate would not benefit from this steric relief of ground state destabilization. Lastly, the small ρ value measured for maleimide hydrolysis suggests that hydroxide may be insensitive to the ring substituent's electronic nature given its strong nucleophilic nature.

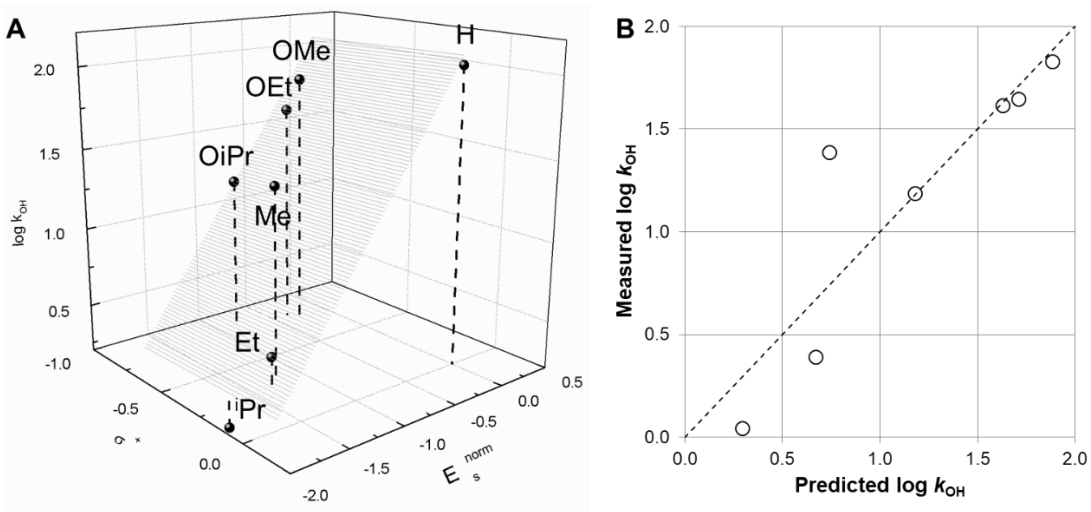


Figure 2.13. A) Two-parameter fitting of log k_{OH} values against both electronic (σ^+) and steric (E_s^{norm}) substituent parameters with the fitted plane shown in grey and yielding respective reaction constants of $\rho = -0.1 \pm 0.5$ and $\delta = 0.9 \pm 0.2$. B) Correlation of measured log k_{OH} values against values predicted from fitting against both electronic (σ^+) and steric (E_s^{norm}) substituent parameters. The dashed line is shown as a reference and indicates theoretical perfect correlation. $R^2 = 0.7956$.

Two-parameter fitting of the rate data from the thiol addition reaction of substrates **S1-S5** yielded a steric reaction constant $\delta = 0.2 \pm 0.4$ and an electronic reaction constant $\rho = 1.1 \pm 0.4$. These constants define the plane shown in Figure 2.14A with the correlation of predicted to observed rate constants shown in Figure 2.14B. Unlike maleimide hydrolysis, the ρ and δ constants suggest that maleimide thiol addition is minimally sensitive to the ring substituent's steric characteristics and much more sensitive to its electronic effects. The reduced steric sensitivity of thiolate addition can be explained by the S^- nucleophile's greater polarizability compared to the smaller O^- nucleophile involved in hydrolysis. Additionally, the thiolate addition's larger electronic sensitivity might be because the substituent is directly conjugated with the reactive C=C bond in the maleimide ring (Scheme 2.5).

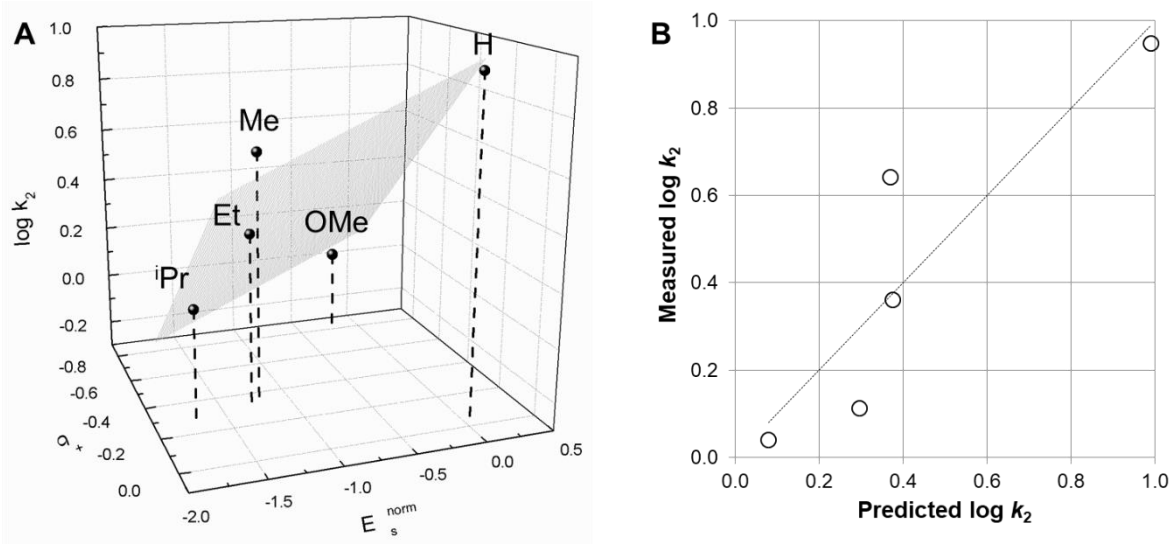


Figure 2.14 A) Two-parameter fitting of log k_2 values against both electronic (σ^+) and steric (E_s^{norm}) substituent parameters with the fitted plane shown in grey and yielding respective reaction constants of $\rho = 1.1 \pm 0.4$ and $\delta = 0.2 \pm 0.2$. B) Correlation of measured log k_2 values against values predicted from fitting against both electronic (σ^+) and steric (E_s^{norm}) substituent parameters. The dashed line is shown as a reference and indicates theoretical perfect correlation.

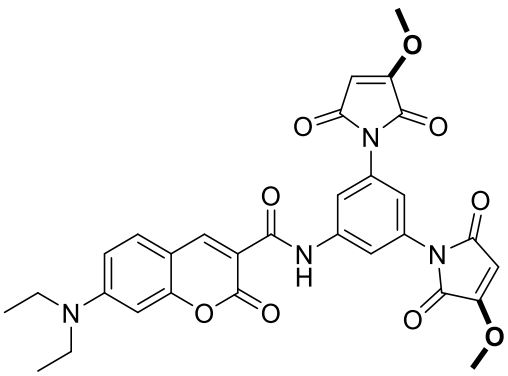
Nevertheless, the rough correlation shown in Figure 2.14 and the fact that the thiol addition reaction forms two regioisomers cautions against overinterpreting the data. The kinetic data reported here are based on the appearance of total product. Therefore, it only

represents the sum of the rate constants for the formation of both regioisomers. However, the rate constants associated with the formation of each of these regioisomers would likely be affected differently by each maleimide ring substituent. In order to correct the observed second-order rate constants to obtain relative rate constants for geminal and vicinal addition, knowledge of the ratio of these regioisomers under assay conditions is required. This could not be obtained due to the solubility issues encountered during the product studies. The uncertainty raised by this regioselectivity contributes to the approximation in using only one pair of substituent parameters to study substrates that potentially form two regioisomers. Despite this, we can still qualitatively conclude that the substituent's electronic effects have a greater influence on maleimide thiol addition than its steric effects given the solvent-dependent regioselectivity that we were able to measure.

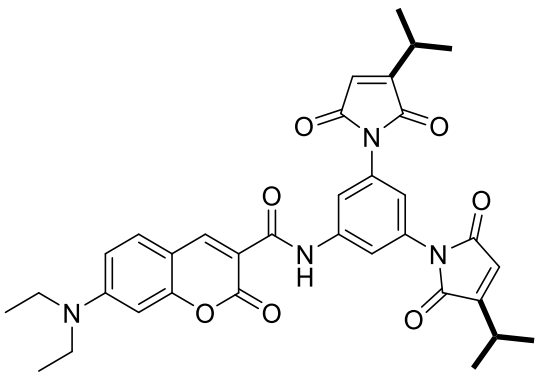
2.3.4. Towards improving current FlARe probes

The current FlARe probe design uses two methoxymaleimides, which reacts minimally with GSH and makes the probe selective for the genetically fused dC10 tag. This attenuation against adventitious thiols is supported by the low thiol addition rate constant for the methoxy substituted maleimide and BME found in this study (**S5**, 1.3 $M^{-1}s^{-1}$). Notably, the isopropyl substituted maleimide, **S4**, has a similar thiol addition rate constant but a much lower hydrolysis rate constant (1.1 $M^{-1}s^{-1}$). This is exemplified by the much lower k_{OH}/k_2 = 0.8 ratio reported for **S4** as compared to the k_{OH}/k_2 = 40.5 observed for **S5**. Therefore, we hypothesized that a di-isopropylmaleimide FlARe probe, **KT12**, would be more stable against hydrolysis than **YC20**, the current di-methoxymaleimide coumarin based FlARe design, while sharing the same selectivity for the dC10 tag (Figure 2.15). Even though maleimide thiol addition is much faster than maleimide hydrolysis, improving probe stability is still beneficial

as the label must reach the target within the cell. Maleimides that decompose to the maleamic acid during this transport process will both contribute to background fluorescence and not react with the target peptide, decreasing the amount of protein labelled.



YC20

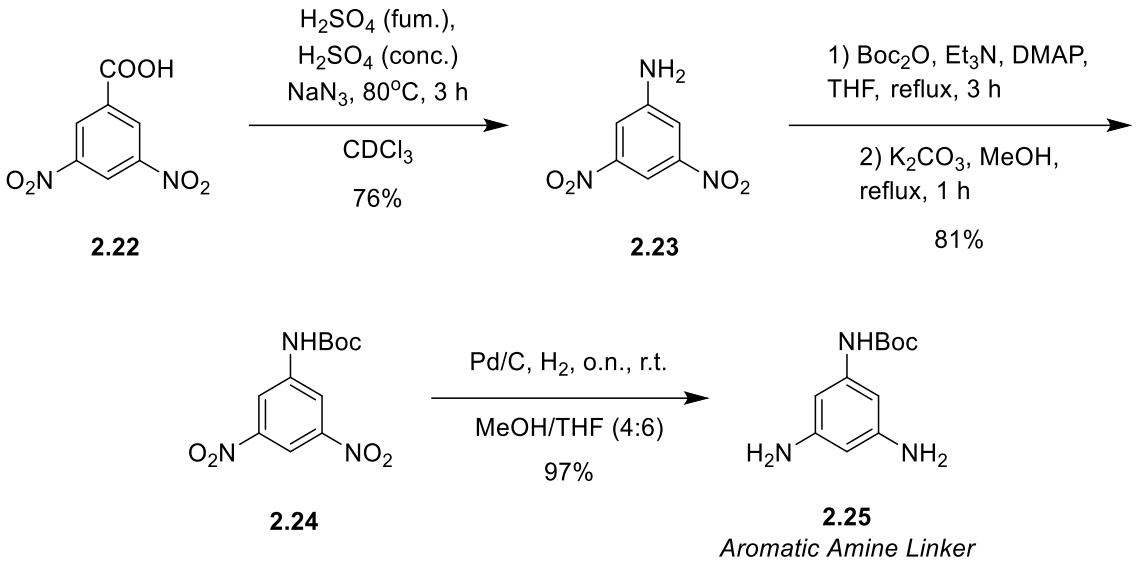


KT12

Figure 2.15. Structures of **YC20** and its isopropylmaleimide variant **KT12**.

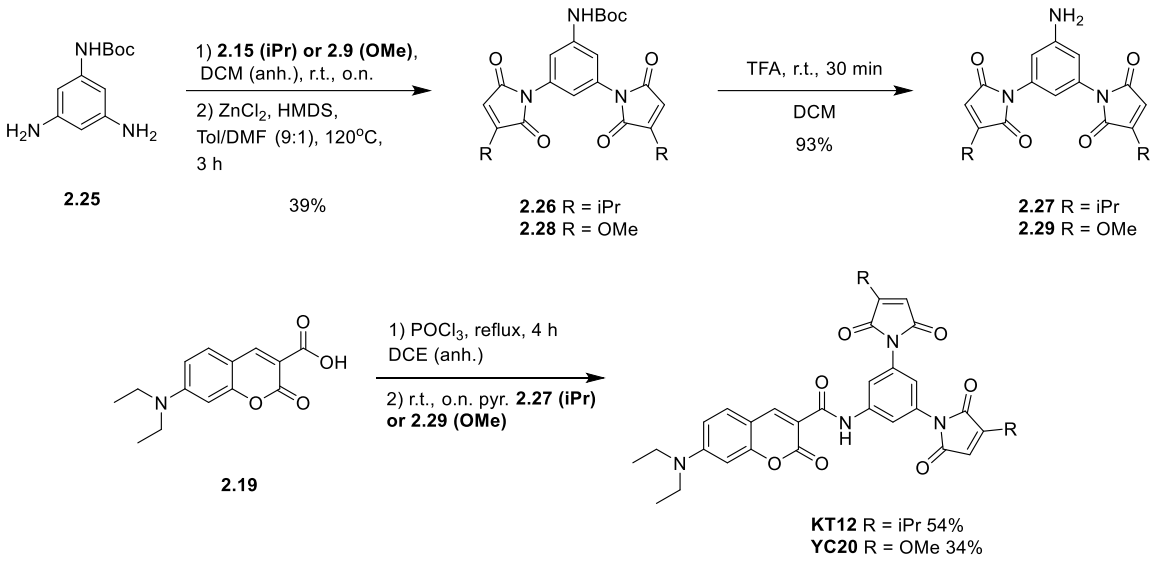
2.3.4.1. Synthesis

A convergent strategy was taken to synthesize **KT12** by assembling together isopropylmaleic anhydride (compound **2.15**), 7-*N*-diethylaminocoumarin-3-carbxylic acid (compound **2.19**), and an aromatic diamine scaffold. The synthesis of **2.15** and **2.19** have already been outlined (*vida supra*) while the aromatic diamine linker was synthesized from methods outlined in previous studies with minor optimizations.[2] Specifically, a Schmidt reaction was first performed on 3,5-dinitrobenzoic acid to obtain 3,5-dinitroaniline in good yield (Scheme 2.10). The amine of **2.23** was then protected with a single Boc group after first reacting **2.23** with excessive amounts of Boc_2O, followed by the selective removal of one Boc group using K_2CO_3 to obtain **2.24**, also in good yield. Finally, **2.24** was subjected to palladium-catalyzed hydrogenation to obtain the aromatic diamine linker **2.25** in near quantitative yield.



Scheme 2.10. Synthesis of aromatic amine linker 2.25.

The strategy used to synthesize **YC20** was adopted to make the new di-isopropylmaleimide FlARe probe (Scheme 2.11). Thus, the aromatic diamine linker **2.25** was first reacted with isopropylmaleic anhydride, **2.15** and the maleimide ring closure was subsequently performed with $ZnCl_2$ and HMDS in decent yield. The aniline's Boc group was removed using TFA to obtain the key 1,3-di(isopropylmaleimide) aniline intermediate **2.27** in near quantitative yield. Finally, **2.19** was converted to the acid chloride using $POCl_3$, which was then reacted with **2.27** to form the final FlARe probe, **KT12**, in good yield.



Scheme 2.11. Synthesis of the final KT12 FlARe probe.

2.3.4.2. Properties of KT12

Precipitation was observed during the initial kinetic experiments with **KT12** but not with **YC20**. This is problematic as precipitation influences the amount of measured fluorescence due to scattering and potentially aggregation-caused quenching.[7] This observation suggested that **YC20** is more soluble than **KT12**. Thus, for the purposes of comparing both FlARe probes, the DMSO concentration in the assays was raised from 5% to 20%.

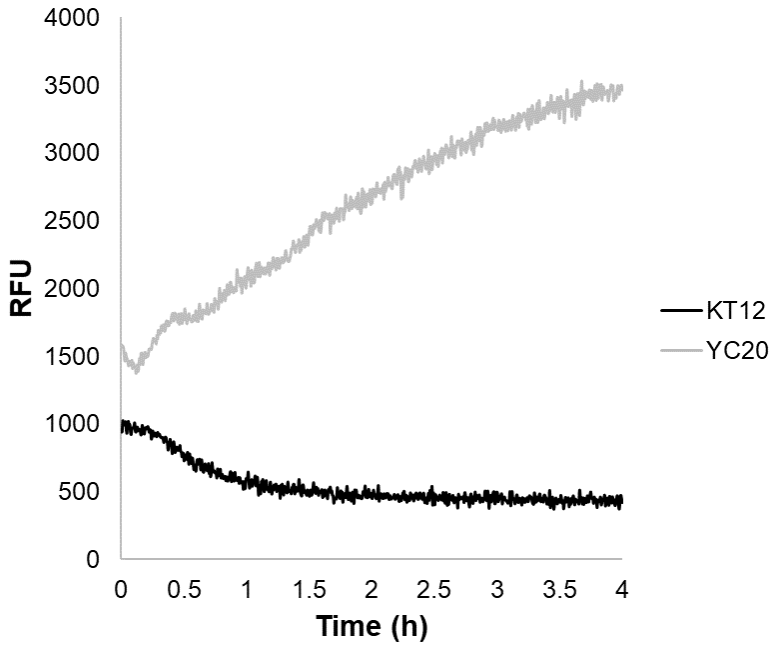


Figure 2.16. Stability of 25 μM **KT12** (black) or **YC20** (grey) in the absence of thiol. Reaction conditions: 125 μM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), 20% DMSO, 37 °C.

Using the new assay conditions, **KT12** and **YC20** was first evaluated for its stability against hydrolysis by monitoring the fluorescence increase over time in the absence of any thiol. When evaluating FlARe probes against GSH, a five-fold excess of tris(2-carboxyethyl)phosphine (TCEP) is used to ensure the complete reduction of GSH in solution. To maintain the same conditions, 125 μM TCEP was also used when evaluating the stability of **KT12** and **YC20** despite the absence of GSH. Figure 2.16 shows that **KT12** exhibits no fluorescence increase over 4 h, while **YC20** showed about a two-fold increase within that time. Thus, **KT12** seems to be more stable in aqueous solutions than **YC20**, which is consistent with the above substituent study suggesting that the isopropylmaleimide hydrolyzes slower than the methoxymaleimide.

Next, we examined **KT12's** reactivity with GSH by assessing its reactivity with excessive concentrations of the tripeptide. Figure 2.17 shows that at cellular concentrations of

GSH, ranging from 5 – 10 mM,[43] **KT12** shows substantial reactivity as even at GSH concentrations as low as 1 mM, **KT12** was found to react. Given the large excess of GSH compared to **KT12**, pseudo-first order rate constants were measured (Figure 2.17) using the method of initial rates (See Experimental, section 2.5.4.2). From this, a second order rate constant of $k_2 = 1.72 \pm 0.05 \times 10^{-2}$ $M^{-1}s^{-1}$ was calculated for the thiol addition of GSH to **KT12**. This evaluation was repeated for **YC20** and a $k_2 = 4.9 \pm 0.3 \times 10^{-2}$ $M^{-1}s^{-1}$ was obtained, demonstrating an almost three times increase in reactivity with GSH compared to **KT12** (Figure 2.17). Previous studies studying the reaction of GSH with the unsubstituted maleimide variant of **YC20** at pH 7.5 and 20 °C report a much higher rate constant of 1.1×10^3 $M^{-1}s^{-1}$.[35] The low rate constants for both **YC20** and **KT12** compared to the unsubstituted maleimide FlARe probe mirrors the low rate constants for the reaction between **S4** (^{i}Pr) and **S5** (OMe) and BME compared to the unsubstituted **S1** substrate. The similar rate constants for the methoxy **S5** and isopropyl **S4** maleimide substrates with BME also supports the small difference in rate constants between **KT12** and **YC20**, regarding their reaction with GSH.

Although numerous similarities can be drawn, it is also instructive to evaluate the differences that arise between the thiol addition of GSH to the various dimaleimide FlARe probes and the thiol addition of BME to the monomaleimide substrates studied above. For example, the rate constants for the reaction of **KT12** or **YC20** with GSH are two orders of magnitude smaller than the rate constants of BME reacting with **S4** or **S5**, respectively. However, the unsubstituted maleimide FlARe probe is estimated to react at least three orders of magnitude faster than the addition of BME to **S1**. Furthermore, **YC20** was still found to react three times faster than **KT12**, despite the negligible difference in reactivity with BME for **S5** and **S4**. This suggests that the difference in the small molecule thiol and the addition

of a second maleimide may amplify the relative reactivities observed for the thiol addition of BME to monomaleimide coumarin substrates.

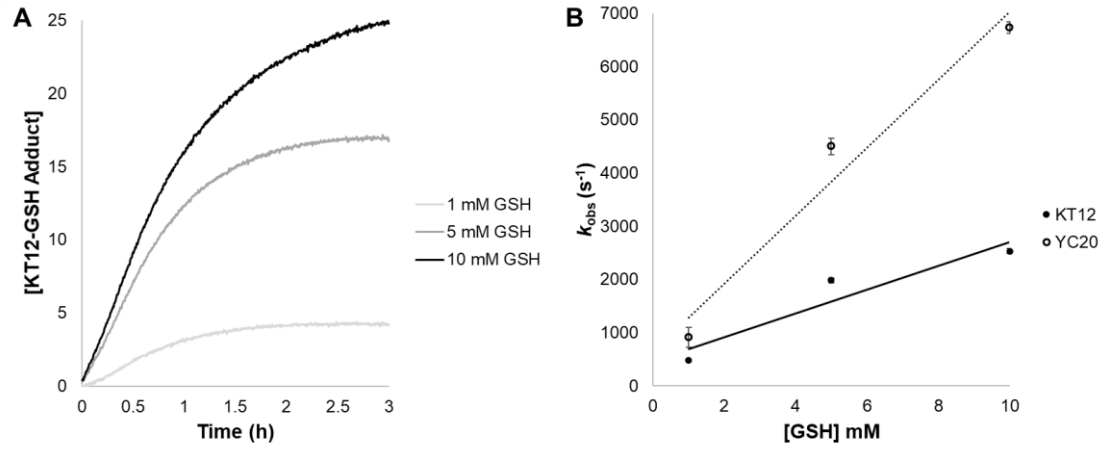
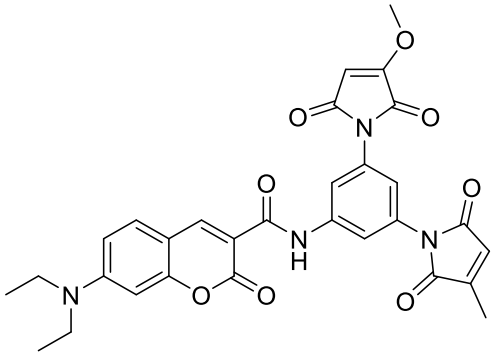


Figure 2.17. A) Kinetics of the reaction between 25 µM **KT12** and 1-10 mM GSH under pseudo-first-order rate conditions. Reaction conditions: 125 µM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), 20% DMSO, 37 °C. B) Plot of k_{obs} vs. [GSH] for **KT12** and **YC20** with lines of best fit for **KT12** (solid line) and **YC20** (dotted line). Pseudo-first-order rate constants, k_{obs}, were determined using method initial rates from (A). **KT12** $k_2 = 1.72 \pm 0.05 \times 10^{-2}$ $M^{-1}s^{-1}$, $R^2 = 0.8972$. **YC20** $k_2 = 4.9 \pm 0.3 \times 10^{-2}$ $M^{-1}s^{-1}$, $R^2 = 0.9615$.

Lastly, 25 µM **KT12** and **YC20** were incubated with equimolar amounts of MBP-dC10* in a side-by-side comparison of *in vitro* protein labelling (Figure 2.19). The dC10* tag is an improved variant of the dC10 target peptide that was developed by Miroslava Strmiskova and further explained in Chapter 3. A $k_2 = 13.7 \pm 0.7$ $M^{-1}s^{-1}$ for **YC20** was calculated, which is two times faster than its reaction with the old MBP-dC10 ($k_2 = 7$ $M^{-1}s^{-1}$). This is consistent with the results found by Miroslava Strmiskova, who also observed a two-fold increase in reactivity when the similar asymmetric coumarin-based FlARe probe, **YC9**, was reacted with the dC10* tag (Figure 2.18).[44,45] Evidently, compared to **KT12** ($k_2 = 3.3 \pm 0.2$ $M^{-1}s^{-1}$), **YC20** reacted about four times faster with MBP-dC10*. This difference in reactivity mirrors **YC20's** increased reactivity with GSH, although the sub-magnitude improvement continues to reflect the small differences regarding substituent effects between the methoxy and isopropyl

substituents. However, it still suggests that regarding dimaleimide FlARe probes, minute differences that were previously unobservable using similar monomaleimide fluorogens may be emphasized.



YC9

Figure 2.18. Structure of YC9.

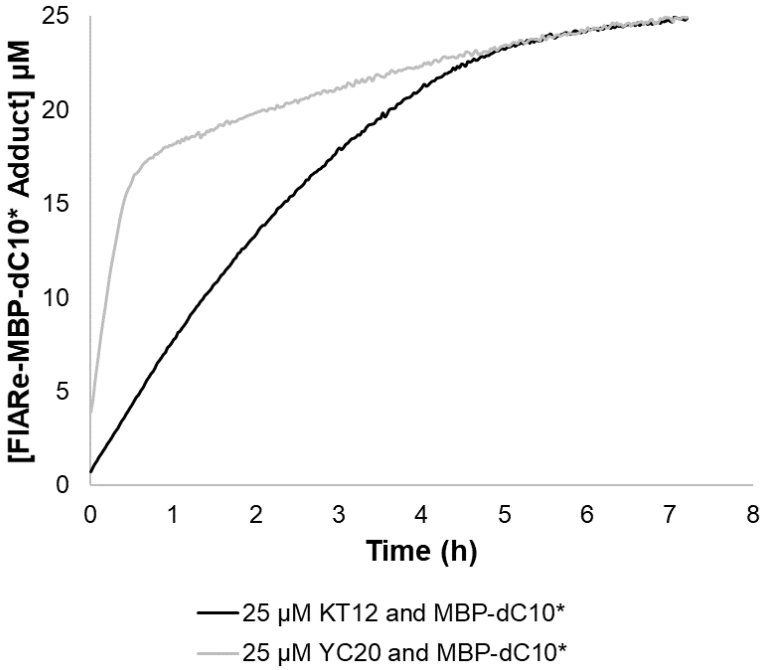


Figure 2.19. Reaction between 25 µM **KT12** or **YC20** and 25 µM MBP-dC10*. Reaction conditions: 125 µM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), 20% DMSO, 37 °C. Second order rate constants, k_2, were determined using the method of initial rates. **KT12** k_2 = 3.3 ± 0.2 $M^{-1}s^{-1}$. **YC20** k_2 = 13.7 ± 0.7 $M^{-1}s^{-1}$.

2.3.4.3. Comparison to the Current FlARe probe design

Table 2.5 summarizes the properties of **YC20** and **KT12** regarding their relative solubility, reactivity with MBP-dC10* and GSH in terms of second-order rate constants and a qualitative assessment of their respective hydrolytic stability. We observed that **YC20** exhibits higher solubility in aqueous media and reacts faster with MBP-dC10*. However, it also reacts faster with GSH and is less stable. Contrary to this, **KT12**, is more stable in aqueous conditions and reacts slower with GSH. Consequently, **KT12** also reacts slower with MBP-dC10* and reports solubility issues that are not found with **YC20**.

Table 2.5. Comparison of **KT12** and **YC20**. Beneficial and desired characteristics are **bolded**. Solubility was determined purely qualitatively based on whether or not precipitation was observed after the reaction in the old assay conditions (125 μM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), **10% DMSO**).

	KT12 (^{i}Pr)	**YC20 (OMe)**
Solubility	Less	**More**
MBP-dC10* Reactivity	k_2 = 3.3 ± 0.2 M^{-1} s^{-1}	$\boldsymbol{k_2}$ **= 13.7 ± 0.7** $\mathbf{M^{-1}}$ $\mathbf{s^{-1}}$
GSH Reactivity	$\boldsymbol{k_2}$ **= 1.72 ± 0.05 × 10^{-2} M^{-1}s^{-1}**	k_2 = 4.9 ± 0.3 × 10^{-2} $M^{-1}s^{-1}$
Stability vs. hydrolysis	**More**	Less

Although **KT12** is more stable against hydrolysis and reacts slower with GSH, which would ideally translate to lower background fluorescence, its lower solubility means that less probe can be used during imaging experiments. Unless more DMSO is introduced, which can prove harmful to cells,[7] **KT12's** lower solubility may result in lower brightness if there is insufficient probe in solution for protein labelling. Additionally, its slower reactivity with MBP-dC10* suggest that longer labelling times are required. This could lead to more cell death as the cells must remain outside optimal growth conditions during labelling. Currently, **YC20** requires 45 minutes for enough labelling in HEK293 cells.[2] Given **KT12's** lower

reaction constant, the isopropyl variant may take approximately four times as long for sufficient labelling (3 h). Furthermore, **KT12's** thiol addition rate constant with MBP-dC10* is 200 times faster than its rate constant with GSH. However, **YC20's** rate constant with the protein is 273 times higher than its reaction rate constant with GSH, suggesting that that **YC20** may be more selective than KT12. Therefore, given YC20's greater solubility, faster reactivity with MBP-dC10* and greater selectivity, **KT12** was not deemed a better alternative to **YC20**.

2.4. Conclusion

To summarize, a library of fluorogenic monomaleimide probes containing different maleimide ring substituents were prepared and their thiol addition with BME and base-catalyzed hydrolysis were studied. Regarding maleimide hydrolysis, product studies demonstrated a chemoselective reaction that favours reaction at the carbonyl adjacent to the substituted carbon. Two-parameter analysis of the substituent effects revealed that the hydrolysis rate constant was almost completely dependent on the steric properties of the substituent. On the other hand, the product studies showed that thiol addition regioselectivity depended on the nature of the substituent. Contrary to hydrolysis, two-parameter analysis indicated that the thiol addition rate constant was primarily dependent on the electronic effect of the substituent.

Based on this, we synthesized a variant of the current FlARe probe, **YC20**, carrying two isopropylmaleimides instead of methoxymaleimides (**KT12**), which was hypothesized to improve FlARe probe stability in aqueous media. Although **KT12** was more stable, it lacked the solubility, reaction speed with the dC10* tag and selectivity offered by **YC20**. Although we did not improve the current FlARe probe, the trends in reactivity elucidated herein will

still benefit the design of substituted maleimide derivatives for future maleimide-based bioconjugation or labelling methods.

2.5. Experimental

Compounds highlighted with * indicate that it has **not** been synthesized by Kelvin Tsao and so characterization data was **not** provided. All protocols and characterization data not provided here can be found in reference 1.

2.5.1. Synthesis

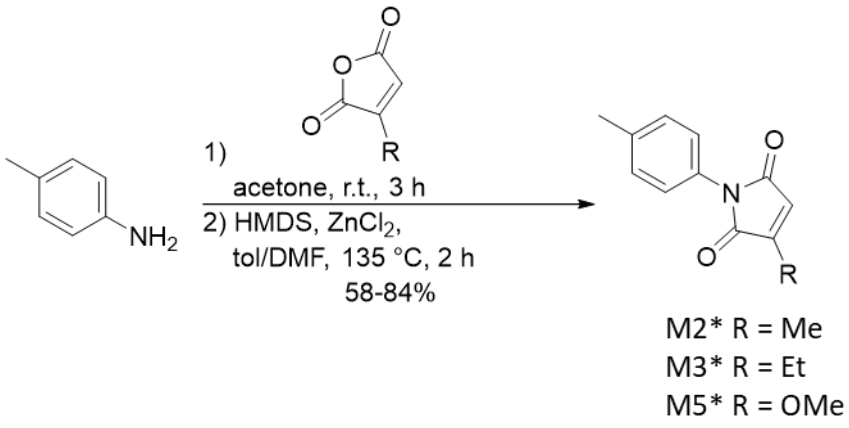
All reagents and solvents for reactions were used as received unless otherwise stated. Dichloromethane, dimethylformamide, methanol and acetonitrile were dried using a solvent purification system from LC Technology Solution Inc.

Reactions were monitored by thin layer chromatography (TLC) using E. Merck silica gel 60F254 precoated aluminum plates and visualized by illumination with a short-wavelength ultraviolet light or long-wavelength visible light. Flash column chromatography was performed on ZEOCHEM® silica gel 60 (ECO 40-63 μm) using ethyl acetate/n-hexane or methanol/dichloromethane as eluting solvents.

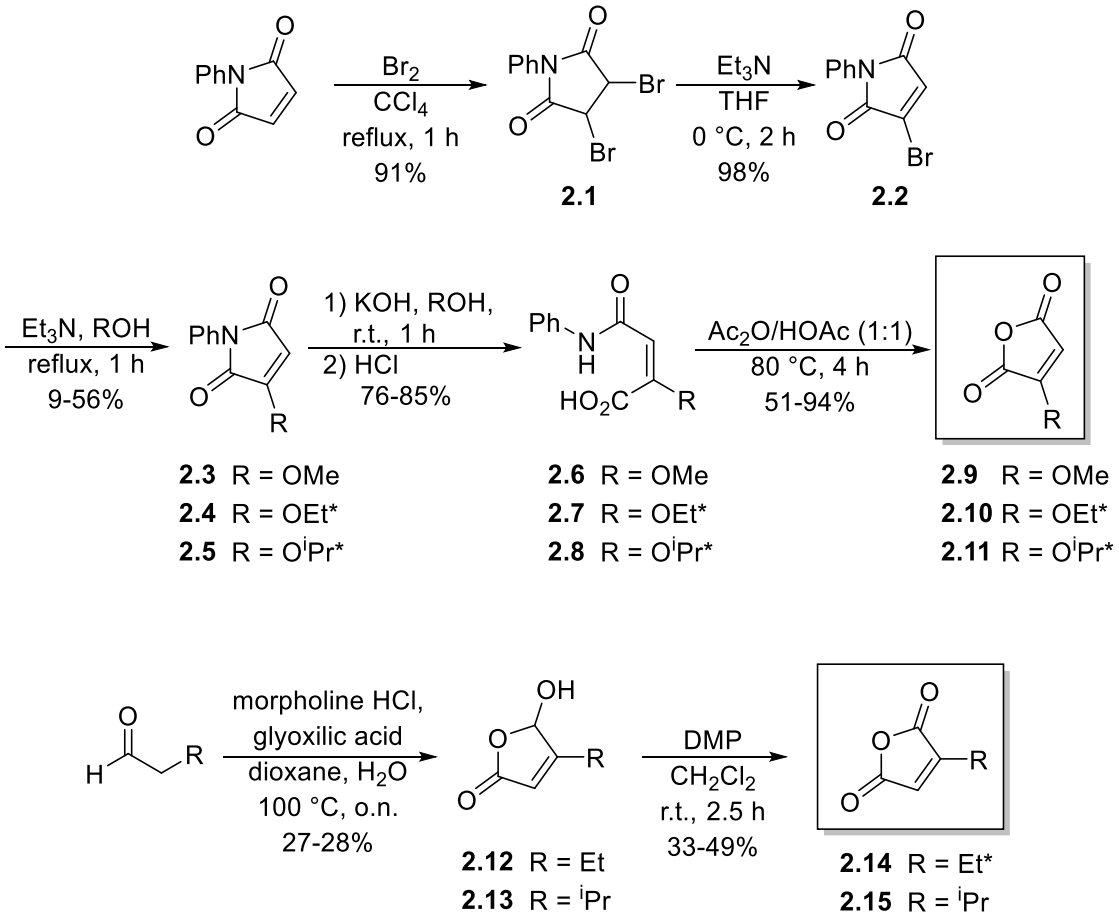
Nuclear magnetic resonance (NMR) spectra were recorded in deuterated DMSO or deuterated chloroform at ambient temperature. The experiments were performed either on a Bruker Avance 400 Fourier Transform Spectrometer operating at 400 MHz for ^{1}H and at 100 MHz for ^{13}C or on a Bruker Avance 500 Fourier Transform Spectrometer operating at 500 MHz for ^{1}H and at 125 MHz for ^{13}C. EI-MS spectra were recorded on a Kratos concept mass spectrometer for both low resolution and high-resolution mass spectra. ESI-MS spectra were

recorded on a Waters Micromass Q-TOF mass spectrometer. Melting points were measured on an EZ-Melt automated melting point apparatus and are uncorrected.

2.5.1.1. Synthesis of Model Compounds (Scheme 2.1) and Substituted Maleic Anhydrides (Scheme 2.6)

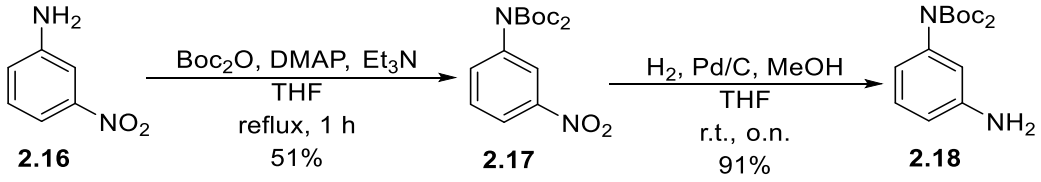


M2, **M3**, and **M5** were not synthesized by Kelvin Tsao. The synthesis and characterization of model compounds **M2**, **M3**, and **M5** can be found in reference 1.



Kelvin Tsao did not synthesize the substituted maleic anhydrides. The synthesis and characterization of the substituted maleic anhydrides and their intermediates can be found in reference 1.

2.5.1.2. Synthesis of Aryl Linker Intermediate (Scheme 2.7)

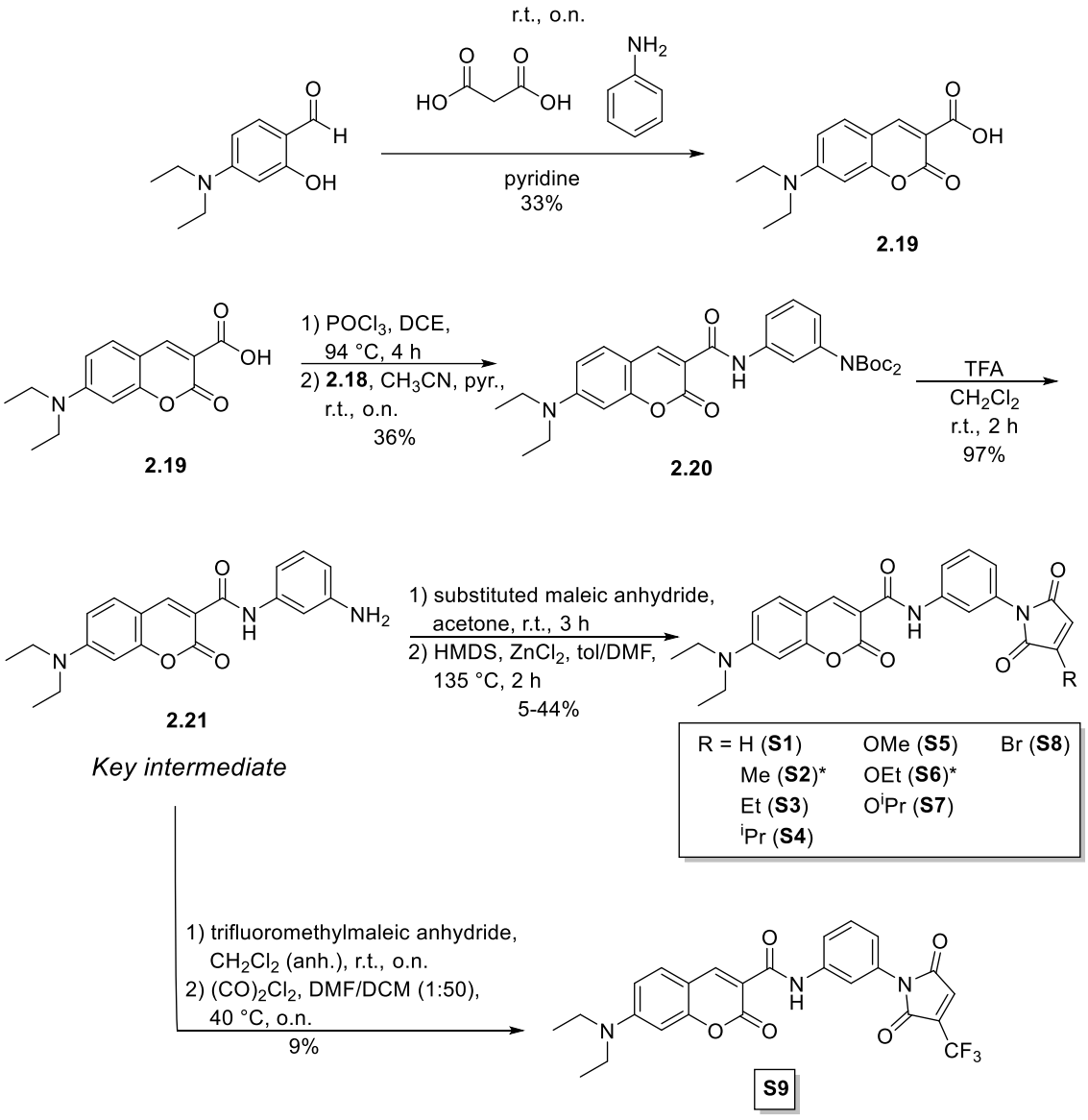


Compound **2.17**:[1] To a solution of 3-nitroaniline (compound **2.16**) (4.00 g, 29.0 mmol) in THF (40 mL) was added Boc_2O (19.0 g, 87.0 mmol), DMAP (26 mg, cat.), and Et_3N (10

mL, 73.0 mmol); this reaction mixture was heated to reflux for 3 h. The solvent was then removed using a rotary evaporator, and the residue was dissolved in Et_2O (40 mL), washed with 1 M HCl (3×10 mL), $NaHCO_3$ (10 mL), and brine (10 mL). The organic layer was dried with $MgSO_4$, filtered, and then concentrated to obtain compound **2.17** as a white solid (4.99 g, 14.8 mmol) in 51% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.19 (d, *J* = 8.0 Hz, 1H), 8.06 (t, *J* = 1.9 Hz, 1H), 7.65 – 7.42 (m, 2H), 1.45 (s, 18H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.4, 140.5, 134.4, 129.4, 123.5, 122.4, 83.8, 27.9. HRMS (ESI): m/z calcd for $C_{16}H_{22}N_2O_6Na$ ([MNa^+]): 361.1376. Found: 361.1360. **mp**: 70.0-71.2 °C.

Compound **2.18**:[1] Compound **2.17** (2.50 g, 7.40 mmol) was dissolved in MeOH/THF (16 mL/24 mL), and Pd/C (625 mg, 25% w/w) was added to the solution, and the mixture was stirred under H_2 overnight. The Pd/C was removed by filtering through Celite, and the filtrate was evaporated to dryness. The residue was purified by column chromatography to yield compound **2.18** as a white solid (2.07 g, 6.72 mmol) in 91% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.11 (t, *J* = 7.9 Hz, 1H), 6.60 (ddd, *J* = 8.0, 2.3, 0.9 Hz, 1H), 6.54 (ddd, *J* = 7.8, 1.9, 0.9 Hz, 1H), 6.48 (t, *J* = 2.1 Hz, 1H), 3.67 (s, 2H), 1.43 (s, 18H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 152.1, 146.9, 140.2, 129.4, 118.2, 114.7, 114.3, 82.5, 28.0. HRMS (ESI): m/z calcd for $C_{16}H_{24}N_2O_4Na$ ([MNa^+]): 331.1634. Found: 331.1645. **mp**: 136.3-137.5 °C.

2.5.1.3. Synthesis of Coumarin Maleimide Substrates (Scheme 2.9)



Compound **2.19**: 7-diethylaminocoumarin-3-carboxylic acid (compound **2.19**) was synthesized according to a literature procedure and spectra was consistent with what was reported.[46] Aniline (1.23 mL, 13.5 mmol) was added dropwise to a solution of 4-(diethylamino)salicylaldehyde (3.14 g, 15.0 mmol) and malonic acid (3.12 g, 30.0 mmol) dissolved in dry pyridine (15 mL). The mixture was stirred at room temperature overnight, after which EtOH (37.5 mL) was added. After stirring at room temperature for 1 h, the

precipitate was collected by vacuum filtration and rinsed with 0.1 M HCl (15 mL), H_2O (15 mL) and minimal Et_2O (5 mL). Compound **2.19** was obtained as an orange solid in 33% (1.29 g, 4.94 mmol) yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.33 (bs, 1H), 8.63 (s, 1H), 7.44 (d, *J* = 9.0 Hz, 1H), 6.71 (dd, *J* = 9.0, 2.3 Hz, 1H), 6.52 (d, *J* = 2.2 Hz, 1H), 3.49 (q, *J* = 7.1 Hz, 4H), 1.26 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.7, 164.6, 158.2, 153.9, 150.4, 132.1, 111.1, 108.7, 105.7, 97.0, 45.5, 12.5. HRMS (EI): m/z calcd for $C_{14}H_{15}NO_4$ ([MNa^+]) 261.1001. Found: 261.1031.

Compound **2.20**: Under nitrogen and anhydrous conditions, acid **2.19** (1.62 mg, 6.20 mmol) was dissolved in dichloroethane (60 mL). $POCl_3$ (2.9 mL, 31.0 mmol) was added, and the solution was heated to reflux and stirred for 4 h. Afterward, the solution was concentrated using a rotary evaporator, and the residue was dried under vacuum. Following this, the residue was dissolved in pyridine (60 mL) and aniline linker **2.18** (2.967 g, 9.30 mmol) was added. The mixture was stirred overnight at room temperature before removing the solvent by evaporation. The crude solid was dissolved in DCM (60 mL) and washed with 0.1 M HCl (20 mL) and brine (20 mL). The organic layer was dried with $MgSO_4$, which was removed through filtration, and the solution was concentrated to dryness. The crude product was purified by flash chromatography in 7:3 hexanes/ethyl acetate, and a yellow solid (1.23 g, 2.23 mmol) was obtained in 36% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.92 (s, 1H), 8.78 (s, 1H), 7.65 (dd, *J* = 7.3, 1.0 Hz, 2H), 7.46 (d, *J* = 9.0 Hz, 1H), 7.32 (t, *J* = 8.3 Hz, 1H), 6.94 – 6.83 (m, 1H), 6.68 (dd, *J* = 9.0, 2.4 Hz, 1H), 6.53 (d, *J* = 2.3 Hz, 1H), 3.47 (q, *J* = 7.1 Hz, 4H), 1.43 (s, 18H), 1.25 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 163.1, 161.1, 157.8, 152.9, 151.8, 148.6, 139.9, 138.8, 131.3, 128.9, 123.6, 119.9, 119.1, 110.2, 110.1, 108.6, 96.6,

82.7, 45.2, 27.9, 12.4. HRMS (ESI): m/z calcd for $C_{30}H_{37}N_3O_7Na$ ([MNa^+]): 574.2529. Found: 574.2571. **mp**: 170.7-171.5 °C.

Preparation of key intermediate compound **2.21**: Protected amine **2.20** (423 mg, 0.951 mmol) was dissolved in DCM (48 mL). TFA (10 mL) was added dropwise to the solution, which was stirred for 2 h at room temperature. The solvent was removed prior to dissolving the crude solid in a minimal amount of DCM and precipitating with acetonitrile. A yellow solid (324.8 mg, 0.924 mmol) was produced in 97% yield. Spectral data was consistent to those reported previously.[32] ^{1}H NMR (400 MHz, DMSO) δ 10.78 (s, 1H), 8.75 (s, 1H), 7.73 (d, *J* = 9.1 Hz, 1H), 7.62 (s, 1H), 7.33 – 7.16 (m, 2H), 6.85 (dd, *J* = 9.1, 2.3 Hz, 1H), 6.79 (dd, *J* = 7.1, 1.9 Hz, 1H), 6.67 (d, *J* = 2.1 Hz, 1H), 4.79 (s, 2H), 3.50 (q, *J* = 7.0 Hz, 4H), 1.15 (t, *J* = 7.0 Hz, 6H). ^{13}C NMR (101 MHz, DMSO) δ 162.3, 160.8, 158.4, 157.4, 152.9, 148.2, 139.2, 131.9, 130.0, 114.8, 114.3, 110.5, 110.4, 109.0, 107.9, 96.0, 44.4, 12.3. LRMS (ESI) m/z (%): 352.2740 (MH^+, 34.1%). HRMS (ESI): m/z calcd for $C_{20}H_{21}N_3O_3Na$ ([MNa+]): 374.1481. Found: 374.1497. **mp**: 130.0 °C.

Preparation of substrates **S1−S8**: Key intermediate **2.21** (53 mg, 150 μmol) was dissolved in acetone (8 mL), and the corresponding maleic anhydride (225 μmol, 1.5 eq.) was added to the solution. The mixture was stirred at 25 °C overnight. The solvents were evaporated under reduced pressure. The resulting residue was suspended in ether and vacuum filtered. The maleamic acid was collected as a dark brown solid and used in the next step of the reaction without further purification. $ZnCl_2$ (225 μmol, 1.5 eq.) was dissolved in toluene−DMF (9:1 mL), and a dilute solution of HMDS (375 μmol, 2.5 eq.) in toluene (2 mL) was added over 20 min. The resulting mixture was then heated to reflux for 3 h, after which the volatiles were evaporated under reduced pressure. The resulting residue was dissolved in

EtOAc (30 mL) and washed with 0.1 M HCl (10 mL) and saturated Na_2CO_3 (10 mL). The organic phase was dried with $MgSO_4$. The drying agent was removed through filtration, and the solution was concentrated to dryness. The compound was purified using flash column chromatography in 6:4 hexanes/ethyl acetate to obtain **S1**−**S8**.

S1: Obtained 21.2 mg (49 μmol) of white solid in 35% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.99 (s, 1H), 8.77 (s, 1H), 7.87 (t, *J* = 2.0 Hz, 1H), 7.78 – 7.63 (m, 1H), 7.49 – 7.41 (m, 2H), 7.09 (ddd, *J* = 7.9, 1.9, 0.9 Hz, 1H), 6.86 (s, 2H), 6.67 (dd, *J* = 9.0, 2.4 Hz, 1H), 6.53 (d, *J* = 2.2 Hz, 1H), 3.47 (q, *J* = 7.1 Hz, 4H), 1.25 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.4, 163.1, 161.2, 157.8, 152.9, 148.7, 139.2, 134.2, 131.7, 131.4, 129.5, 121.7, 119.8, 118.1, 110.2, 109.9, 108.6, 96.6, 45.2, 12.4. LRMS (EI) m/z (%): 431.1434 (M+, 29.3%). HRMS (ESI): m/z calcd for $C_{24}H_{21}N_3O_5$ ([MNa^+]): 454.1379. Found: 454.1363. **mp**: 167.8−168.0 °C.

S3: Obtained 28.4 mg (61.8 μmol) of yellow solid in 44% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.96 (s, 1H), 8.76 (s, 1H), 7.84 (t, *J* = 1.9 Hz, 1H), 7.68 (dd, *J* = 8.2, 1.2 Hz, 1H), 7.43 (dd, *J* = 16.7, 8.6 Hz, 2H), 7.12 – 7.05 (m, 1H), 6.67 (dd, *J* = 8.9, 2.2 Hz, 1H), 6.53 (d, *J* = 2.0 Hz, 1H), 6.43 (t, *J* = 1.9 Hz, 1H), 3.46 (q, *J* = 7.1 Hz, 4H), 2.55 (qd, *J* = 7.4, 1.9 Hz, 2H), 1.30 – 1.18 (m, *J* = 13.0, 7.2 Hz, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.3, 169.7, 163.1, 161.3, 157.9, 152.9, 151.8, 148.8, 139.2, 132.3, 131.5, 129.5, 126.0, 121.8, 119.7, 118.1, 110.5, 110.2, 108.9, 97.0, 45.4, 19.2, 12.5, 11.4. LRMS (ESI) m/z (%): 482.0051 (MNa^+, 100%). HRMS (ESI): m/z calcd for $C_{26}H_{25}N_3O_5Na$ ([MNa^+]): 482.1692. Found: 482.1678. **mp**: 172.1−172.3 °C.

S4: Obtained 9.50 mg (19.3 μmol) of yellow solid in 14% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.96 (s, 1H), 8.77 (s, 1H), 7.85 (t, *J* = 2.0 Hz, 1H), 7.68 (dd, *J* = 8.2, 1.1 Hz, 1H),

7.43 (ddd, *J* = 12.3, 9.4, 6.6 Hz, 2H), 7.10 (dd, *J* = 8.0, 1.0 Hz, 1H), 6.68 (dd, *J* = 9.0, 2.3 Hz, 1H), 6.54 (d, *J* = 2.2 Hz, 1H), 6.38 (d, *J* = 1.6 Hz, 1H), 3.47 (q, *J* = 7.1 Hz, 4H), 2.94 (dq, *J* = 6.8, 5.4 Hz, 1H), 1.30 – 1.22 (m, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.0, 169.7, 163.2, 161.3, 158.0, 156.0, 153.0, 148.8, 139.2, 132.3, 131.5, 129.5, 124.9, 121.8, 119.7, 118.2, 110.4, 110.2, 108.8, 96.9, 45.4, 30.1, 21.1, 12.6. LRMS (ESI) m/z (%): 496.1886 (MNa^+, 100%). HRMS (EI): m/z calcd for $C_{27}H_{27}N_3O_5Na$ ([MNa^+]) 496.1848. Found: 496.1886. **mp**: 154.1−154.3 °C.

S6: Obtained 29.5 mg (62.0 μmol) of a bright yellow solid in 44% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.95 (s, 1H), 8.76 (s, 1H), 7.81 (s, 1H), 7.71 (d, *J* = 8.1 Hz, 1H), 7.43 (dd, *J* = 18.0, 8.6 Hz, 2H), 7.09 (d, *J* = 7.9 Hz, 1H), 6.66 (dd, *J* = 8.9, 2.2 Hz, 1H), 6.52 (d, *J* = 2.0 Hz, 1H), 5.52 (s, 1H), 4.21 (q, *J* = 7.1 Hz, 2H), 3.46 (q, *J* = 7.0 Hz, 4H), 1.53 (t, *J* = 7.1 Hz, 3H), 1.24 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.2, 164.6, 163.2, 161.3, 159.9, 157.9, 153.0, 148.8, 139.2, 131.8, 131.5, 129.6, 121.9, 119.8, 118.3, 110.3, 110.1, 108.7, 96.8, 96.5, 68.7, 45.3, 14.1, 12.6. LRMS (ESI) m/z (%): 476.4018 (MH^+, 7.2%). LRMS (ESI) m/z (%): 498.3858 (MNa^+, 100%). HRMS (ESI): m/z calcd for $C_{26}H_{25}N_3O_6Na$ ([MNa^+]): 498.1641. Found: 498.1630. **mp**: 238.5−238.7 °C.

S7: Obtained 16.7 mg (34.1 μmol) of a yellow solid in 24% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.95 (s, 1H), 8.76 (s, 1H), 7.81 (s, 1H), 7.71 (d, *J* = 8.2 Hz, 1H), 7.53 – 7.37 (m, 2H), 7.09 (d, *J* = 7.9 Hz, 1H), 6.67 (dd, *J* = 9.0, 2.1 Hz, 1H), 6.53 (d, *J* = 2.0 Hz, 1H), 5.47 (s, 1H), 4.54 (dt, *J* = 12.3, 6.1 Hz, 1H), 3.46 (q, *J* = 7.1 Hz, 4H), 1.47 (d, *J* = 6.1 Hz, 6H), 1.25 (t, *J* = 7.0 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.6, 164.9, 163.2, 161.3, 159.0, 158.0, 153.0, 148.8, 139.2, 131.9, 131.5, 129.6, 121.9, 119.8, 118.3, 110.3, 110.1, 108.7, 96.8, 96.2,

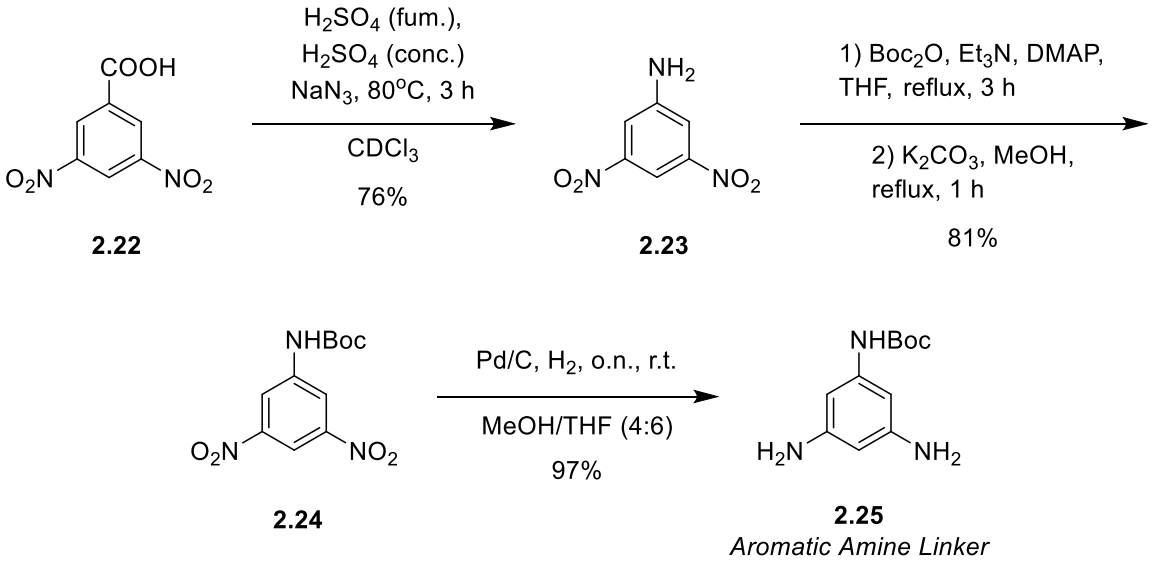
76.8, 45.3, 21.3, 12.6. LRMS (ESI) m/z (%): 511.8859 (MNa^+, 100%). HRMS (ESI): m/z calcd for $C_{27}H_{27}N_3O_6Na$ ([MNa^+]): 512.1797. Found: 512.1776. **mp**: 197.9 – 199.0 °C.

S8: Obtained 15 mg (29 μmol) of a white solid in 21% yield; ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.00 (s, 1H), 8.76 (s, 1H), 7.89 (t, *J* = 1.9 Hz, 1H), 7.68 (dd, *J* = 8.2, 1.1 Hz, 1H), 7.50 – 7.37 (m, 2H), 7.12 – 7.05 (m, 1H), 7.03 (s, 1H), 6.67 (dd, *J* = 9.0, 2.4 Hz, 1H), 6.52 (d, *J* = 2.2 Hz, 1H), 3.46 (q, *J* = 7.1 Hz, 4H), 1.25 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.4, 164.2, 163.2, 161.4, 158.0, 153.1, 148.8, 139.4, 132.0, 132.0, 131.6, 131.6, 129.7, 121.8, 120.2, 118.2, 110.4, 109.9, 108.7, 96.8, 45.3, 12.6. LRMS (ESI) m/z (%): 532.0435 (MNa^+, 100%). HRMS (ESI):m/z calcd for $C_{24}H_{20}N_3O_5BrNa$ ([MNa^+]): 532.0484. Found: 532.0467. **mp**: 136.2−136.4 °C.

S9: Under N_2, trifluoromethylmaleic anhydride (0.450 mmol, 43.7 μL) in 10 mL of anhydrous DCM was added to a solution of compound **2.21** (100 mg, 0.300 mmol) dissolved in 5 mL of anhydrous DCM and stirred overnight. The solution was evaporated, and the residue was resuspended in 10 mL of Et2O, after which the ether was removed and the residue dried. A round-bottom flask fitted with a condenser was purged and filled with N_2 before adding $(CO)_2Cl_2$ (130 μL, 1.46 mmol) and 0.3 mL of DMF through the condenser to limit water exposure; an additional 25 mL of DCM was then used to wash the condenser. The mixture was heated to reflux and stirred overnight before solvent was removed by evaporation. The final product was purified from this mixture by flash chromatography using 6:4 hexanes/EtOAc as the eluent. A yellow solid (13.2 mg, 26.4 μmol) was obtained in 9.3% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.03 (s, 1H), 8.77 (s, 1H), 7.95 (t, *J* = 1.9 Hz, 1H), 7.66 (d, *J* = 9.3 Hz, 1H), 7.46 (dd, *J* = 8.5, 7.2 Hz, 1H), 7.16 (d, *J* = 1.6 Hz, 1H), 7.08 (dd, *J* = 7.9, 1.0 Hz, 1H), 6.68 (dd, *J* = 9.0, 2.4 Hz, 1H), 6.53 (d, *J* = 2.2 Hz, 1H), 3.47 (q, *J* = 7.1 Hz, 2H),

1.26 (t, J = 7.1 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 165.9, 163.1, 161.4, 157.9, 153.0, 148.8, 139.4, 133.4, 131.5, 130.8, 129.7, 121.6, 120.4, 118.1, 110.3, 109.7, 108.6, 96.7, 45.2, 12.4. LRMS (ESI) m/z (%): 522.0706 (MNa^+, 100%). HRMS (ESI): m/z calcd for $C_{25}H_{20}N_3O_5F_3Na$ ([MNa^+]): 522.1240. Found: 522.1253. **mp**: 173.2−173.3 °C.

2.5.1.4. Synthesis of Aryl Diamine Linker (Scheme 2.10)



Compound **2.23**: Compound **2.23** was synthesized using a previously established protocol. Spectra are consistent with those that were reported.[26] Fuming H_2SO_4 (18 mL) and concentrated H_2SO_4 was added slowly to a solution of compound **2.22** (4.0 g, 18.9 mmol) dissolved in chloroform (25 mL). NaN_3 (1.4 g, 21.7 mmol) was added and the solution was heated to 80 °C and stirred for 3 h. After this, the solution was cooled to room temperature and poured over water and ice to form an orange precipitate. The precipitate was collected by vacuum filtration and dissolved in EtOAc (60 mL). The organic layer was washed with H_2O (20 mL) and brine (20 mL) and then dried with $MgSO_4$. **2.23** was collected after removing the solvent using a rotary evaporator and obtained in 76% (2.63 g, 14.4 mmol) as an orange solid. ^{1}H NMR (400 MHz, DMSO) δ 7.88 (s, 1H), 7.72 (d, J = 1.7 Hz, 2H), 6.51 (s, 2H). ^{13}C NMR

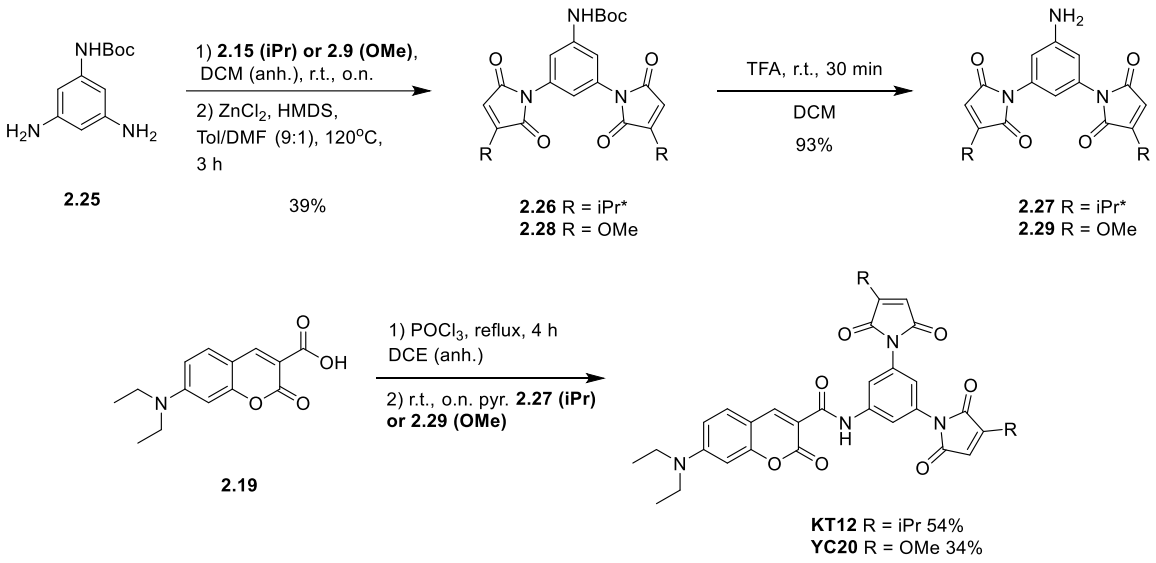
(100 MHz, DMSO) δ 151.2, 148.8, 112.4, 103.4. HRMS (EI): calcd for $C_6H_5N_3O_4$: 183.0280, found: 183.0280.

Compound **2.24**: Compound **2.24** was synthesized using a previously established protocol and spectra are consistent with those that were reported.[26] To a solution of **2.23** (2.00 g, 10.9 mmol) in THF (55 mL) was added Boc_2O (7.15 g, 32.8 mmol), Et_3N (2.98 mL, 27.3 mmol), and DMAP (75 mg, 5.6 mol%). The solution was heated to reflux and stirred for 3 h, after which the solvent was removed by evaporation. The residue was dissolved in Et_2O (150 mL) and washed with 1 M HCl (50 mL), sat. $NaHCO_3$ (50 mL), brine (50 mL) and then the solvent was removed from the organic layer by evaporation. The residue was re-dissolved in MeOH to which K_2CO_3 (4.53 g, 32.8 mmol) was added before heating the solution to reflux and stirring it for 1 h. Afterwards, the solution was partitioned between EtOAc (30 mL) and H_2O (60 mL). The organic layer was washed with 1 M HCl (20 mL) and brine (20 mL) before it was dried with $MgSO_4$, filtered, and then evaporated. The crude was purified by flash chromatography using a 9:1 hexane/EtOAc elution system to obtain compound **2.24** was a brown solid in 81% yield (2.51 g, 7.42 mmol). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.68 (t, J = 2.0 Hz, 1H), 8.63 (d, J = 2.0 Hz, 2H), 7.01 (s, 1H), 1.56 (s, 9H).

Compound **2.25**: Compound **2.25** was synthesized using a previously established protocol and spectra are consistent with those that were reported.[2] To a solution of compound **2.24** (1.5 g, 5.30 mmol) in a mixture of MeOH/THF (12 ml/18 mL) was added Pd/C (25 %w/w), under N_2. The N_2 in the flask was replaced with H_2 and the reaction was allowed to stir at room temperature overnight under H_2. Afterwards, the mixture was filtered through celite and the solvent was removed by evaporation to obtain pure compound **2.25** in 97% (1.14 g, 5.14 mmol) yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.40 (s, 1H), 6.20 (s, 2H), 5.75 (s, 1H),

3.68 (bs, 4H), 1.49 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 152.8, 147.9, 140.4, 97.4, 96.6, 80.4, 28.5. HRMS (EI): calcd for $C_{11}H_{17}N_3O_2$: 223.1321, found: 223.1326.

2.5.1.5. Synthesis of KT12 and YC20 (Scheme 2.11)



Compound **2.26**: To a solution of compound **2.25** (442 mg, 1.98 mmol) in anhydrous DCM was added compound **2.15** (610 mg, 4.35 mmol). The mixture was stirred at room temperature overnight. Afterwards, the solvent was removed by evaporation and the residue was dissolved in a mixture of toluene (5 mL) and DMF (1 mL), before adding $ZnCl_2$ (807 mg, 5.93 mmol) and a solution of HMDS (1.89 mL, 8.90 mmol) in toluene (4 mL) to the reaction vessel. The mixture was set to reflux and stirred for 3 h before removing the solvent by evaporation. Any remaining liquid was diluted with EtOAc (60 mL) and washed with 0.1 M HCl (20 mL), sat. Na_2CO_3 (20 mL) and brine (20 mL). After drying the organic layer with $MgSO_4$, filtering, and then removing the solvent by rotary evaporation, **2.26** was purified by flash chromatography using 2:8 EtOAc/hexanes as the eluent. **2.26** was obtained as a pale-yellow solid in 39% yield (365 mg, 0.792 mmol). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 (s, 2H), 7.16 (s, 1H), 6.70 (s, 1H), 6.36 (s, 2H), 2.91 (dt, *J* = 13.4, 6.7 Hz, 2H), 1.50 (s, 9H), 1.26 (d,

J = 6.8 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.6, 169.2, 156.1, 152.4, 139.6, 132.8, 125.8, 125.0, 117.0, 114.5, 81.2, 28.4, 26.1, 21.0. HRMS (ESI): m/z calcd for $C_{25}H_{29}N_3O_6Na$ ($[MNa^+]$): 490.1954. Found: 490.1953.

Compound **2.27**: **2.16** (365 mg, 0.782 mmol) was dissolved in DCM (10 mL) and TFA (10 mL) was added dropwise to the solution. The reaction was stirred at room temperature for 30 min and then dried by evaporation. The residue was washed with DCM (2 × 20 mL) and evaporated to remove any residual TFA before purifying by flash chromatography using 4:6 EtOAc/hexanes as the eluent. **2.27** was obtained as a yellow powder in 93% yield (267 mg, 0.726 mmol). ^{1}H NMR (400 MHz, $CDCl_3$): *δ* (ppm) 6.87 (s, 1 H), 6.70 (s, 2 H), 6.35 (s, 2 H), 4.02 (s, 2 H), 2.90 (m, 2 H), 1.26 (d, *J* = 6.2 Hz, 12 H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.8, 169.5, 156.0, 147.3, 133.0, 125.0, 112.8, 111.4, 26.1, 21.0. HRMS (ESI): m/z calcd for $C_{20}H_{21}N_3O_4$ ($[MNa^+]$): 390.1430. Found: 390.1405.

KT12: Under N_2, compound **2.19** (50.0 mg, 0.191 mmol) was dissolved in DCE (anh., 4 mL). To this, $POCl_3$ (5 eq., 0.088 mL, 0.957 mmol) was added followed by heating the reaction to 94 °C and stirring for 4 h. The solvent was removed by evaporation before dissolving the residue in pyridine (4 mL) and adding compound **2.27** (79.3 mg, 0.191 mmol) and stirring overnight. Afterwards, the reaction was diluted with toluene (20 mL) and then dried by evaporation. The residue was dissolved in DCM (30 mL) and washed with 0.1 M HCl (2x10 mL) and brine (20 mL). This was followed by drying the organic layer with $MgSO_4$, filtering, and then evaporation before purifying the compound by flash chromatography using 4:6 EtOAc/hexanes as the eluent. **KT12** was obtained as a yellow solid in 54% yield (63.3 mg, 0.104 mmol). ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.03 (s, 1H), 8.76 (s, 1H), 7.83 (s, 2H), 7.45 (d, *J* = 8.0 Hz, 1H), 7.23 (s, 1H), 6.66 (d, *J* = 7.4 Hz, 1H), 6.52 (s, 1H),

6.39 (s, 2H), 3.46 (q, *J* = 6.8 Hz, 4H), 2.92 (dq, *J* = 12.6, 6.3 Hz, 2H), 1.30 – 1.22 (m, 18H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.5, 169.2, 163.1, 161.4, 158.0, 156.04, 153.1, 148.9, 139.5, 132.6, 131.6, 125.0, 118.3, 116.5, 110.4, 109.8, 108.7, 96.8, 45.3, 26.1, 21.0, 12.6. HRMS (ESI): m/z calcd for C34H34N4O7 ([MNa^+]): 633.2325. Found: 633.2371.

Compound **2.28** and compound **2.29** was taken from a previous study. Its synthesis is like that of compound **2.26** and **2.27**, in which the methoxymaleic anhydride (compound **2.9**) was used instead of the isopropyl maleic anhydride (compound **2.15**). The full synthesis and characterization of these two compounds can be found in reference 2.

YC20: YC20 was synthesized by following the protocol outlined in reference 2. The synthetic procedure is outlined above for **KT12**, except compound **2.29** was used instead of compound **2.27**. Starting from 26.1 mg (100 μmol) of compound **2.19**, **YC20** was obtained as a yellow solid in 34% yield (20 mg, 34 μmol) ^{1}H NMR (300 MHz, $CDCl_3$) δ 11.04 (s, 1H), 8.77 (s, 1H), 7.84 (d, *J* = 1.9 Hz, 2H), 7.46 (d, *J* = 9.0 Hz, 1H), 7.23 (s, 1H), 6.67 (d, *J* = 8.9 Hz, 1H), 6.54 (s, 1H), 5.57 (s, 2H), 4.01 (s, 6H), 3.47 (d, *J* = 7.3 Hz, 4H), 1.25 (s, 6H).

2.5.2. Product Studies

Three N-*p*-methylphenylmaleimides (**M2**, **M3**, and **M5**), synthesized as model compounds for substrates **S2**, **S3**, and **S5**, were subjected to thiol addition and hydrolysis reactions, and the products of those reactions were analyzed by ^{1}H NMR. For thiol addition, solutions were prepared of 150 mM of each model compound in 1 mL of *d*6-DMSO containing 1.0 eq. of BME (10.5 μL). A solution containing only the model compound served as a blank. After allowing each solution to react at room temperature overnight, 500 μL of the reaction mixture was transferred to an NMR tube and ^{1}H NMR spectra were recorded.

For the hydrolysis product studies, 10% NaOH was added to the model compounds dissolved in acetonitrile to make a final NaOH solution of 1% NaOH. The mixture was allowed to react for 3 h before removing the acetonitrile by rotary evaporation. The remaining solution was diluted with 10 mL of H_2O, and the respective maleamic acid was precipitated using 1 M HCl.

Hydrolyzed M2:[29] Obtained 51.8 mg (0.236 mmol) of a white solid in 48% yield. ^{1}H NMR (400 MHz, DMSO) δ 12.93 (s, 1H), 10.10 (s, 1H), 7.50 (d, *J* = 8.4 Hz, 2H), 7.10 (d, *J* = 8.3 Hz, 2H), 6.09 (d, *J* = 1.5 Hz, 1H), 2.24 (s, 3H), 1.98 (d, *J* = 1.2 Hz, 3H).^{13}C NMR (100 MHz, DMSO) δ (ppm) 170.1, 162.5, 142.6, 136.4, 132.5, 129.1, 123.3, 1193, 20.6, 20.5. HRMS (EI): m/z calcd for $C_{12}H_{13}NO_3$ ([M^+]) 219.08954. Found: 219.08679. **mp**: 159.6−159.7 °C.

Hydrolyzed M3: Obtained 66.7 mg (0.286 mmol) of a yellow solid in 63% yield. ^{1}H NMR (400 MHz, DMSO) δ 10.08 (s, 1H), 7.48 (d, *J* = 8.4 Hz, 2H), 7.10 (d, *J* = 8.3 Hz, 2H), 5.99 (s, 1H), 2.29 (q, *J* = 7.4 Hz, 2H), 2.23 (s, 3H), 1.05 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (100 MHz, DMSO) δ 170.7, 162.7, 149.1, 136.6, 132.8, 129.4, 120.8, 119.4, 27.3, 20.7, 11.9. HRMS (EI): m/z calcd for $C_{13}H_{15}NO_3$ ([M^+]) 233.10519. Found: 233.10346. **mp**: 132.9−133.0 °C.

Hydrolyzed M5: Obtained 101 mg (0.429 mmol) of a white solid in 93% yield. ^{1}H NMR (400 MHz, DMSO) δ 10.04 (s, 1H), 7.45 (d, *J* = 8.4 Hz, 2H), 7.10 (d, *J* = 8.3 Hz, 2H), 5.54 (s, 1H), 3.68 (s, 3H), 2.23 (s, 3H). ^{13}C NMR (100 MHz, DMSO) δ 164.5, 163.5, 160.3, 136.7, 132.7, 129.4, 119.5, 96.7, 56.7, 20.7. HRMS (EI): m/z calcd for $C_{12}H_{13}NO_4$ ([M^+]) 235.08446. Found: 235.08303. **mp**: 179.8−179.9 °C.

*2.5.3. Expression and Purification of Test Protein MBP-dC10**

The plasmid that coded for maltose binding protein tagged with dC10* (MBP-dC10*) has been designed and created in a published study outlining the rational design in improving the reactivity of the original dC10 tag by introducing positively charged residues around the cysteine.[47] MBP-dC10* was expressed and purified according to the protocol outlined in the same study. This involved transforming the expression plasmid into the BL21-Gold (DE3) strain of *E. coli* cells.[47] A successfully transformed colony (screened using ampicillin resistance) was then chosen and grown in 3 mL of LB medium containing 100 μg/mL ampicillin overnight. Next, this pre-culture was used to inoculate 500 mL LB medium (0.2% w/v glucose, 100 μg/mL ampicillin), which was then incubated at 37 °C and shaken at 200 rpm. Upon reaching an OD_{600} of 0.4-0.6, IPTG was added to a final concentration of 0.3 mM and the culture was induced for 3 h by shaking at 200 rpm and 37 °C. Following induction, the cells were harvested by centrifugation (4000 g, 20 min or 2000 g, 40 min) and the pellet was re-suspended in MBP binding buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.4). The suspended cells were then lysed by sonication (30 s pulses at 100% intensity, 1 min pause between cycles, 3 cycles) using a Hielscher sonicator. The soluble protein was separated from insoluble matter by centrifugation (6400 g, 20 min) and during this time the amylose resin was prepared. To prepare the amylose resin, a 2 mL volume of amylose resin was added to the column and washed with 5 column volumes (CV) of binding buffer. The soluble protein fraction was then incubated with the amylose resin for 2 h at 4 °C with gentle shaking, after which the flow-through was collected by gravity chromatography, followed by 3 washes (3 CV/wash) with binding buffer. MBP-dC10* was eluted with 2.5 CV of elution buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM D-maltose, pH 7.4) and then buffer exchanged

into 50 mM HEPES buffer (I = 100 mM, pH 7.4) to remove maltose. Protein purity was confirmed by SDS-PAGE, and protein yields generally ranged around 12 mg from 500 mL of expression culture, according to the Bradford Assay (Figure 2.20).

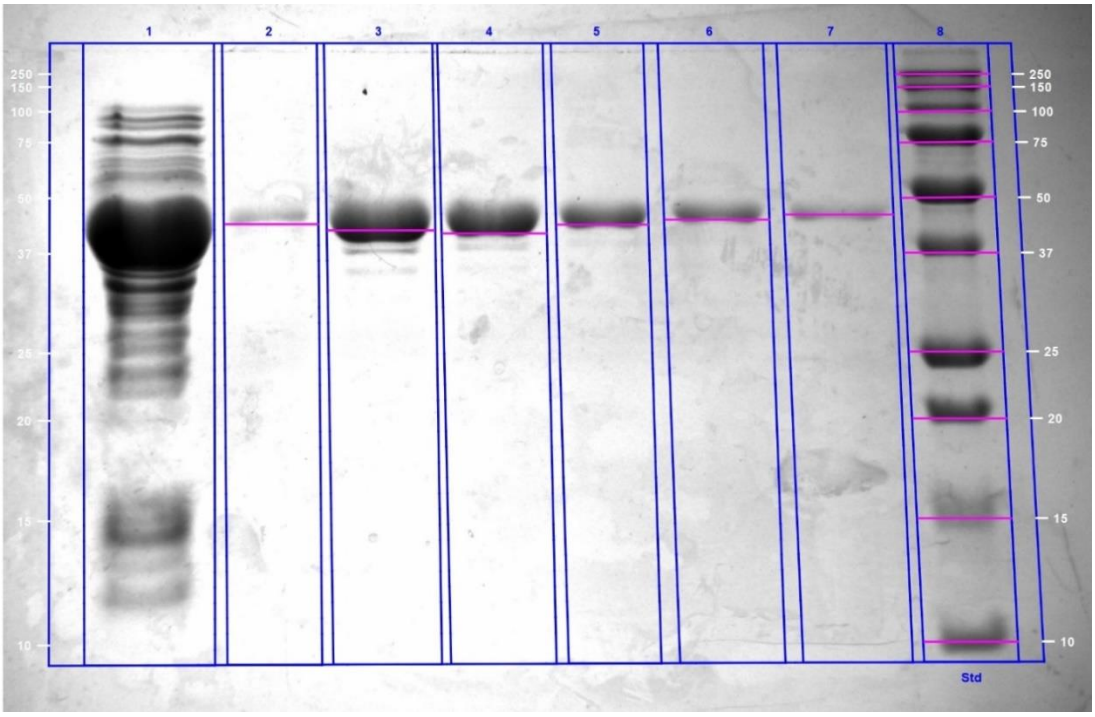


Figure 2.20. SDS-PAGE gel following MBP-dC10* purification. Lanes: 1) Column flow-through 2) Wash with binding buffer (*vido supra*) 3-7) Eluted fractions using 10 mM maltose elution buffer 8) Precision Plus Protein™ All Blue Prestained Protein Standards from BioRad. Gel analyzed and annotated with Image Lab. Bands in lanes 3-7 were analyzed to have a MW from 41.2-45.6 kDa. MW of MBP-dC10 = 44.9 kDa.

2.5.4. Kinetic Studies.

2.5.4.1. Hydrolysis and Thiol Addition of Coumarin Maleimide Substrates

Kinetic experiments were performed at either 22 (hydrolysis) or 37 °C (thiol addition) using either a Synergy H4 Hybrid Multi-Mode Microplate Reader with excitation and emission monochromators set at a 9-nm bandpass or an OLIS RSM 1000 stopped-flow fluorimeter. Buffered solutions containing different concentrations of BME were prepared in a 96-well plate or loaded into the stopped-flow filling syringe immediately before recording.

Plate reader reactions were initiated by the addition of stock DMSO solutions of the coumarin *N*-arylmaleimide substrates (**S1**−**S7**), while stopped-flow reactions were initiated by mixing with a solution of substrate diluted in reaction buffer. Hydrolysis reactions were studied using 25 μM coumarin *N*-arylmaleimide substrate in final solutions containing 5% DMSO and buffers at pH 5.0 (50 mM acetate buffer), 6.1 (50 mM MES buffer), 6.6 (50 mM MES buffer), 7.2 (50 mM MOPS buffer), 7.7 (50 mM HEPES buffer), 8.1 (50 mM Tris buffer), 8.6 (100 mM Tris buffer), 9.0 (100 mM CHES buffer), 9.5 (50 mM CHES buffer), 10.0 (100 mM CAPS buffer), and 10.5 (50 mM CAPS buffer). Ionic strength was adjusted to 0.075 M with KCl. For thiol addition experiments, the final reaction solution contained 25 μM coumarin *N*-arylmaleimide substrate and 0.250, 0.500, 0.750, 1.00, 1.25, 1.87, 2.50 mM BME in 50 mM HEPES buffer (pH 7.4) with 5% DMSO. Samples were excited at 440 nm, and fluorescence intensity was monitored at 485 nm as a function of time. All curves of time-dependent fluorescence increase were followed over >5 half-lives and fitted by nonlinear regression to mono-exponential equation 2.1.

Equation 2.1: $F_t = F_0 + F(1 - e^{-k_{obs}t})$

where F_t is the fluorescence intensity at time t, F_0 is the fluorescence intensity at time zero, F is the increase in fluorescence intensity, and k_{obs} is the measured pseudo-first-order rate constant. For hydrolysis reactions, the second-order rate constant for hydroxide-catalyzed hydrolysis (k_{OH}) was calculated by plotting k_{obs} values against hydroxide concentration and fitting to equation 2.2

Equation 2.2: $k_{obs} = k_w + k_{OH}[-OH]$

where k_w is the rate constant for pH-independent reaction with water, observed below pH 7.

For thiol addition reactions, second-order rate constants (k_2) were calculated by plotting k_{obs} values against BME concentration and fitting to equation 2.3

Equation 2.3: $k_{obs} = k_0 + k_2[BME]$

where k_0 is the rate constant for the increase of fluorescence, extrapolated to zero BME. Under the conditions of the thiol addition reactions (pH 7.4) this was always observed to be negligible.

2.5.4.2. Reaction of FlARe probes with GSH or MBP-dC10*

A Synergy H4 Hybrid Multi-Mode Microplate Reader with excitation and emission monochromators set at a 20-nm bandpass was used to carry out kinetic experiments at 37 ºC. A 96-well plate was prepared with solutions containing 25 µM MBP-dC10* or 1-10 mM GSH in 50 mM HEPES buffer (I = 100 mM adjusted with NaCl, pH 7.4) with 20% DMSO and 125 µM TCEP. Blanks consisted of solutions containing just 50 mM HEPES buffer (I = 100 mM adjusted with NaCl, pH 7.4) with 20% DMSO and 125 µM TCEP instead. Labelling reagent was added to a final concentration of 25 µM immediately before recording. Samples were excited at 435 nm and emission intensity was followed at 485 nm as a function of time. Second-order rate constants and pseudo-first-order rate constants were calculated using the method of initial rates. This was done by first normalizing the relative fluorescence intensity at each time point in the reaction to the maximal fluorescence reached at the end of the kinetic run. After normalizing the fluorescence, the maximum fluorescence (ie. 100% fluorescence) was assumed to be the maximum concentration of product that could be reached (ie. 25 µM of product). This allowed a [product] vs. time curve to be constructed. The rate constants were then calculated by taking the slope of the first 10% of the reaction (from 0 – 2.5 µM of product)

on the [product] vs. time curve (v_{init}) and determining the k_{obs} using equation 2.4 or k_2 using equation 2.5.

Equation 2.4: $$k_{obs} = \frac{v_{init}}{[FlARe\ probe]_{init}}$$

Equation 2.5: $$k_2 = \frac{v_{init}}{[FlARe\ probe]_{init} \times [MBP{-}dC10{*}]_{init}}$$

2.5.5. Approximation of the Isopropoxy Taft Parameter

To our knowledge, there is no reported value for the Taft steric parameter (E_s) for the isopropoxy group. Given that these values are principally additive, an approximate value of 0.42 was estimated using extrapolations from literature values (Table 2.6).

Changing an alpha methylene to oxygen converts a propyl group to an ethoxy group, which increases the E_s parameter by 1.26. Applying the same alteration to isobutyl yields a value of 0.33 for the corresponding isopropoxy group.

Inserting an oxygen changes methyl to methoxy, which increases the E_s to 0.99. Alternatively, inserting oxygen into ethyl gives an ethoxy group, with an increase of 0.97. With this reasoning, applying the same insertion to isopropyl yields a value of 0.52 or 0.50 for isopropoxy.

Adding a methyl group to propyl at the β position gives isobutyl, with a decrease in E_s of 0.57. Thus, applying the same branching to ethoxy yields a value of 0.33.

By taking the average of these four values (0.33, 0.52, 0.50 and 0.33) gives the E_s value of 0.42 for isopropoxy. Since the E_s value of Me is -1.24 with respect to the E_s value of H, the E_s value of isopropoxy with respect to H would therefore be -0.82, which is used in the LFER analysis.

Table 2.6. Selected Taft steric parameters taken from Taft, R. W., Jr. Newman, M. S., Ed.; John Wiley & Sons, Inc.: New York, 1956.[42]

R	E_s
Pr	-0.36
OEt	0.9
^{i}Bu	-0.93
Me	0
OMe	0.99
^{i}Pr	-0.47
Et	-0.07

2.6. References in Chapter 2

(1) Chen, Y.; Tsao, K.; De Francesco, É.; Keillor, J. W. (2015) Ring Substituent Effects on the Thiol Addition and Hydrolysis Reactions of N-Arylmaleimides. *J. Org. Chem.* 80, 12182–12192.

(2) Chen, Y.; Clouthier, C. M.; Tsao, K.; Strmiskova, M.; Lachance, H.; Keillor, J. W. (2014) Coumarin-Based Fluorogenic Probes for No-Wash Protein Labeling. *Angew. Chem. Int. Ed. Engl.* 53, 13785–13788.

(3) Girouard, S.; Houle, M.-H.; Grandbois, A.; Keillor, J. W.; Michnick, S. W. (2005) Synthesis and Characterization of Dimaleimide Fluorogens Designed for Specific Labeling of Proteins. *J. Am. Chem. Soc.* 127, 559–566.

(4) Machida, M.; Machida, M. I.; Kanaoka, Y. (1977) Hydrolysis of N-Substituted Maleimides: Stability of Fluorescence Thiol Reagents in Aqueous Media. *Chem. Pharm. Bull. (Tokyo).* 25, 2739–2743.

(5) Hansch, C.; Leo, A.; Taft, R. W. (1991) A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* 91, 165–195.

(6) Hammett, L. P. (1937) The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* 59, 96–103.

(7) Taft, R. W. (1952) Polar and Steric Substituent Constants for Aliphatic and o-Benzoate Groups from Rates of Esterification and Hydrolysis of Esters. *J. Am. Chem. Soc.* 74, 3120–3128.

(8) Brown, H. C.; Okamoto, Y. (1958) Electrophilic Substituent Constants. *J. Am. Chem.*

Soc. 80, 4979–4987.

(9) Heitz, J. R.; Anderson, C. D.; Anderson, B. M. (1968) Inactivation of Yeast Alcohol Dehydrogenase by N-Alkylmaleimides. *Arch. Biochem. Biophys.* 127, 627–636.

(10) Matsui, S.; Aida, H. (1978) Hydrolysis of Some N-Alkylmaleimides. *J. Chem. Soc. Perkin Trans. 2* 12, 1277.

(11) Bednar, R. A. (1990) Reactivity and PH Dependence of Thiol Conjugation to N-Ethylmaleimide: Detection of a Conformational Change in Chalcone Isomerase. *Biochemistry* 29, 3684–3690.

(12) Lee, C. C.; Samuels, E. R. (1964) The Kinetics of Reaction Between L-Cysteine Hydrochloride and Some Maleimides. *Can. J. Chem.* 42, 168–170.

(13) Schopfer, L. M.; Salhany, J. M. (1998) Spectroscopic and Kinetic Characterization of Eosin-5-Maleimide. *Anal. Biochem.* 257, 139–148.

(14) Gregory, J. D. (1955) The Stability of N-Ethylmaleimide and Its Reaction with Sulfhydryl Groups. *J. Am. Chem. Soc.* 77, 3922–3923.

(15) Roberts, E.; Rouser, G. (1958) Spectrophotometric Assay for Reaction of N-Ethylmaleimide with Sulfhydryl Groups. *Anal. Chem.* 30, 1291–1292.

(16) Corrie, J. E. T. (1994) Thiol-Reactive Fluorescent Probes for Protein Labelling. *J. Chem. Soc. Perkin Trans.* 20, 2975–2982.

(17) Kanaoka, Y.; Sekine, T.; Machida, M.; Soma, Y.; Tanizawa, K.; Ban, Y. (1964) Studies on Protein-Sulfhydryl Reagents. I. Synthesis of Benzimidazole Derivatives of Maleimide ; Fluorescent Labeling of Maleimide. *Chem. Pharm. Bull. (Tokyo).* 12, 127–

134.

(18) Kanaoka, Y. (1977) Organic Fluorescence Reagents in the Study of Enzymes and Proteins. *Angew. Chemie Int. Ed. English* 16, 137–147.

(19) Chen, Y.; Tsao, K.; Keillor, J. W. (2014) Fluorogenic Protein Labelling: A Review of Photophysical Quench Mechanisms and Principles of Fluorogen Design. *Can. J. Chem.* 398, 1–10.

(20) Orange, C.; Specht, A.; Puliti, D.; Sakr, E.; Furuta, T.; Winsor, B.; Goeldner, M. (2008) Synthesis and Photochemical Properties of a Light-Activated Fluorophore to Label His-Tagged Proteins. *Chem. Commun.* 10, 1217–1219.

(21) Song, A.; Wang, X.; Lam, K. S. (2003) A Convenient Synthesis of Coumarin-3-Carboxylic Acids via Knoevenagel Condensation of Meldrum's Acid with Ortho-Hydroxyaryl Aldehydes or Ketones. *Tetrahedron Lett.* 44, 1755–1758.

(22) Corrie, J. E. T. (1990) A Convenient Synthesis of N-(7-Dimethylamino-4-Methylcoumarin-3-Yl)-Maleimide Incorporating a Novel Variant of the Pechmann Reaction. *J. Chem. Soc. Perkin Trans. 1* 7, 2151.

(23) Chaurasia, C. S.; Kauffman, J. M. (1990) Synthesis and Fluorescent Properties of a New Photostable Thiol Reagent "BACM." *J. Heterocycl. Chem.* 27, 727–733.

(24) Sippel, T. O. (1981) New Fluorochromes for Thiols: Maleimide and Iodoacetamide Derivatives of a 3-Phenylcoumarin Fluorophore. *J. Histochem. Cytochem.* 29, 314–316.

(25) Machida, M.; Machida, M. I.; Sekine, T.; Kanaoka, Y. (1977) Fluorescence Properties

and Reactions of a Novel Fluorescence Thiol Reagent, N-(7-Dimethylamino-4-Methyl-3-Coumarinyl)Maleimide(DACM). *Chem. Pharm. Bull. (Tokyo).* 25, 1678–1684.

(26) Caron, K.; Lachapelle, V.; Keillor, J. W. (2011) Dramatic Increase of Quench Efficiency in “Spacerless” Dimaleimide Fluorogens. *Org. Biomol. Chem.* 9, 185–197.

(27) Faturaci, Y.; Coşkun, N. (2012) Substituent Effects on the Regioselectivity of Maleamic Acid Formation and Hydrogen Chloride Addition to N-Aryl Maleimides. *Turkish J. Chem.* 36, 749–758.

(28) Sortino, M.; Postigo, A.; Zacchino, S. (2013) Effects of Chirality on the Antifungal Potency of Methylated Succinimides Obtained by Aspergillus Fumigatus Biotransformations. Comparison with Racemic Ones. *Molecules* 18, 5669–5683.

(29) Botvinik, M. M.; Konovolo, I. M. (1965) Reactions of 1-Acyl Derivatives of Histidine with Serine Derivatives. *J. Gen. Chem. USSR* 35, 1127.

(30) Awuah, E.; Capretta, A. (2011) Development of Methods for the Synthesis of Libraries of Substituted Maleimides and α,β-Unsaturated-γ-Butyrolactams. *J. Org. Chem.* 76, 3122–3130.

(31) Mabire, A. B.; Robin, M. P.; Quan, W. D. W.-D.; Willcock, H.; Stavros, V. G.; O’Reilly, R. K. (2015) Aminomaleimide Fluorophores: A Simple Functional Group with Bright, Solvent Dependent Emission. *Chem. Commun.* 51, 9733–9736.

(32) Wang, M.; Wen, J.; Qin, Z.; Wang, H. (2015) A New Coumarin-Rhodamine FRET System as an Efficient Ratiometric Fluorescent Probe for Hg2+ in Aqueous Solution and in Living Cells. *Dye. Pigment.* 120, 208–212.

(33) Tedaldi, L. M.; Smith, M. E. B.; Nathani, R. I.; Baker, J. R. (2009) Bromomaleimides: New Reagents for the Selective and Reversible Modification of Cysteine. *Chem. Commun.* 43, 6583–6585.

(34) Ryan, C. P.; Smith, M. E. B.; Schumacher, F. F.; Grohmann, D.; Papaioannou, D.; Waksman, G.; Werner, F.; Baker, J. R.; Caddick, S. (2011) Tunable Reagents for Multi-Functional Bioconjugation: Reversible or Permanent Chemical Modification of Proteins and Peptides by Control of Maleimide Hydrolysis. *Chem. Commun.* 47, 5452–5454.

(35) Guy, J.; Castonguay, R.; Campos-Reales Pineda, N. B.; Jacquier, V.; Caron, K.; Michnick, S. W.; Keillor, J. W. (2010) De Novo Helical Peptides as Target Sequences for a Specific, Fluorogenic Protein Labelling Strategy. *Mol. Biosyst.* 6, 976–987.

(36) Yang, J. R.; Langmuir, M. E. Synthesis and Properties of a Maleimide Fluorescent Thiol Reagent Derived from a Naphthopyranone. *Journal of Heterocyclic Chemistry*. 1991, pp 1177–1180.

(37) Ambrus, J. I.; Halliday, J. I.; Kanizaj, N.; Absalom, N.; Harpsoe, K.; Balle, T.; Chebib, M.; McLeod, M. D. (2012) Covalent Attachment of Antagonists to the A7 Nicotinic Acetylcholine Receptor: Synthesis and Reactivity of Substituted Maleimides. *Chem. Commun.* 48, 6699–6701.

(38) Serafimova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.; Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frödin, M.; Taunton, J. (2012) Reversible Targeting of Noncatalytic Cysteines with Chemically Tuned Electrophiles. *Nat. Chem. Biol.* 8, 471–476.

(39) Raycroft, M. A. R.; Racine, K.; Rowley, C. N.; Keillor, J. W. (2018) Mechanisms of Alkyl and Aryl Thiol Addition to N-Methylmaleimide. *J. Org. Chem.* 83, 11674–11685.

(40) Jencks, W. P.; Salvesen, K. (1971) Equilibrium Deuterium Isotope Effects on the Ionization of Thiol Acids. *J. Am. Chem. Soc.* 93, 4433–4436.

(41) Lincoln, R.; Greene, L. E.; Krumova, K.; Ding, Z.; Cosa, G. (2014) Electronic Excited State Redox Properties for BODIPY Dyes Predicted from Hammett Constants: Estimating the Driving Force of Photoinduced Electron Transfer. *J. Phys. Chem. A* 118, 10622–10630.

(42) Taft Jr, R. W. *Steric Effects in Organic Chemistry*; Newman, M. S., Ed.; Wiley: New York, (1956).

(43) Jiang, X.; Yu, Y.; Chen, J.; Zhao, M.; Chen, H.; Song, X.; Matzuk, A. J.; Carroll, S. L.; Tan, X.; Sizovs, A.; et al. (2015) Quantitative Imaging of Glutathione in Live Cells Using a Reversible Reaction-Based Ratiometric Fluorescent Probe. *ACS Chem. Biol.* 10, 864–874.

(44) Chen, Y.; Tsao, K.; Acton, S. L.; Keillor, J. W. (2018) A Green BODIPY-Based, Super-Fluorogenic, Protein-Specific Labelling Agent. *Angew. Chem. Int. Ed. Engl.* 57, 12390–12394.

(45) Strmiskova, M. Development of a Dimaleimide-Based Protein Labelling Technique for Fluorogenic, X-Ray Crystallography and NMR Applications, University of Ottawa, (2016).

(46) Ma, Y.; Luo, W.; Quinn, P. J.; Liu, Z.; Hider, R. C. (2004) Design, Synthesis, Physicochemical Properties, and Evaluation of Novel Iron Chelators with Fluorescent Sensors. *J. Med. Chem.* 47, 6349–6362.

(47) Strmiskova, M.; Tsao, K.; Keillor, J. W. (2018) Rational Design of a Highly Reactive Dicysteine Peptide Tag for Fluorogenic Protein Labelling. *Org. Biomol. Chem.* 16, 6332–6340.

2.7. Chapter 2 Appendices

4/15/2020 Rightslink® by Copyright Clearance Center

Copyright Clearance Center RightsLink®

Home | Help | Email Support | Kelvin Tsao

ACS Publications
Most Trusted. Most Cited. Most Read.

Ring Substituent Effects on the Thiol Addition and Hydrolysis Reactions of N-Arylmaleimides

Author: Yingche Chen, Kelvin Tsao, Élise De Francesco, et al
Publication: The Journal of Organic Chemistry
Publisher: American Chemical Society
Date: Dec 1, 2015

Copyright © 2015, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

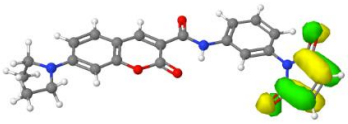
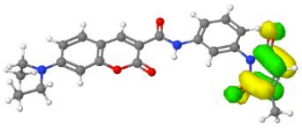
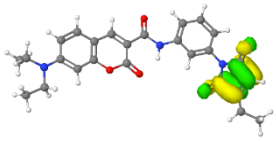
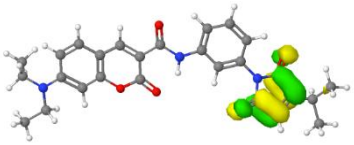
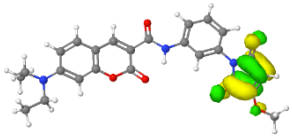
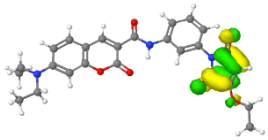
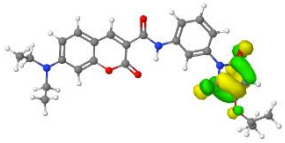
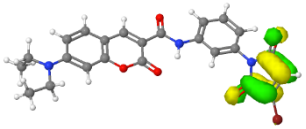
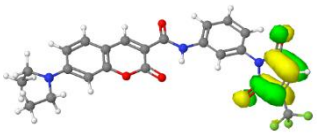
This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

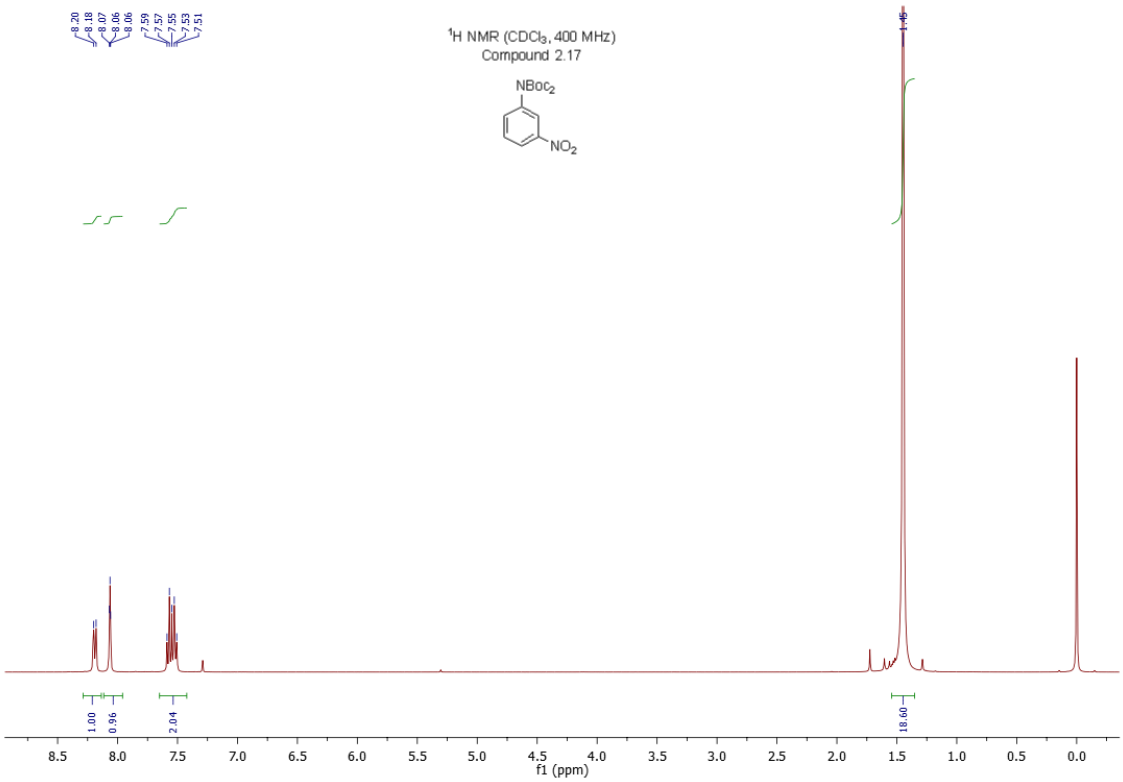
- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

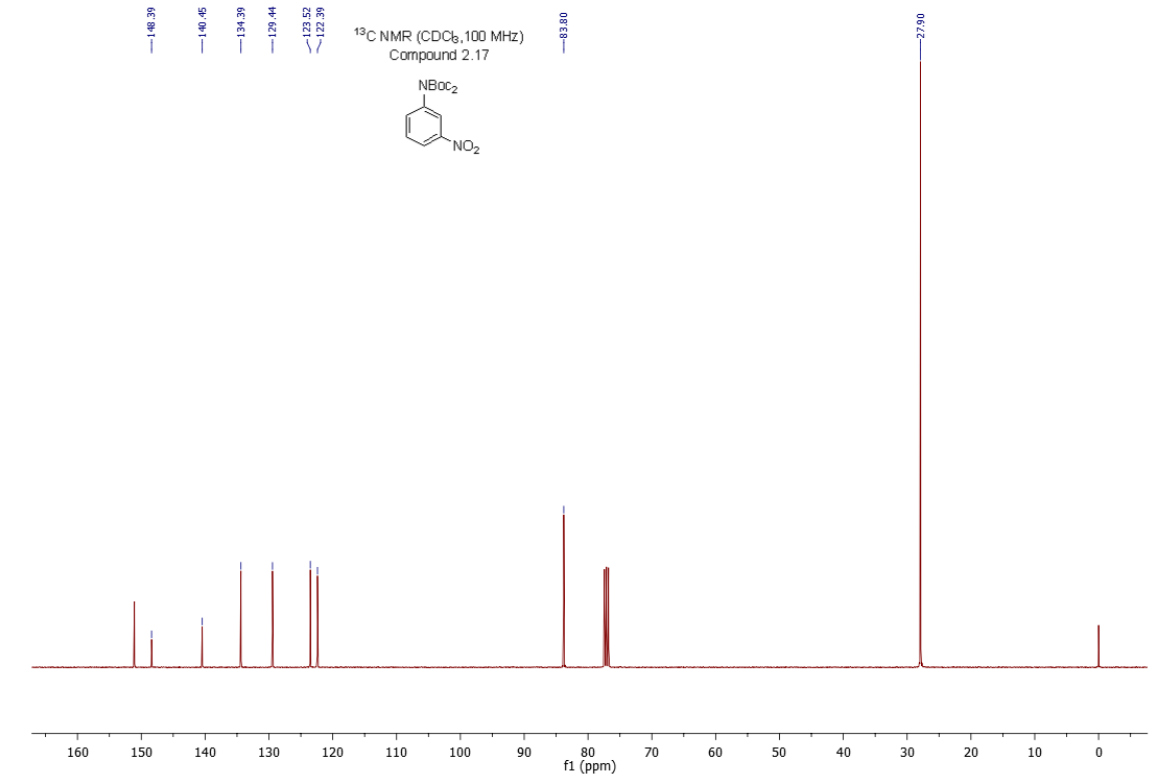
BACK CLOSE WINDOW

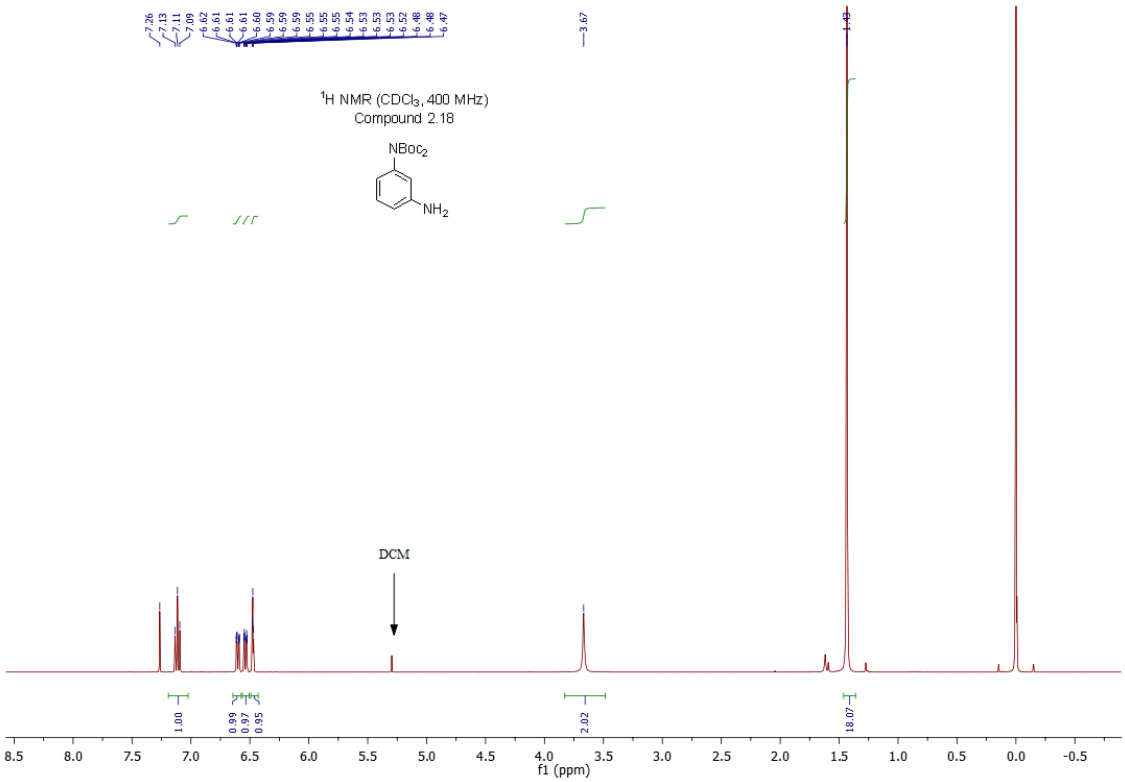
© 2020 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

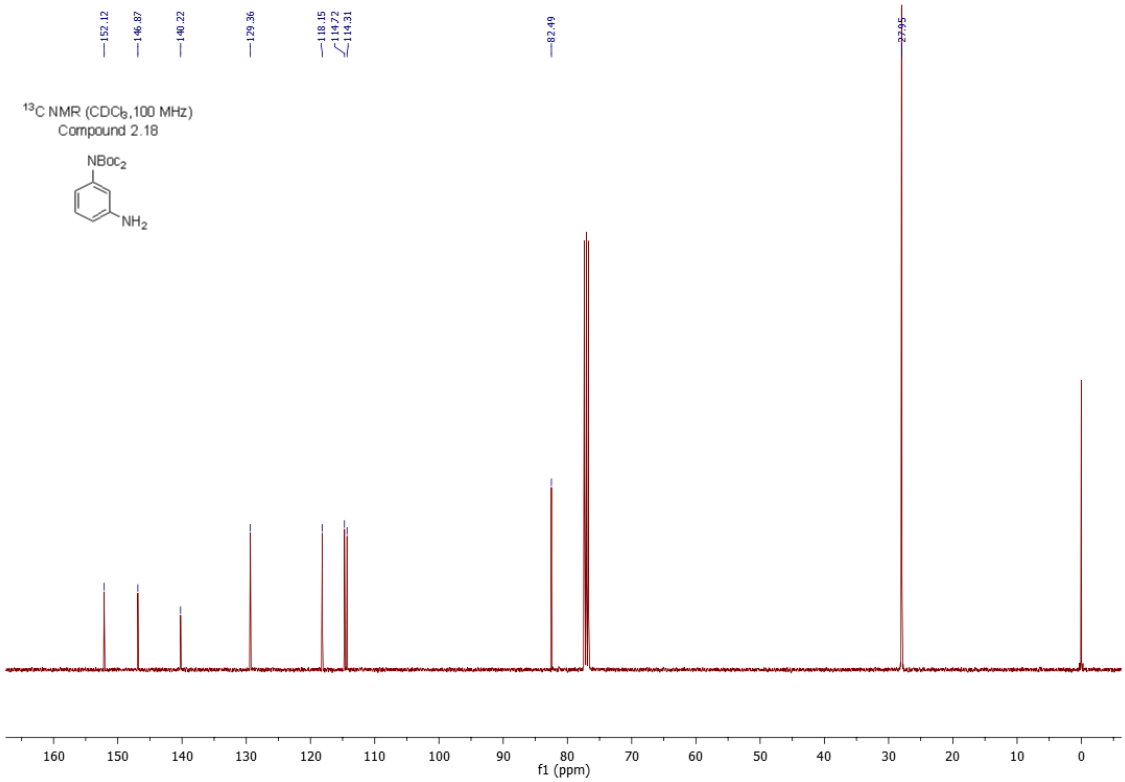
Table 2.7. Hammett σ^+ parameters and LUMO energy levels of substrates **S1**-**S9**. Gaussian 09 at B3LYP/6-31G(d) level of theory were used to perform DFT calculations, which provided the energy levels for the LUMO of the substrates studied in the thiol addition reaction. For Cartesian coordinates and total energy of each minimized structure, see reference 1.

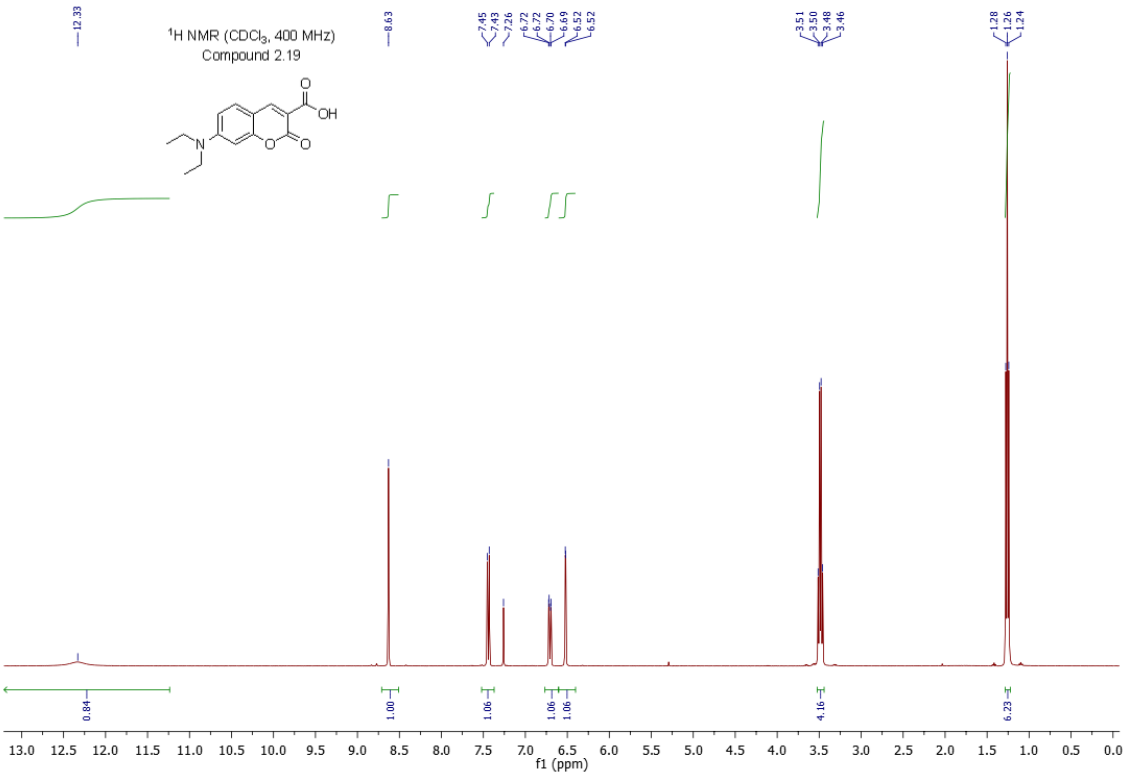
	R	σ^+	Density map of LUMO	Energy (Hartree)
S1	H	0		-0.09921
S2	Me	-0.31		-0.09253
S3	Et	-0.30		-0.09166
S4	iPr	-0.28		-0.09261
S5	OMe	-0.78		-0.08557
S6	OEt	-0.81		-0.08486
S7	OiPr	-0.85		-0.08330
S8	Br	0.15		-0.10565
S9	CF_3	0.61		-0.11380

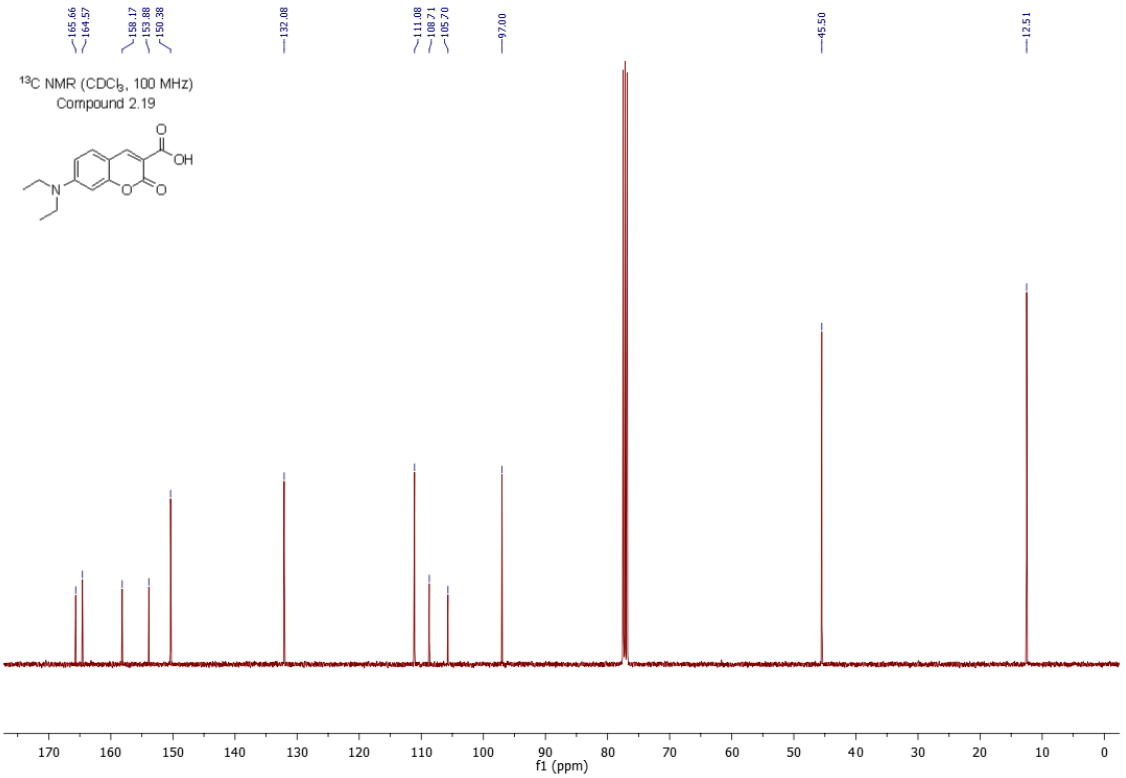

8.20
8.18
8.07
8.06
8.06
7.59
7.57
7.55
7.53
7.51
1.48
^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 2.17
NBoc$_2$
NO$_2$
1.00
0.96
2.04
18.60
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

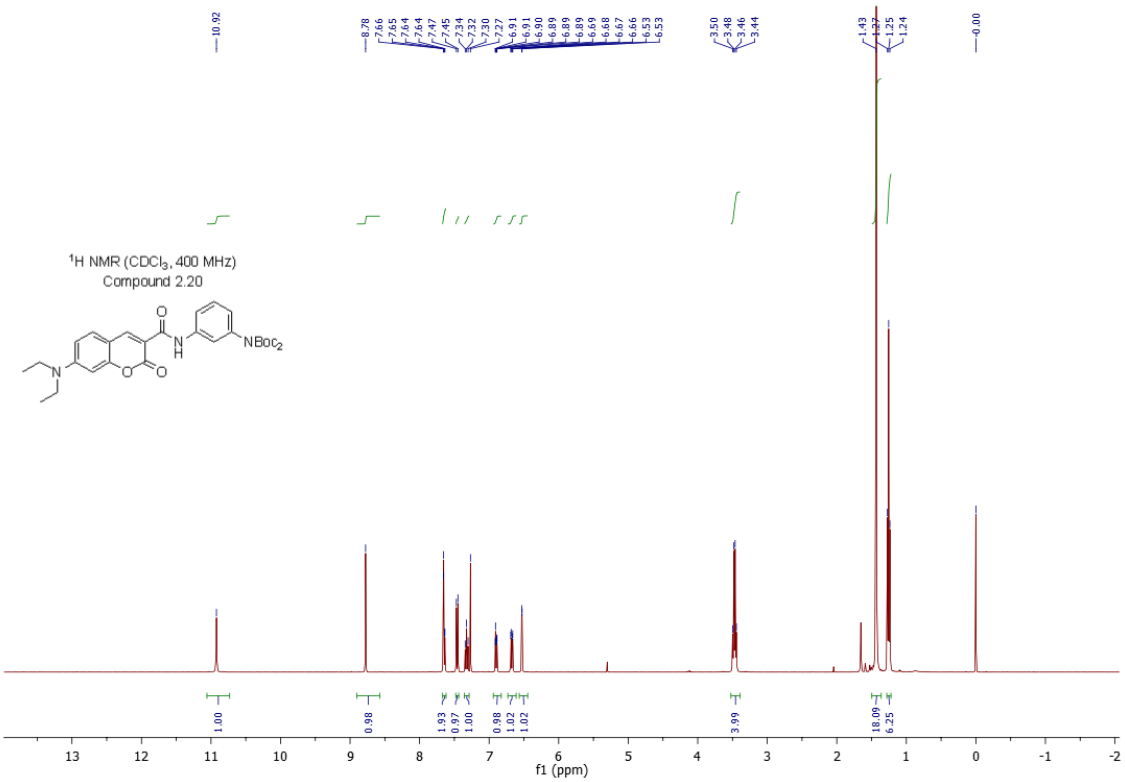

148.39
140.45
134.39
129.44
123.52
122.39
83.80
27.90
^{13}C NMR ($CDCl_3$,100 MHz)
Compound 2.17
NBoc$_2$
NO$_2$
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

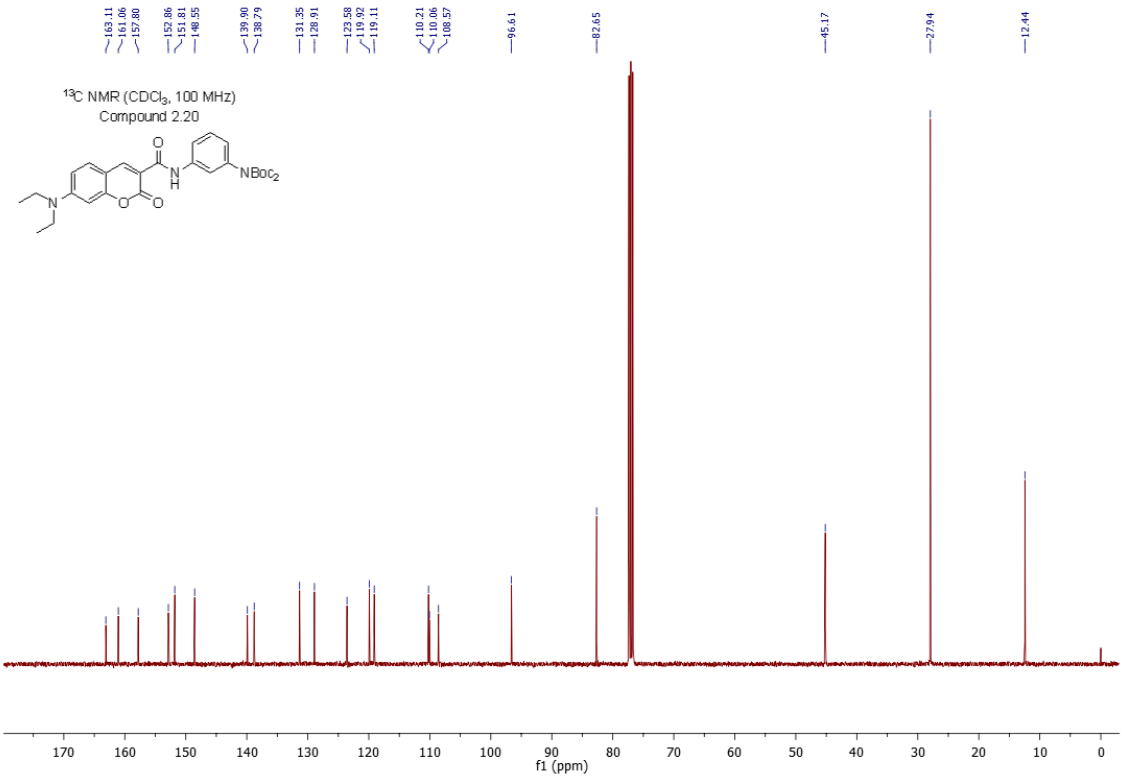

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 2.18
$NBoc_2$
NH_2
DCM
f1 (ppm)

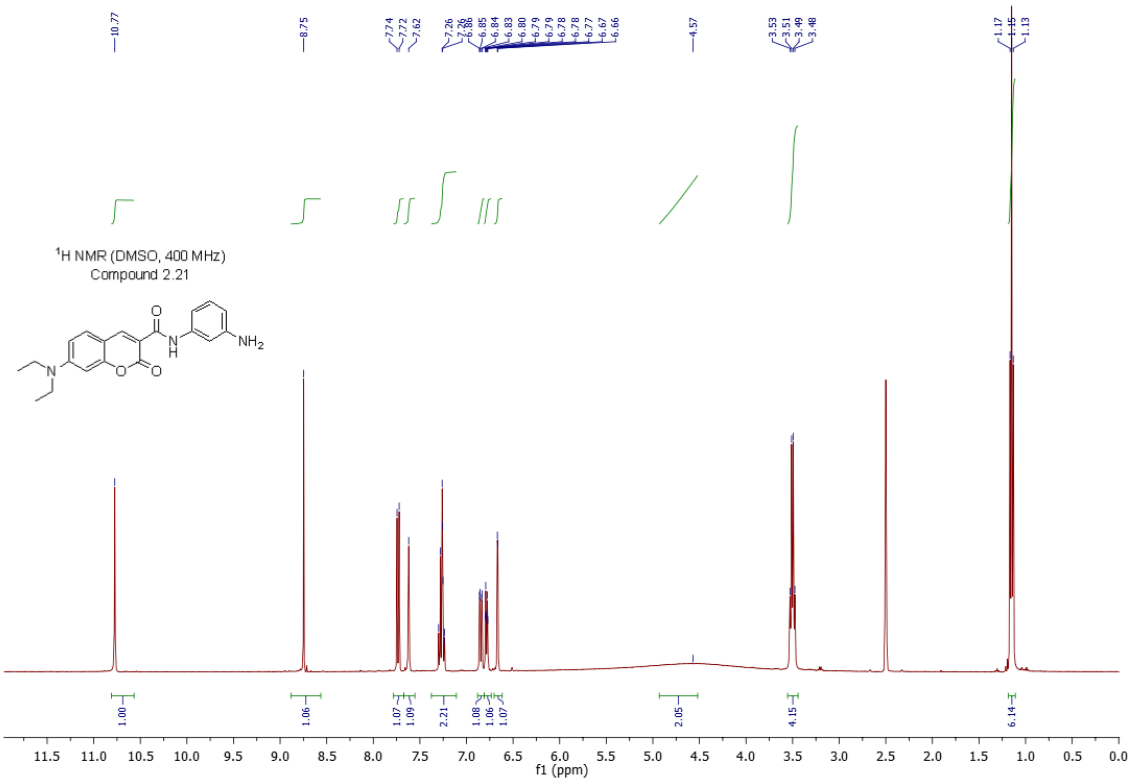

^{13}C NMR ($CDCl_3$,100 MHz)
Compound 2.18
$NBoc_2$
NH_2
f1 (ppm)

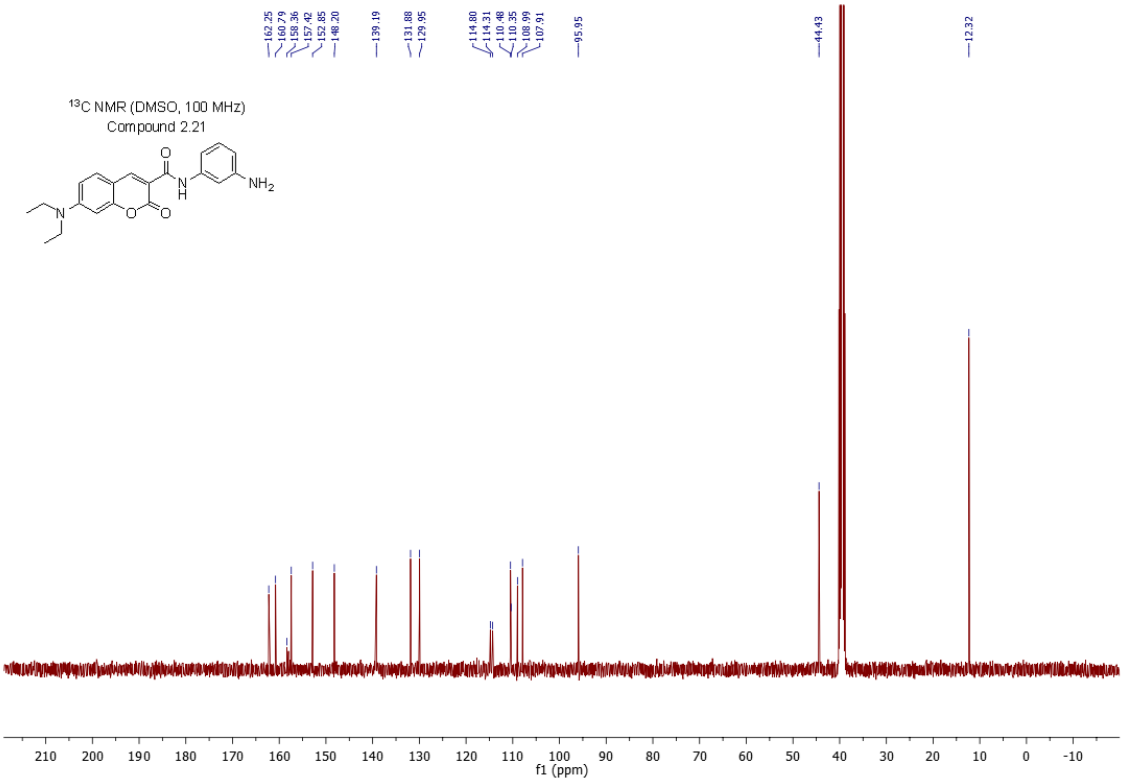

12.33
8.63
7.45
7.43
7.26
6.72
6.72
6.70
6.69
6.52
6.52
3.51
3.50
3.48
3.46
1.28
1.26
1.24
^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 2.19
O
OH
N
O
O
0.84
1.00
1.06
1.06
1.06
4.16
6.23
13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

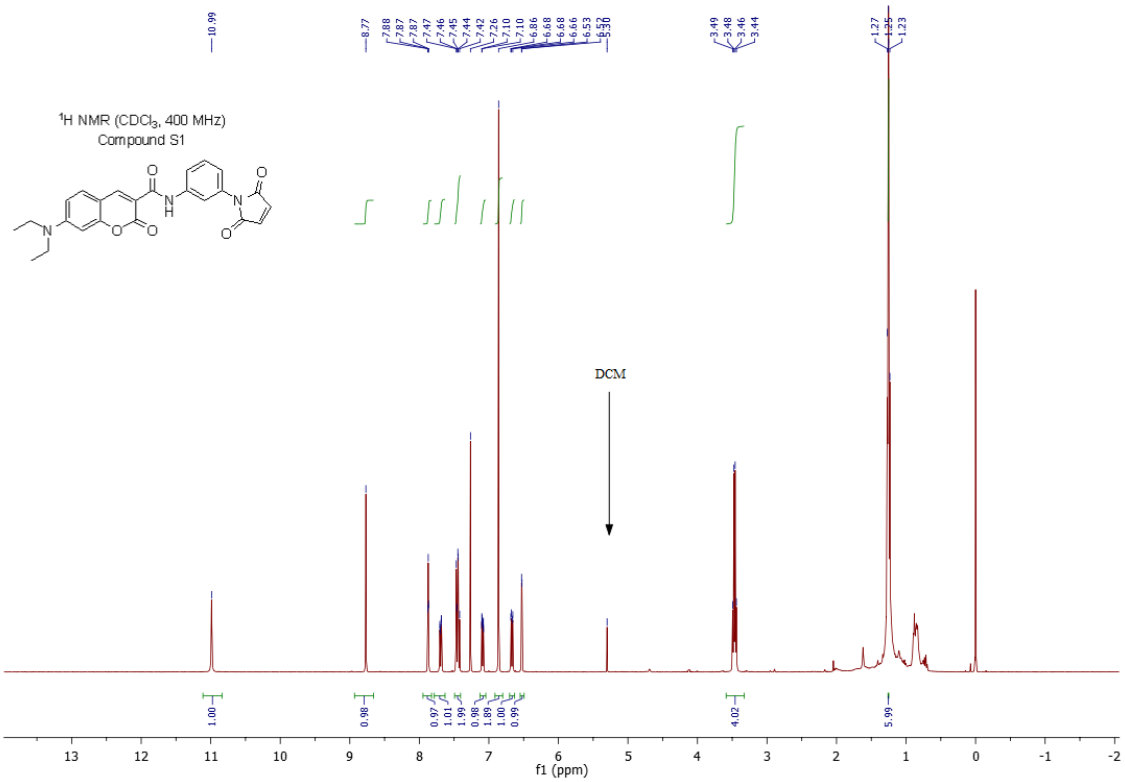

165.66
164.57
158.17
153.88
150.38
132.08
111.08
108.71
105.70
97.00
45.50
12.51
^{13}C NMR ($CDCl_3$, 100 MHz)
Compound 2.19
O
OH
N
O
O
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

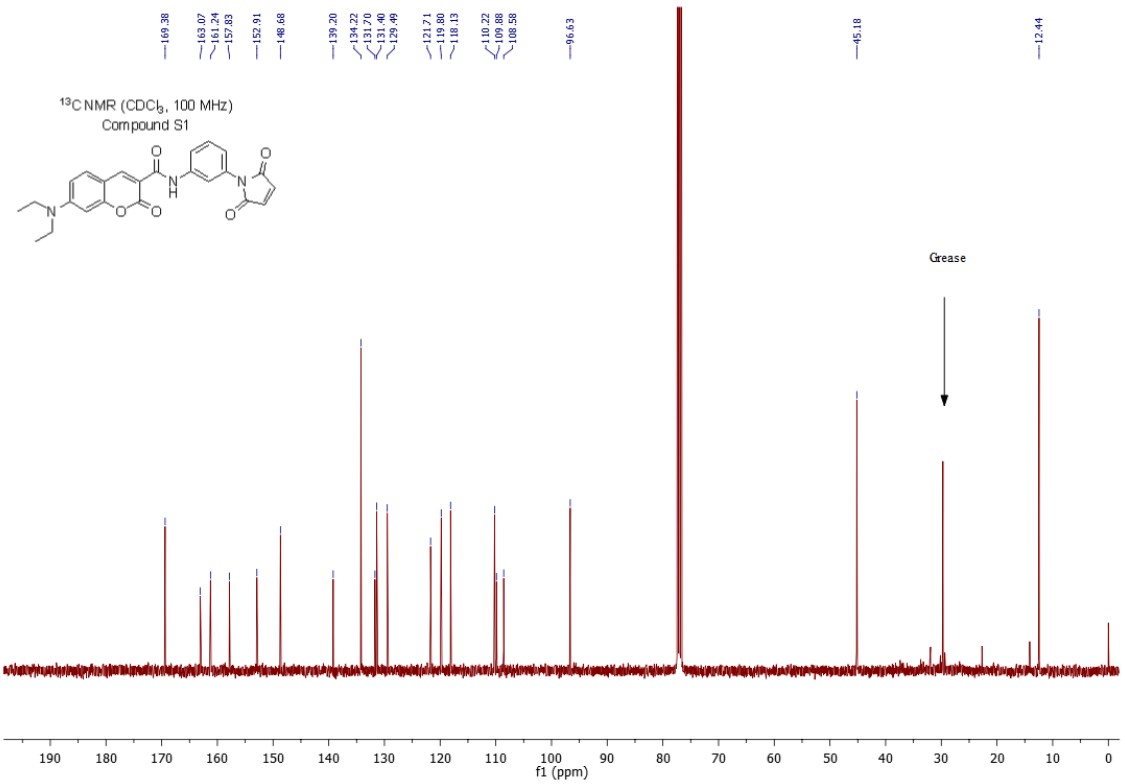

10.92
8.78
7.66
7.65
7.64
7.64
7.47
7.45
7.34
7.32
7.30
7.27
6.91
6.91
6.90
6.89
6.89
6.89
6.69
6.68
6.67
6.66
6.53
6.53
3.50
3.48
3.46
3.44
1.43
1.37
1.25
1.24
0.00
1H NMR (CDCl3, 400 MHz)
Compound 2.20
NBoc2
1.00
0.98
1.93
0.97
1.00
0.98
1.02
1.02
3.99
18.09
6.25
f1 (ppm)

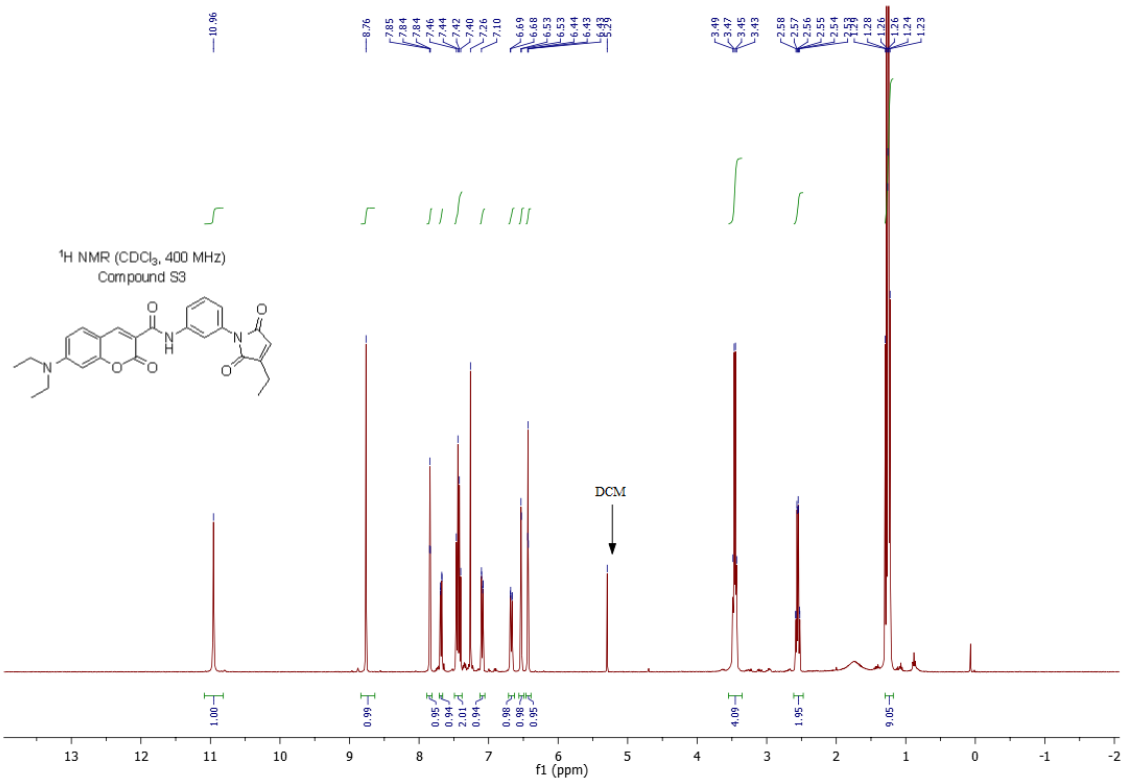

163.11
161.06
157.80
152.86
151.81
148.55
139.90
138.79
131.35
128.91
123.58
119.92
119.11
110.21
110.06
108.57
96.61
82.65
45.17
27.94
12.44
13C NMR (CDCl3, 100 MHz)
Compound 2.20
NBoc2
f1 (ppm)

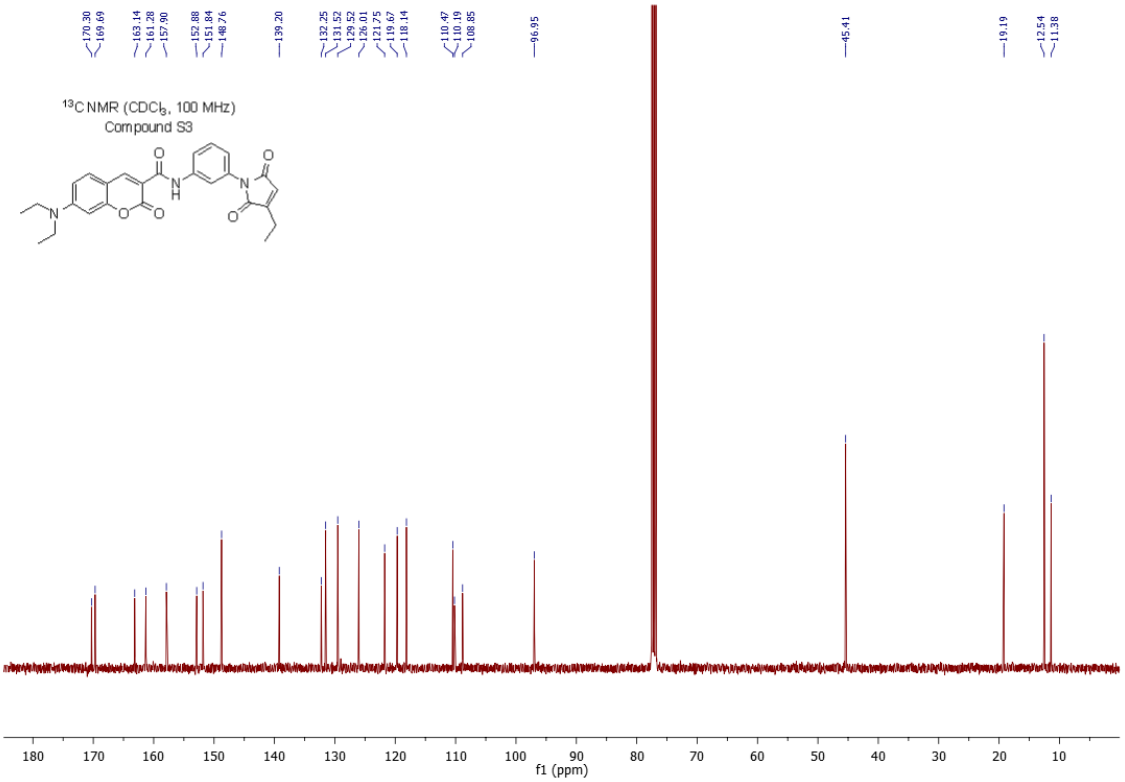

10.77
8.75
7.74
7.72
7.62
7.26
6.85
6.84
6.83
6.80
6.79
6.79
6.78
6.78
6.77
6.67
6.66
4.57
3.53
3.51
3.49
3.48
1.17
1.15
1.13
1H NMR (DMSO, 400 MHz)
Compound 2.21
O
N
H
NH2
N
O
O
1.00
1.06
1.07
1.09
2.21
1.08
1.06
1.07
2.05
4.15
6.14
11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

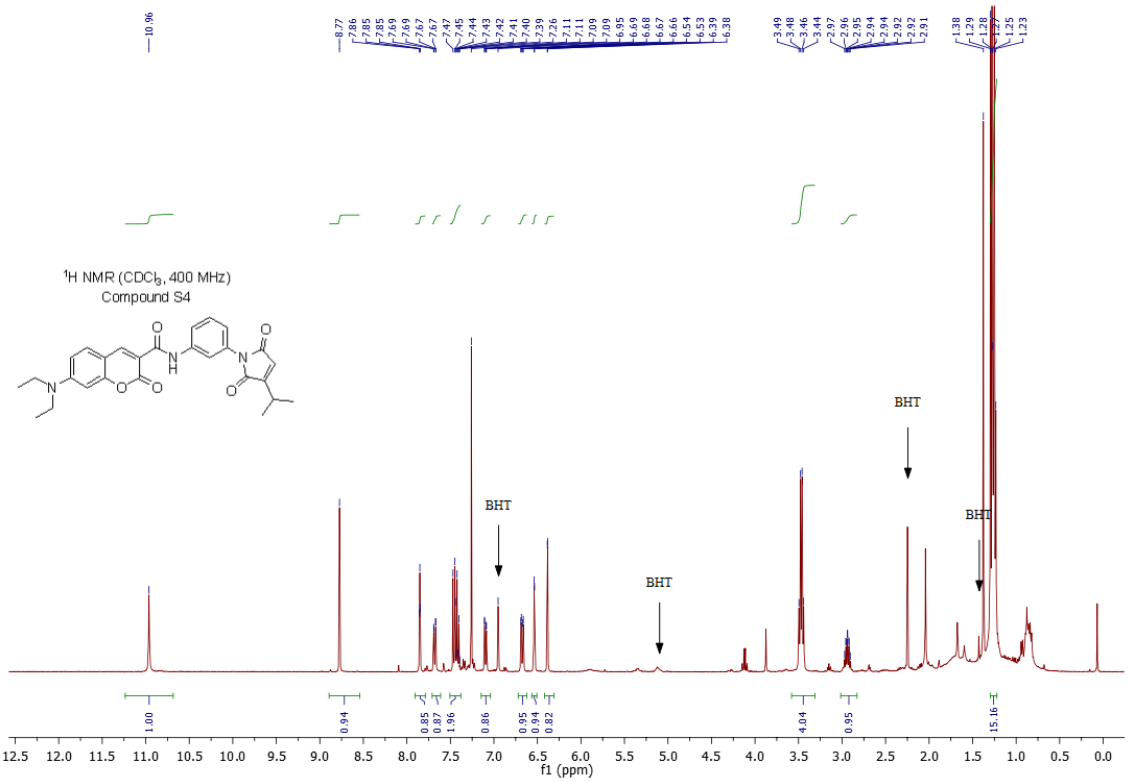

162.25
160.79
158.36
157.42
152.85
148.20
139.19
131.88
129.95
114.80
114.31
110.48
110.35
108.99
107.91
95.95
44.43
12.32
13C NMR (DMSO, 100 MHz)
Compound 2.21
O
N
H
NH2
N
O
O
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

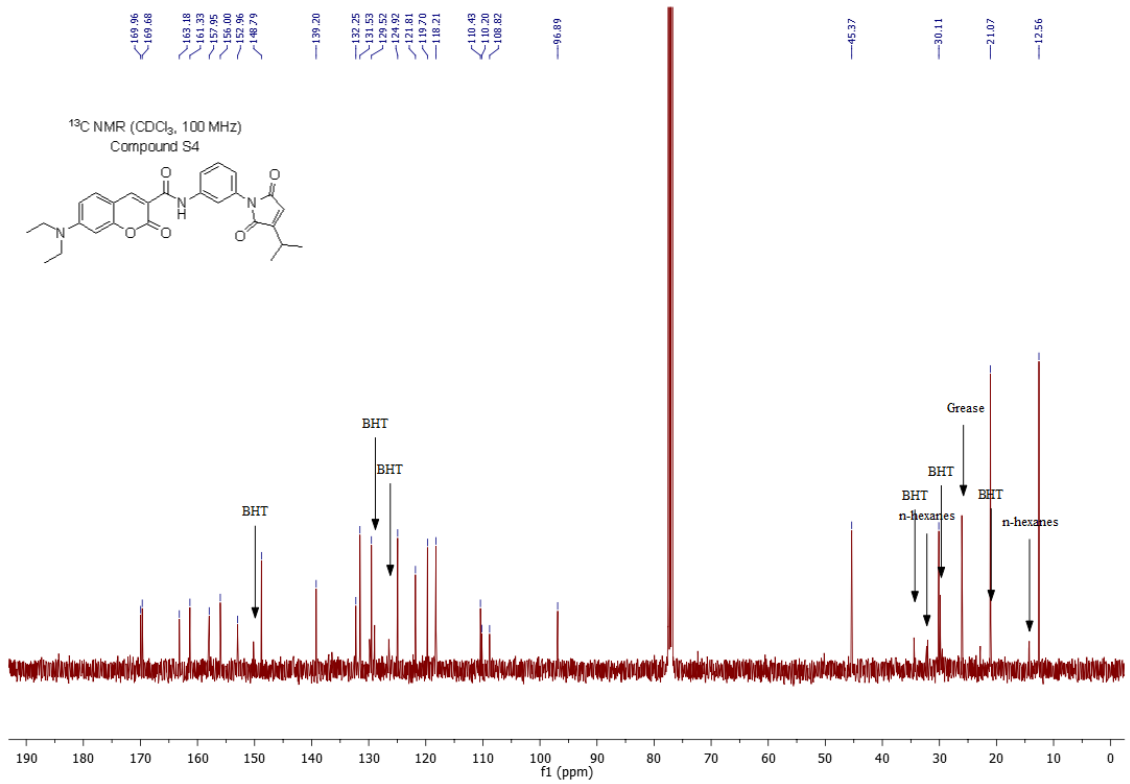

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound S1
DCM
f1 (ppm)

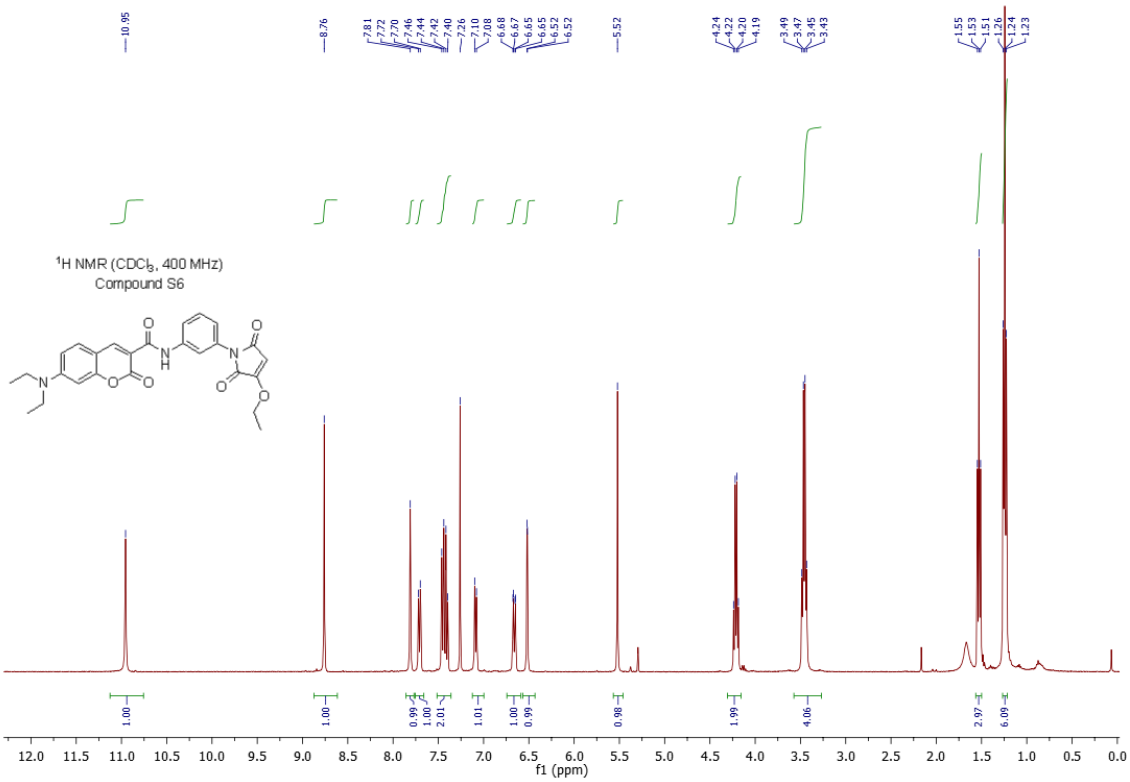

^{13}C NMR ($CDCl_3$, 100 MHz)
Compound S1
Grease
f1 (ppm)

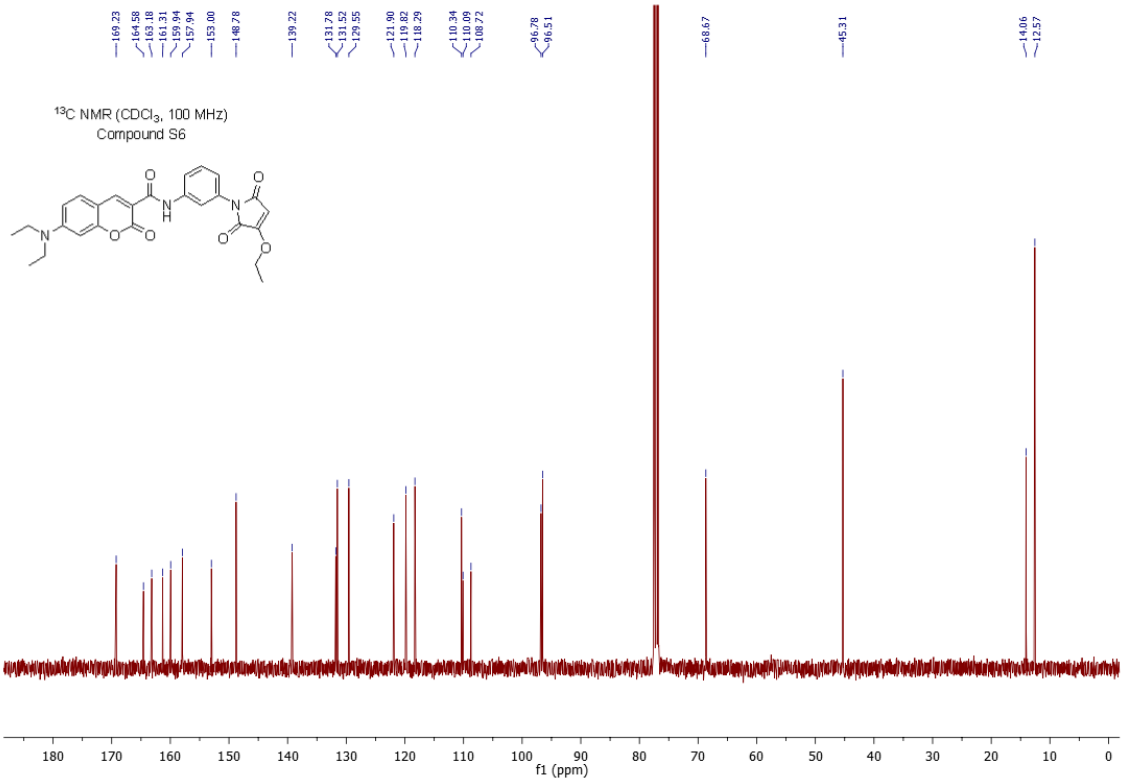

10.96
8.76
7.85
7.84
7.84
7.46
7.44
7.42
7.40
7.26
7.10
6.69
6.68
6.53
6.53
6.44
6.43
6.38
3.49
3.47
3.45
3.43
2.58
2.57
2.56
2.55
2.54
2.53
1.28
1.26
1.26
1.24
1.23
^{1}H NMR ($CDCl_3$, 400 MHz)
Compound S3
DCM
1.00
0.99
0.95
0.94
2.01
0.94
0.98
0.98
0.95
4.09
1.95
9.05
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
f1 (ppm)

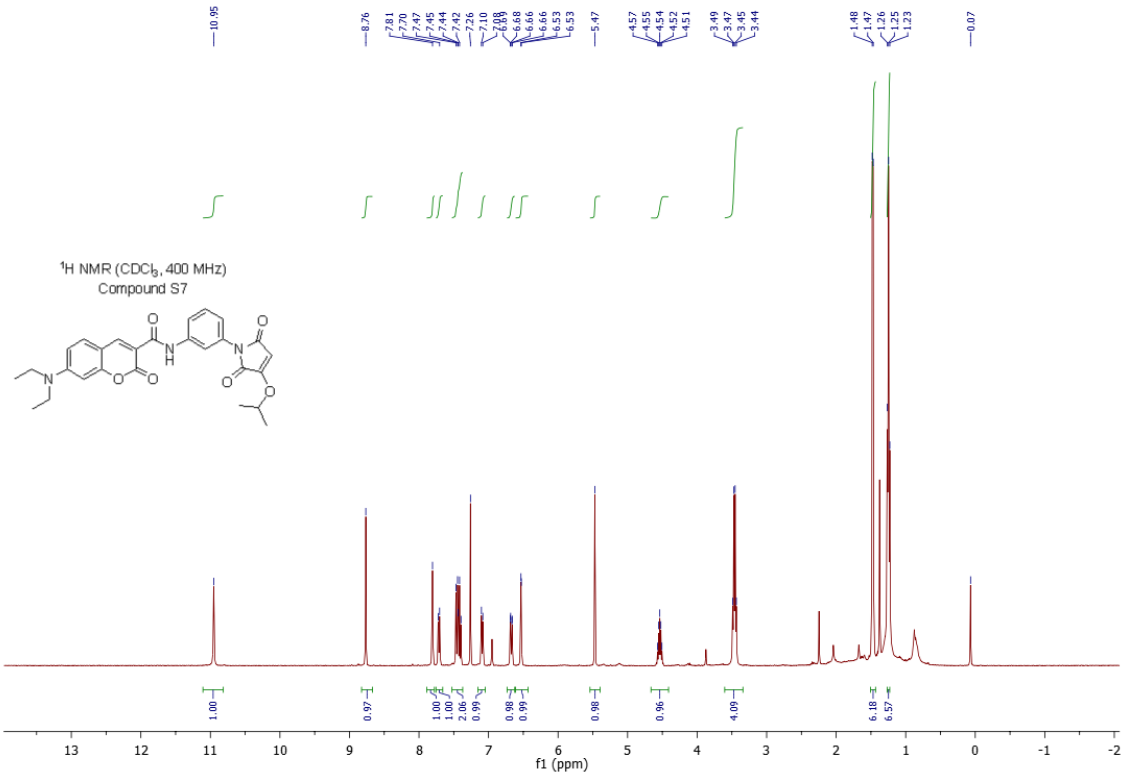

170.30
169.69
163.14
161.28
157.90
152.88
151.84
148.76
139.20
132.25
131.52
129.52
126.01
121.75
119.67
118.14
110.47
110.19
108.85
96.95
45.41
19.19
12.54
11.38
^{13}C NMR ($CDCl_3$, 100 MHz)
Compound S3
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)

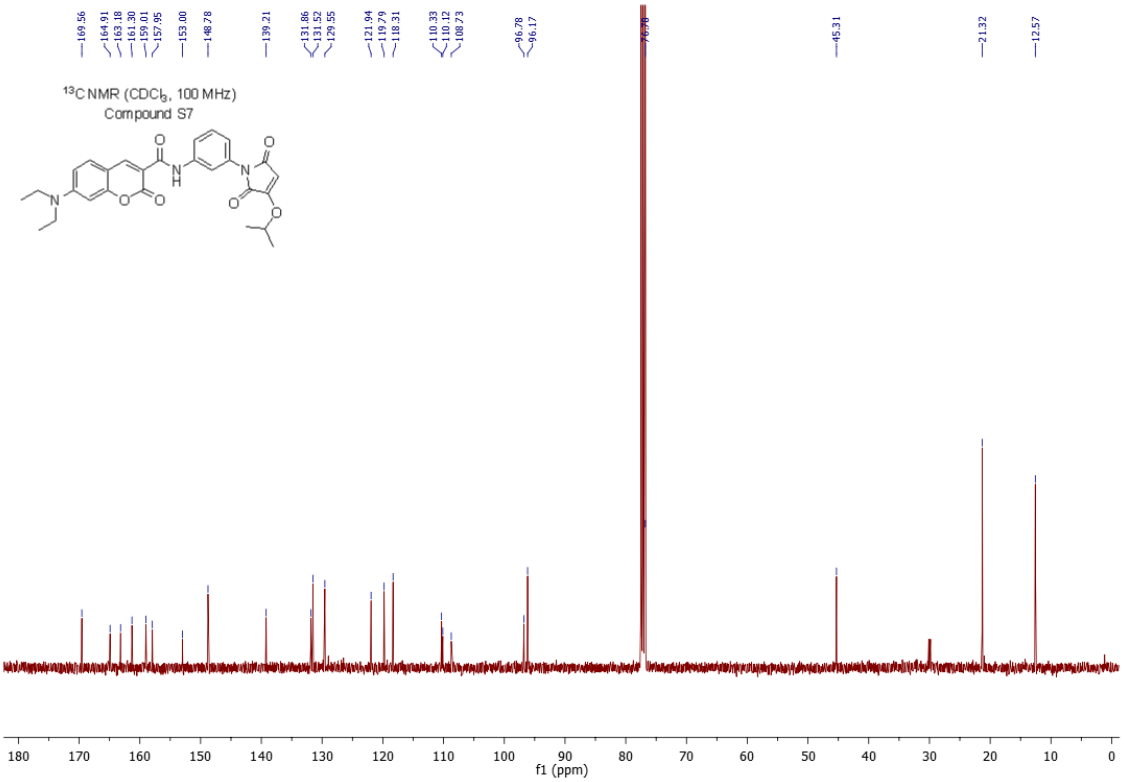

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound S4
BHT
BHT
BHT
BHT
f1 (ppm)

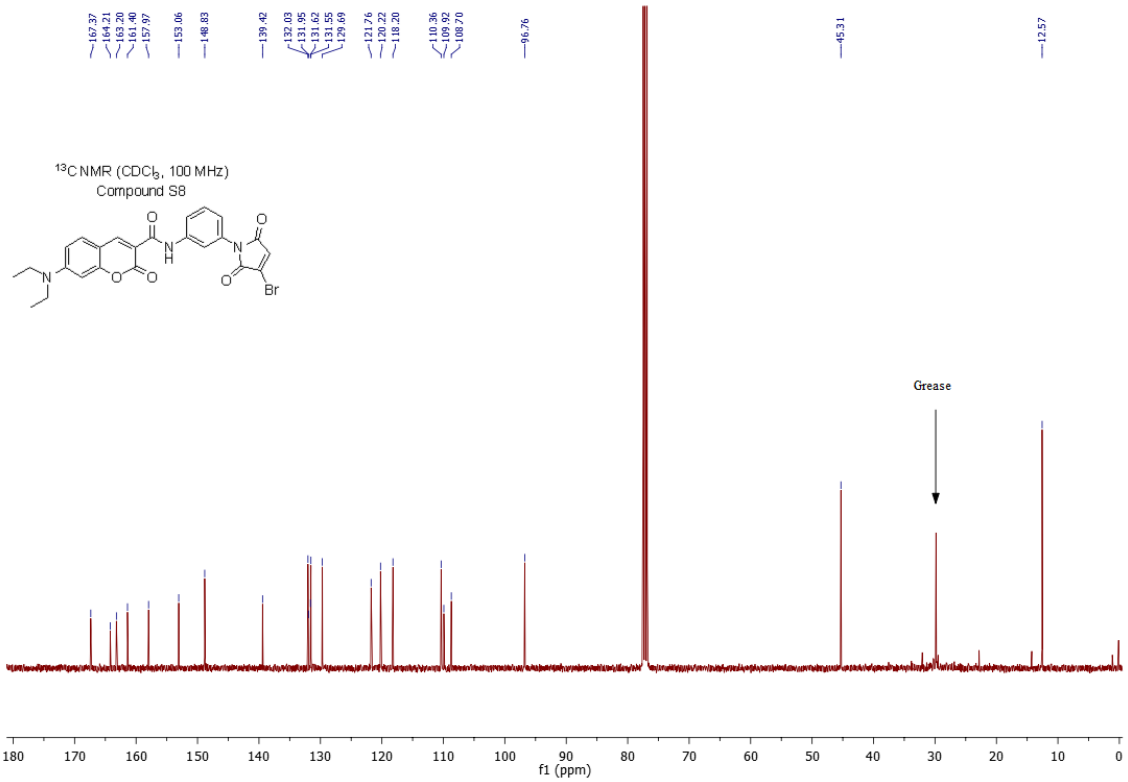

^{13}C NMR ($CDCl_3$, 100 MHz)
Compound S4
BHT
BHT
BHT
BHT
BHT
BHT
Grease
n-hexanes
n-hexanes
f1 (ppm)

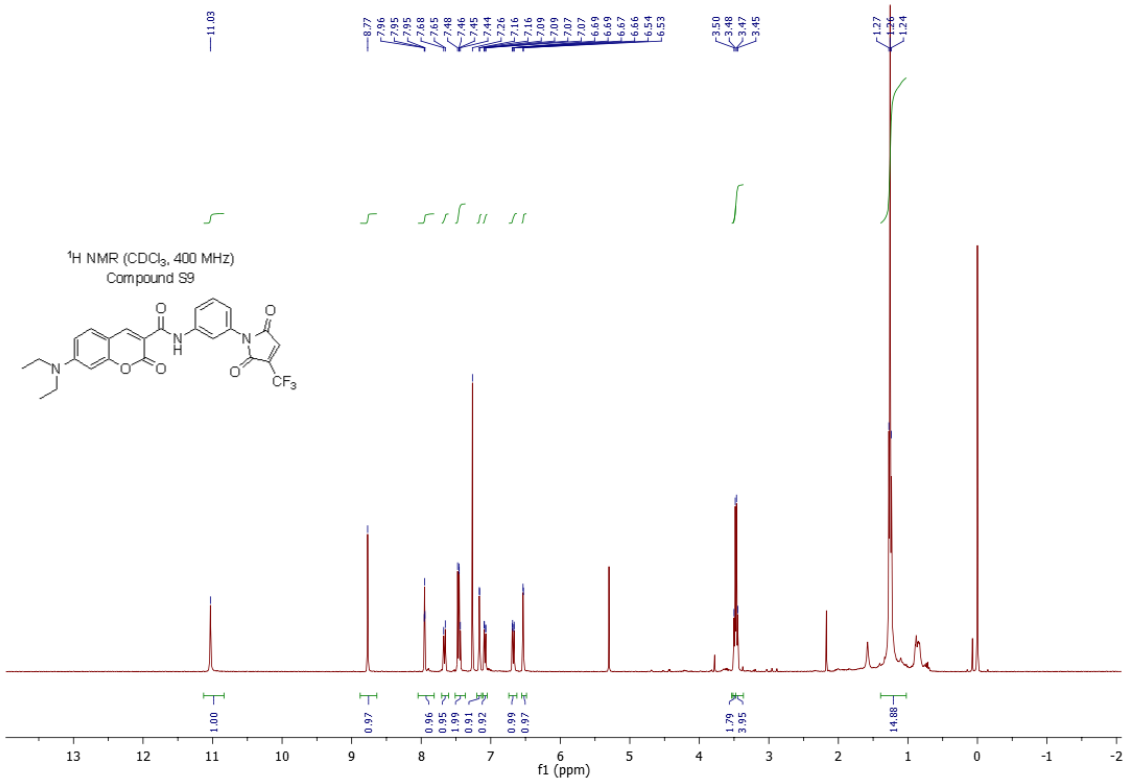

10.95
8.76
7.81
7.72
7.70
7.46
7.44
7.42
7.40
7.26
7.10
7.08
6.68
6.67
6.65
6.65
6.52
6.52
5.52
4.24
4.22
4.20
4.19
3.49
3.47
3.45
3.43
1.55
1.53
1.51
1.26
1.24
1.23
1H NMR (CDCl3, 400 MHz)
Compound S6
1.00
1.00
0.99
1.00
2.01
1.01
1.00
0.99
0.98
1.99
4.06
2.97
6.09
f1 (ppm)

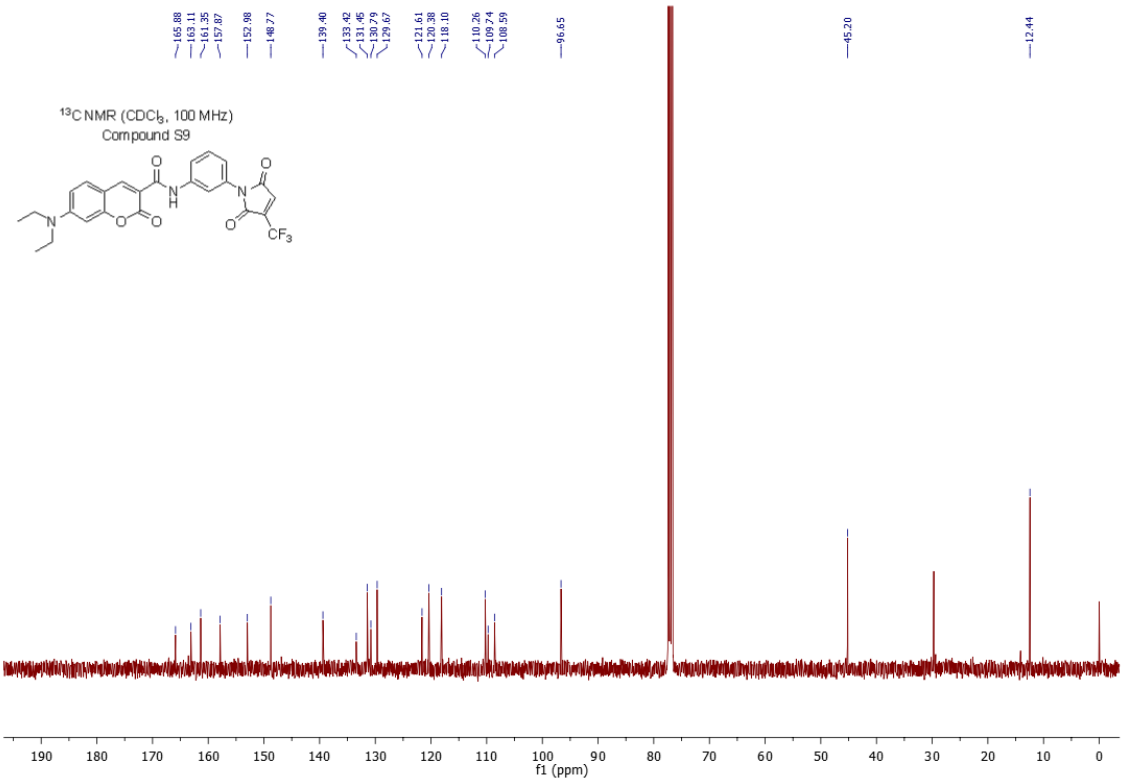

169.23
164.58
163.18
161.31
159.94
157.94
153.00
148.78
139.22
131.78
131.52
129.55
121.90
119.82
118.29
110.34
110.09
108.72
96.78
96.51
68.67
45.31
14.06
12.57
13C NMR (CDCl3, 100 MHz)
Compound S6
f1 (ppm)

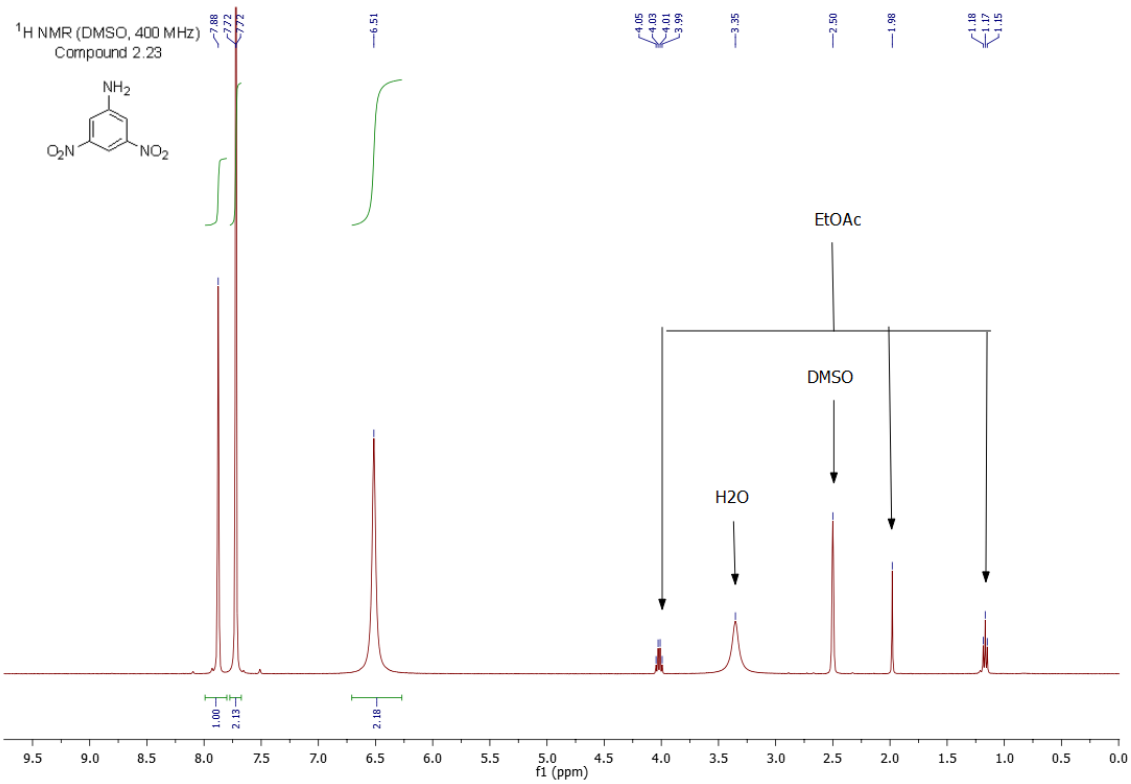

10.95
8.76
7.81
7.70
7.47
7.45
7.44
7.42
7.26
7.10
7.08
6.89
6.68
6.66
6.66
6.53
6.53
5.47
4.57
4.55
4.54
4.52
4.51
3.49
3.47
3.45
3.44
1.48
1.47
1.26
1.25
1.23
0.07
1H NMR (CDCl3, 400 MHz)
Compound S7
1.00
0.97
1.00
1.00
2.06
0.99
0.98
0.99
0.98
0.96
4.09
6.18
6.57
f1 (ppm)

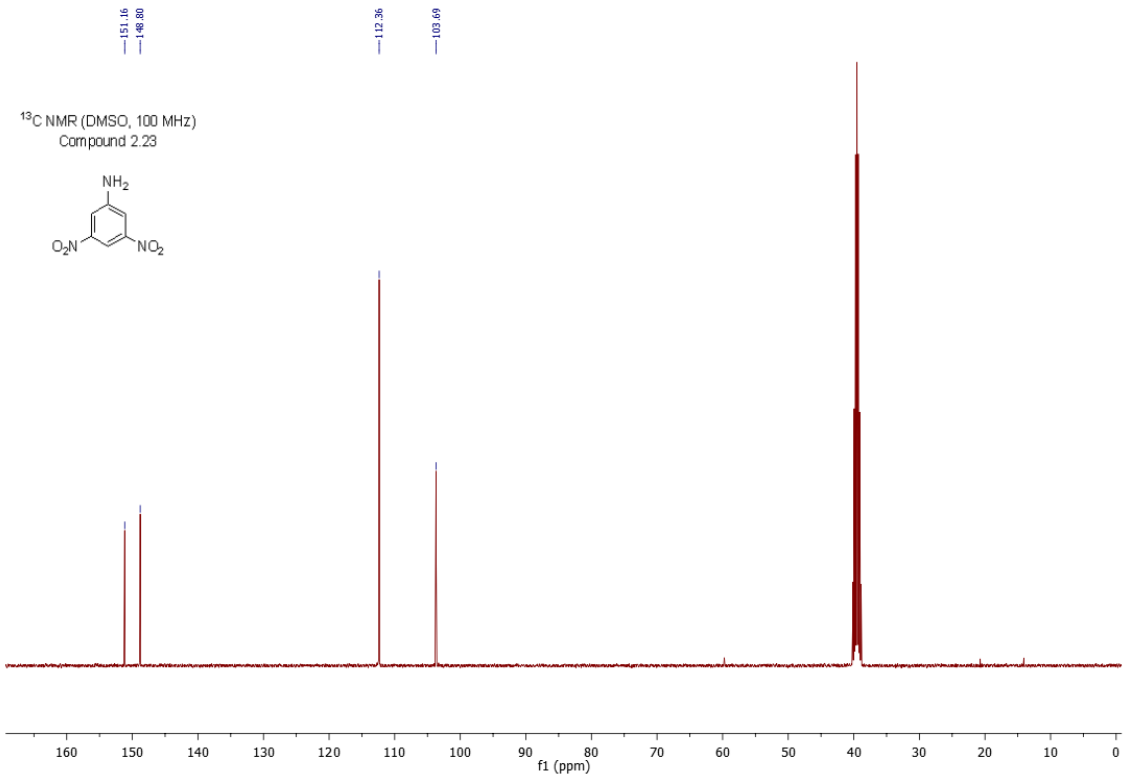

169.56
164.91
163.18
161.30
159.01
157.95
153.00
148.78
139.21
131.86
131.52
129.55
121.94
119.79
118.31
110.33
110.12
108.73
96.78
96.17
76.80
45.31
21.32
12.57
13C NMR (CDCl3, 100 MHz)
Compound S7
f1 (ppm)

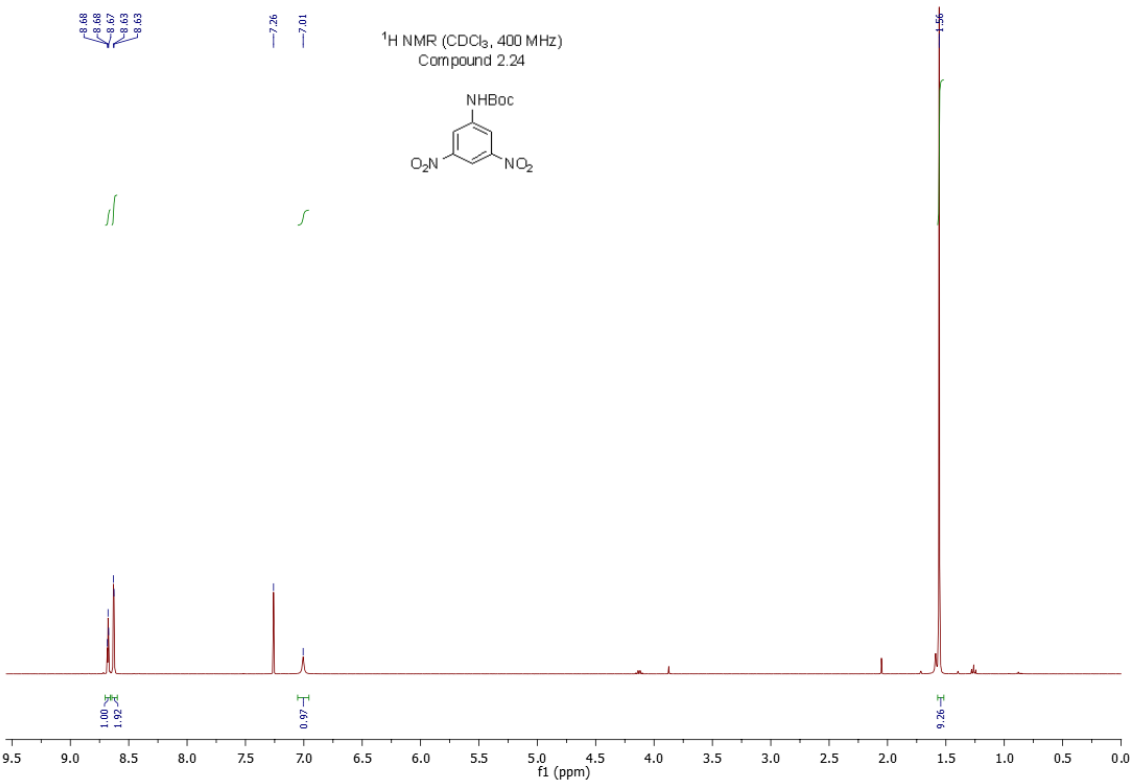

167.37
164.21
163.20
161.40
157.97
153.06
148.83
139.42
132.03
131.95
131.62
131.55
129.69
121.76
120.22
118.20
110.36
109.92
108.70
96.76
45.31
12.57
13C NMR (CDCl3, 100 MHz)
Compound S8
Br
Grease
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

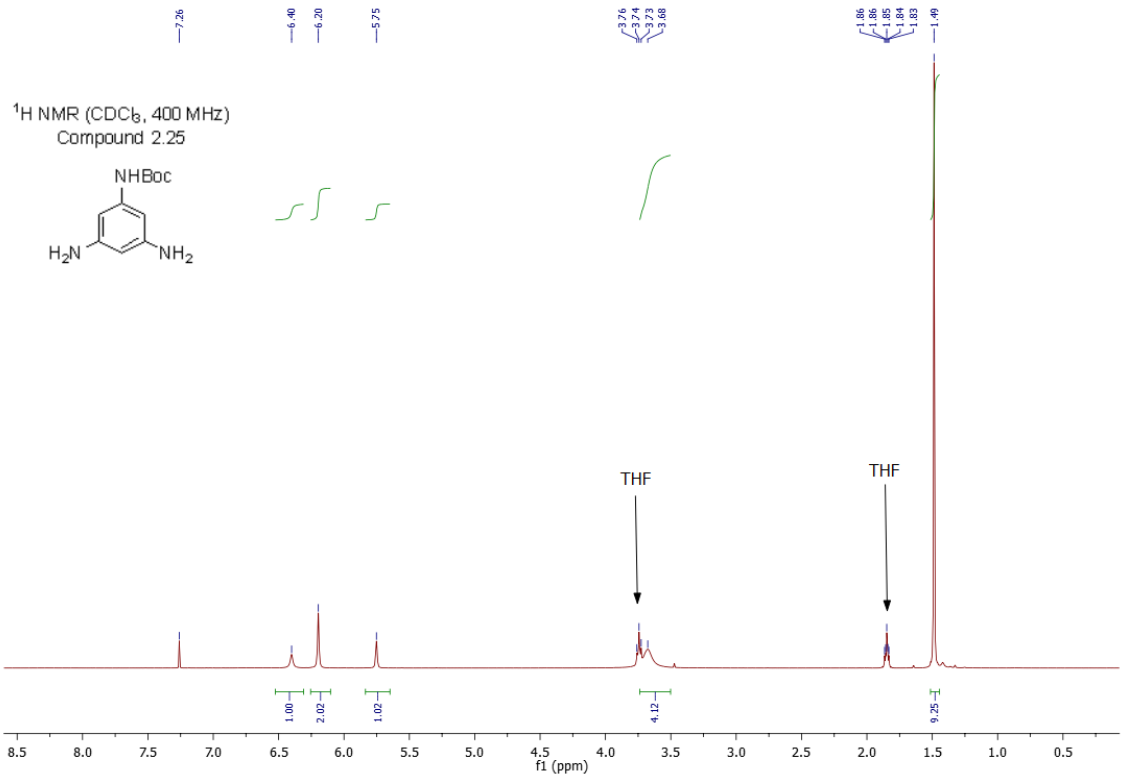

11.03
8.77
7.96
7.95
7.95
7.88
7.65
7.48
7.46
7.45
7.44
7.26
7.16
7.16
7.09
7.09
7.07
7.07
6.69
6.69
6.67
6.66
6.54
6.53
3.50
3.48
3.47
3.45
1.27
1.26
1.24
1H NMR (CDCl3, 400 MHz)
Compound S9
CF3
1.00
0.97
0.96
0.95
1.99
0.91
0.92
0.99
0.97
1.79
3.95
14.88
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
f1 (ppm)

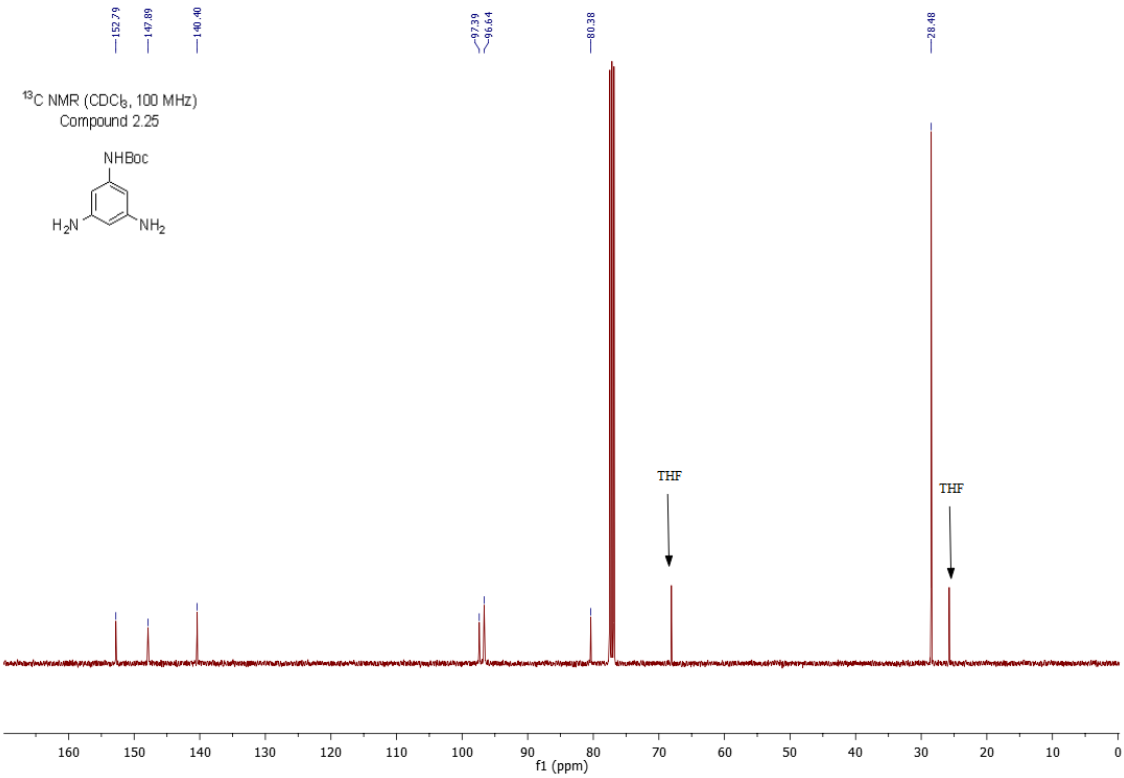

165.88
163.11
161.35
157.87
152.98
148.77
139.40
133.42
131.45
130.79
129.67
121.61
120.38
118.10
110.26
109.74
108.59
96.65
45.20
12.44
13C NMR (CDCl3, 100 MHz)
Compound S9
f1 (ppm)

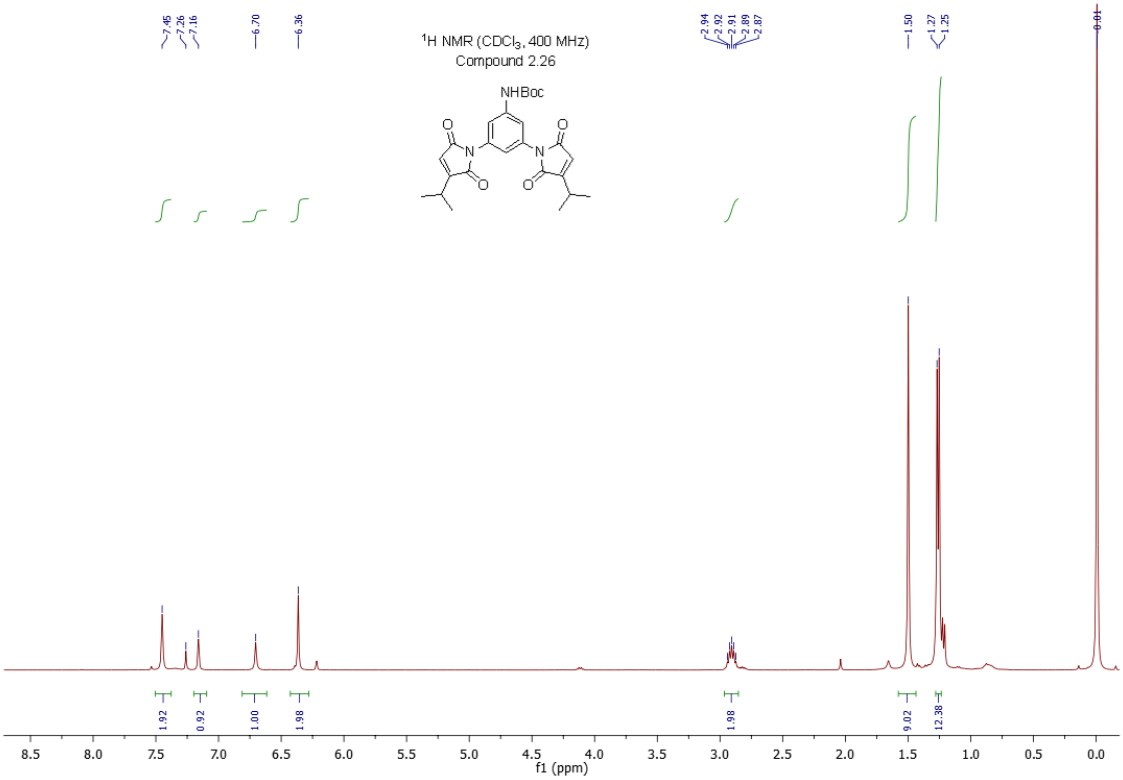

1H NMR (DMSO, 400 MHz)
Compound 2.23
7.88
7.72
7.72
6.51
4.05
4.03
4.01
3.99
3.35
2.50
1.98
1.18
1.17
1.15
EtOAc
DMSO
H2O
1.00
2.13
2.18
f1 (ppm)

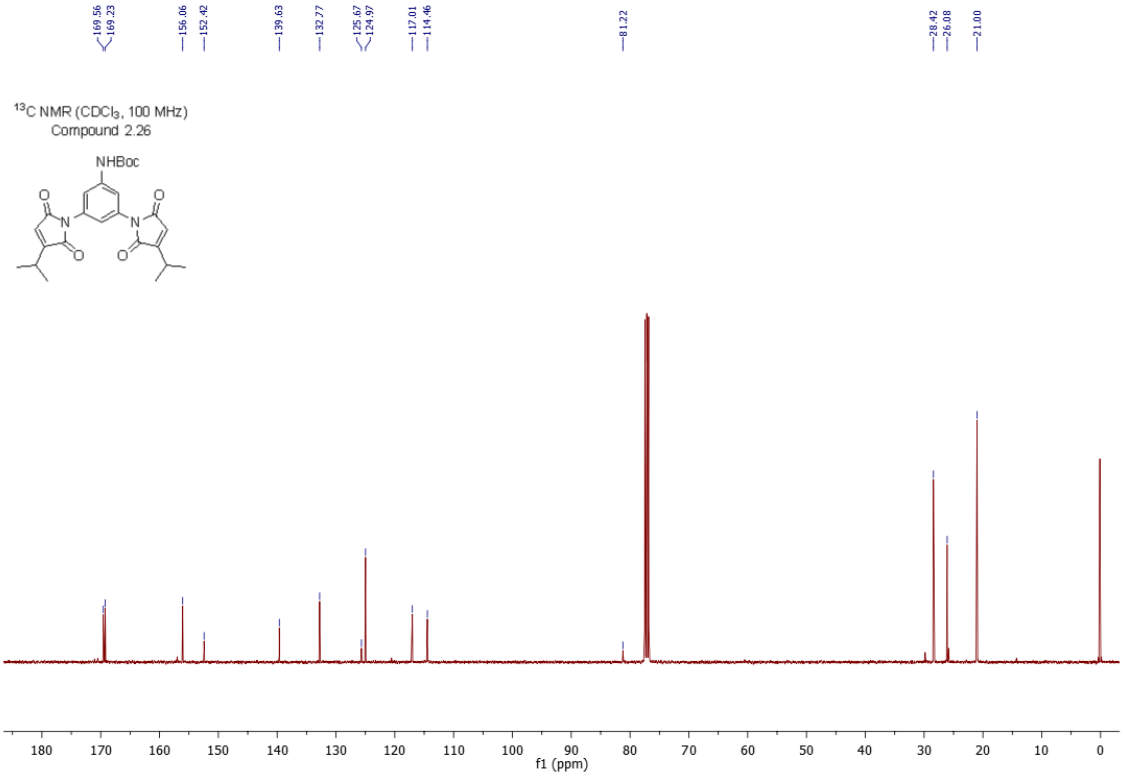

151.16
148.80
112.36
103.69
13C NMR (DMSO, 100 MHz)
Compound 2.23
NH2
O2N
NO2
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

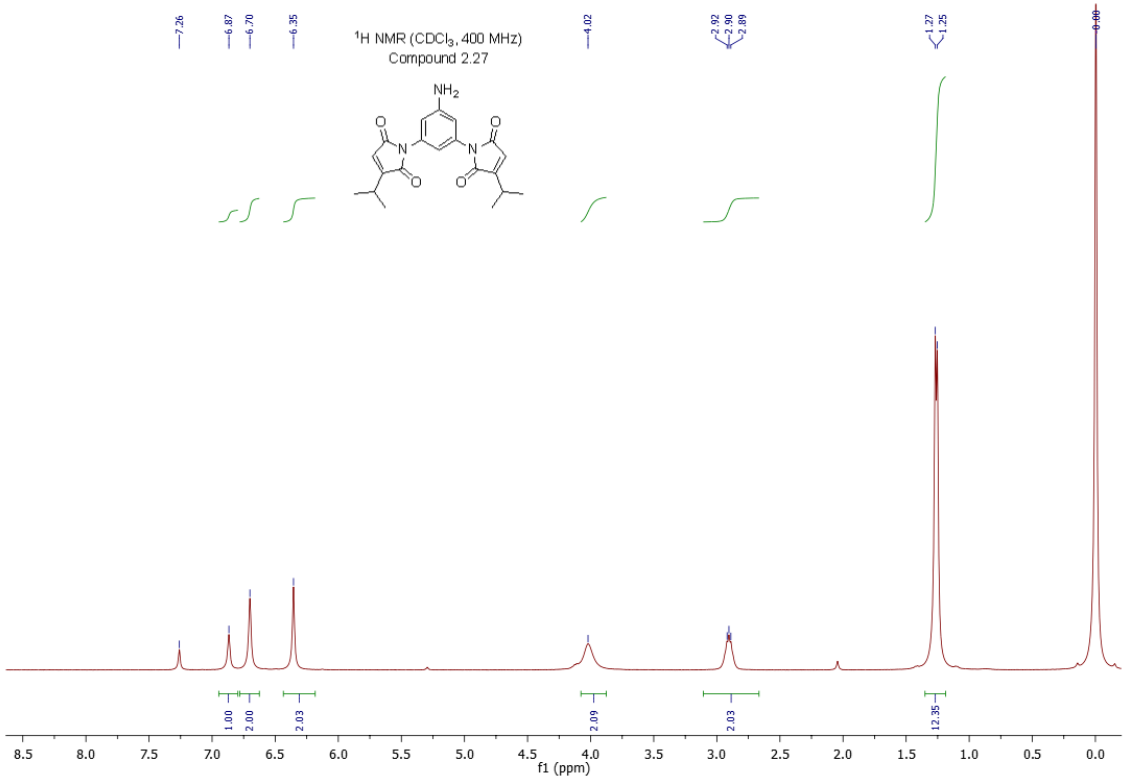

8.68
8.68
8.67
8.63
8.63
7.26
7.01
1.56
1H NMR (CDCl3, 400 MHz)
Compound 2.24
NHBoc
O2N
NO2
1.00
1.92
0.97
9.26
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

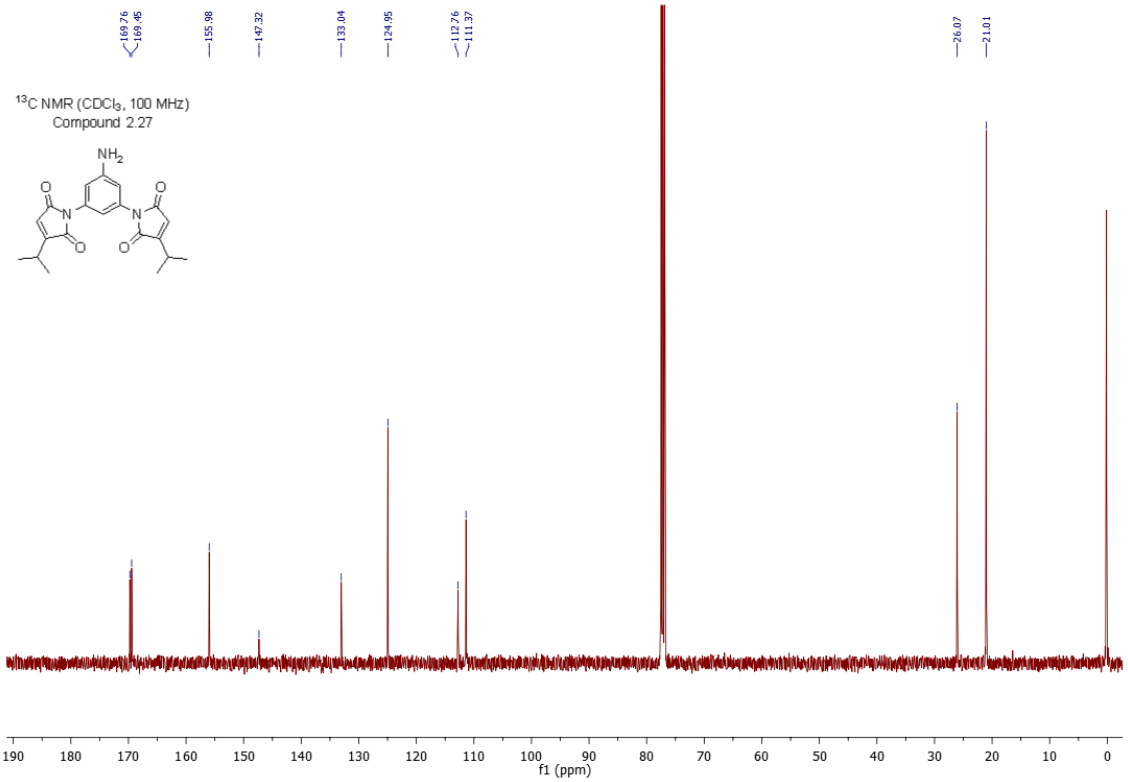

7.26
6.40
6.20
5.75
3.76
3.74
3.73
3.68
1.86
1.86
1.85
1.84
1.83
1.49
^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 2.25
NHBoc
H_2N
NH_2
THF
THF
1.00
2.02
1.02
4.12
9.25
f1 (ppm)

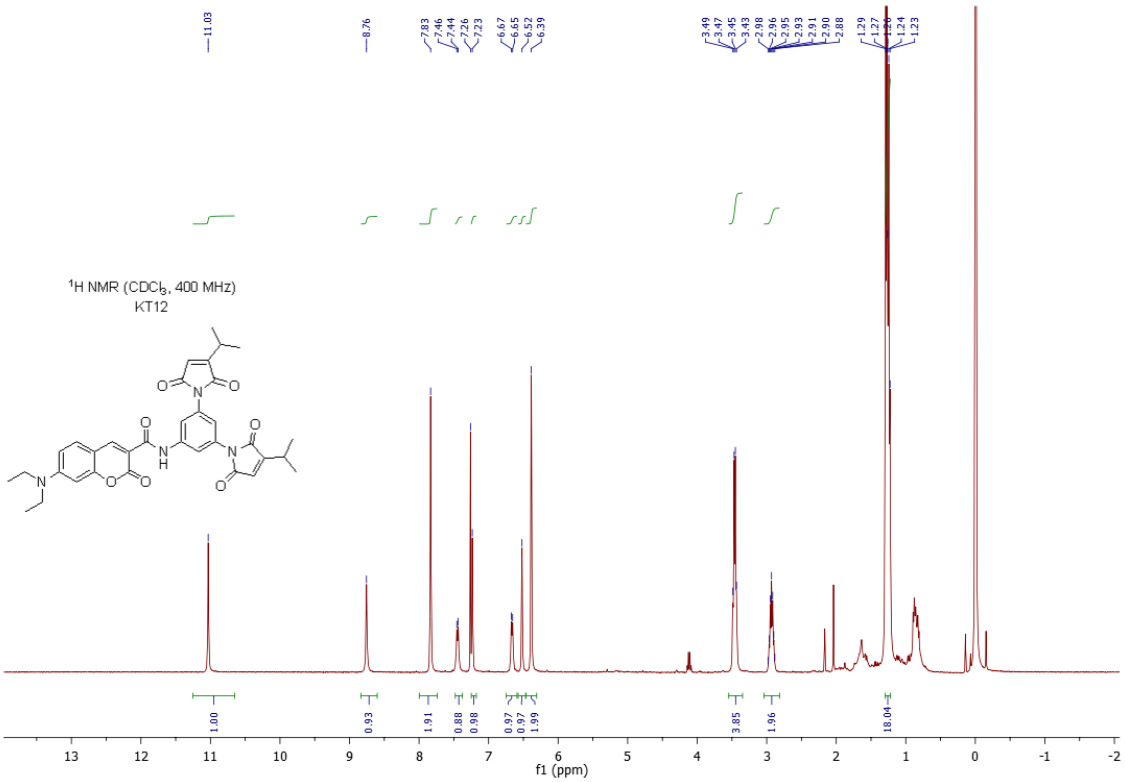

152.79
147.89
140.40
97.39
96.64
80.38
28.48
^{13}C NMR ($CDCl_3$, 100 MHz)
Compound 2.25
NHBoc
H_2N
NH_2
THF
THF
f1 (ppm)

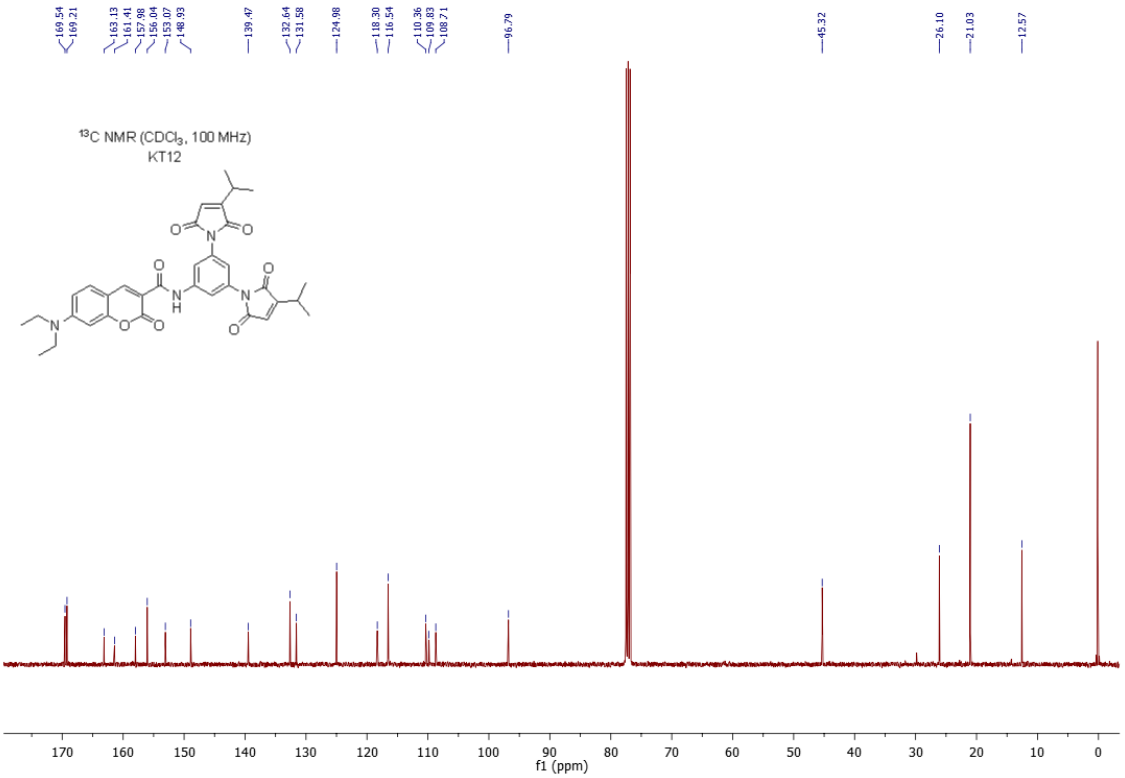

7.45
7.26
7.16
6.70
6.36
2.94
2.92
2.91
2.89
2.87
1.50
1.27
1.25
0.01
^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 2.26
NHBoc
1.92
0.92
1.00
1.98
1.98
9.02
12.38
f1 (ppm)

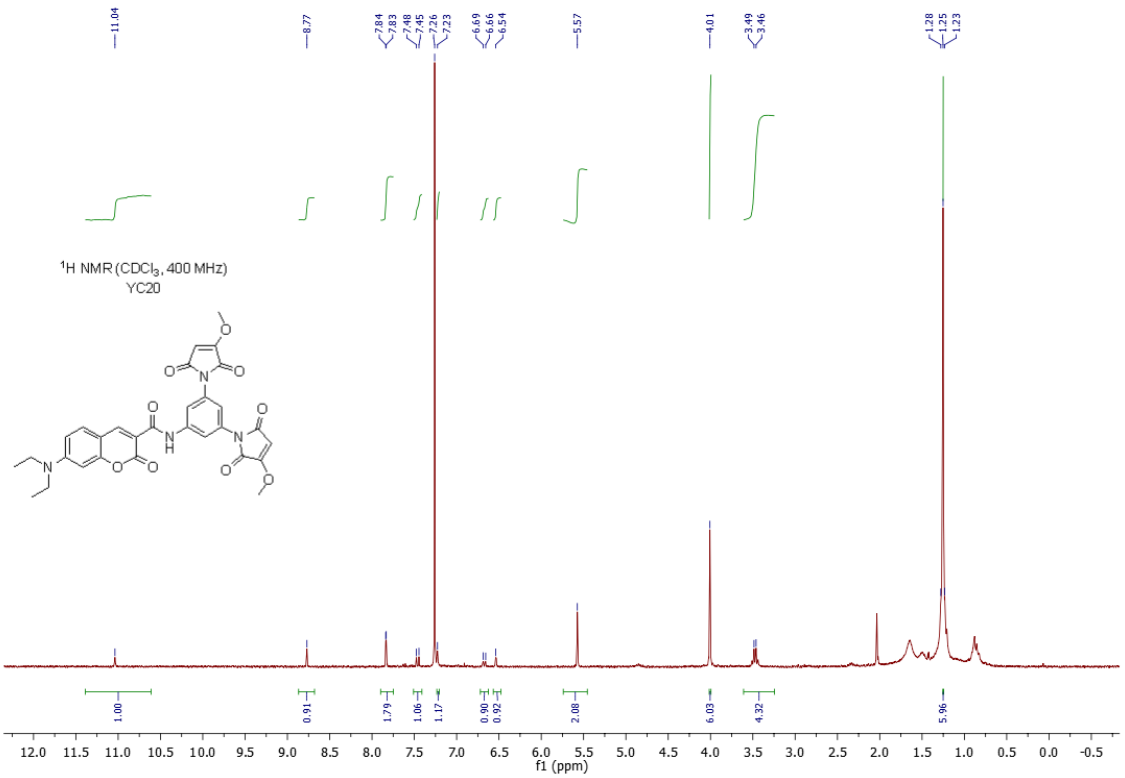

169.56
169.23
156.06
152.42
139.63
132.77
125.67
124.97
117.01
114.46
81.22
28.42
26.08
21.00
^{13}C NMR ($CDCl_3$, 100 MHz)
Compound 2.26
NHBoc
f1 (ppm)

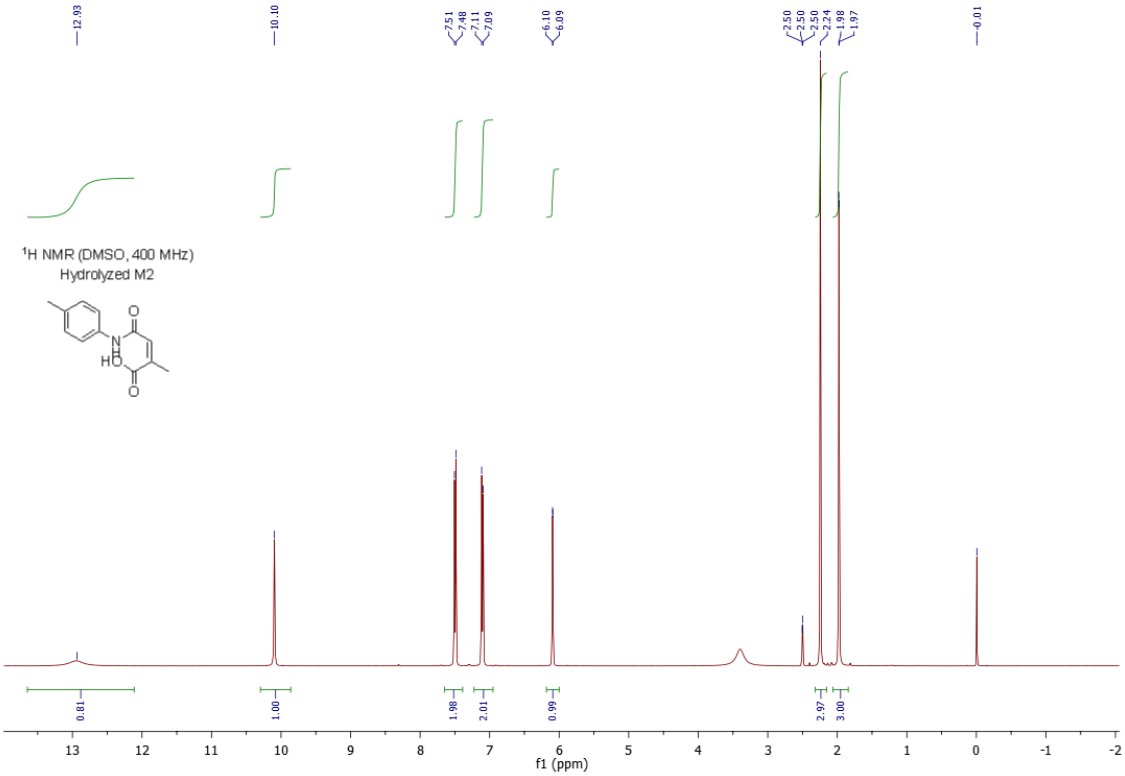

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 2.27
NH_2
7.26
6.87
6.70
6.35
4.02
2.92
2.90
2.89
1.27
1.25
0.00
1.00
2.00
2.03
2.09
2.03
12.35
f1 (ppm)

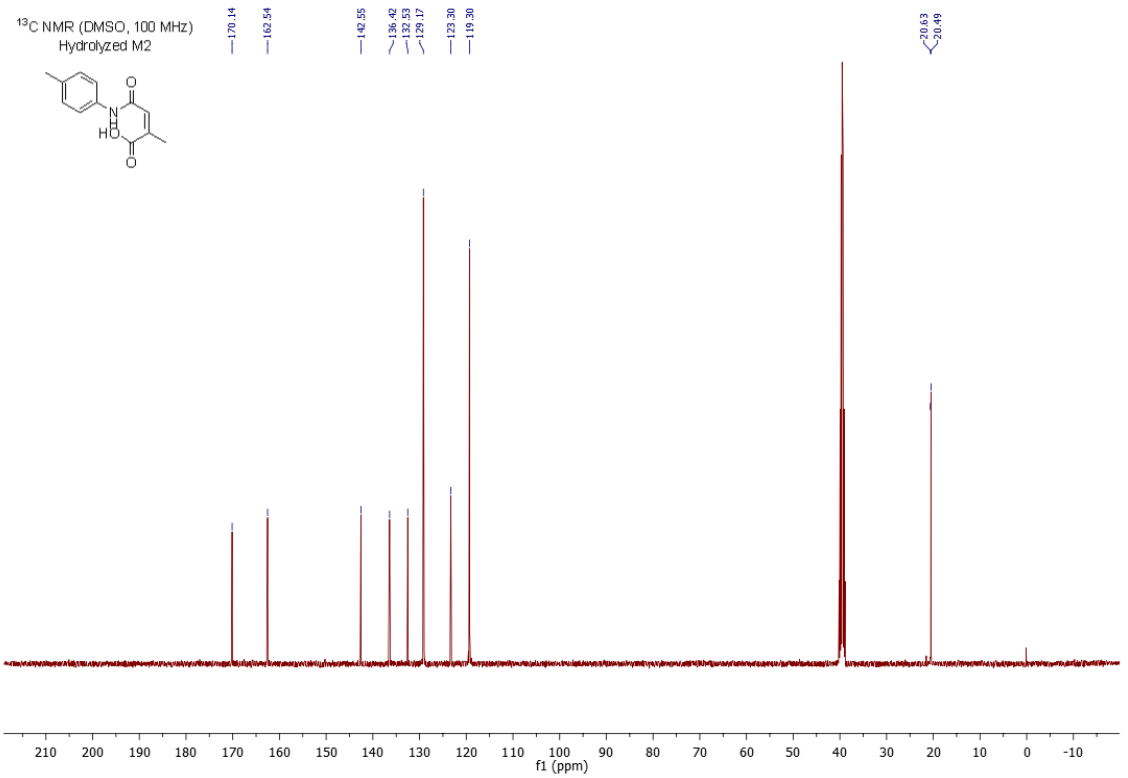

^{13}C NMR ($CDCl_3$, 100 MHz)
Compound 2.27
NH_2
169.76
169.45
155.98
147.32
133.04
124.95
112.76
111.37
26.07
21.01
f1 (ppm)

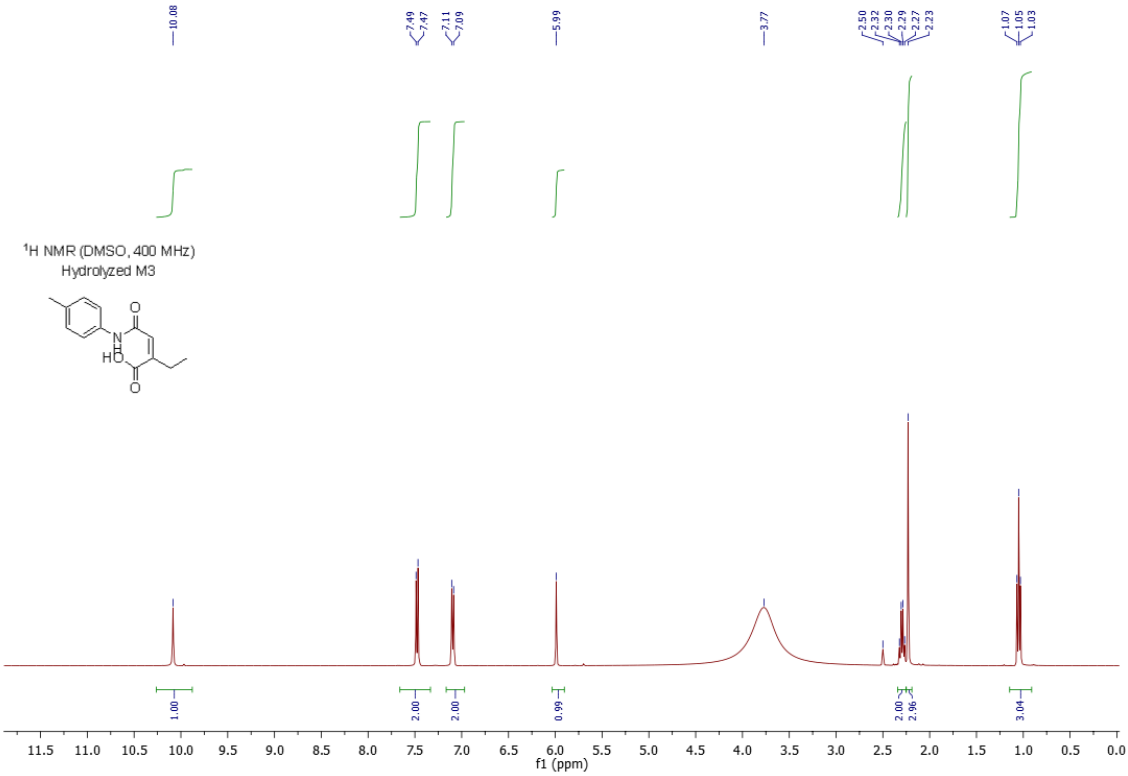

11.03
8.76
7.83
7.46
7.44
7.26
7.23
6.67
6.65
6.52
6.39
3.49
3.47
3.45
3.43
2.98
2.96
2.95
2.93
2.91
2.90
2.88
1.29
1.27
1.26
1.24
1.23
1H NMR (CDCl3, 400 MHz)
KT12
1.00
0.93
1.91
0.88
0.98
0.97
0.97
1.99
3.85
1.96
18.04
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
f1 (ppm)

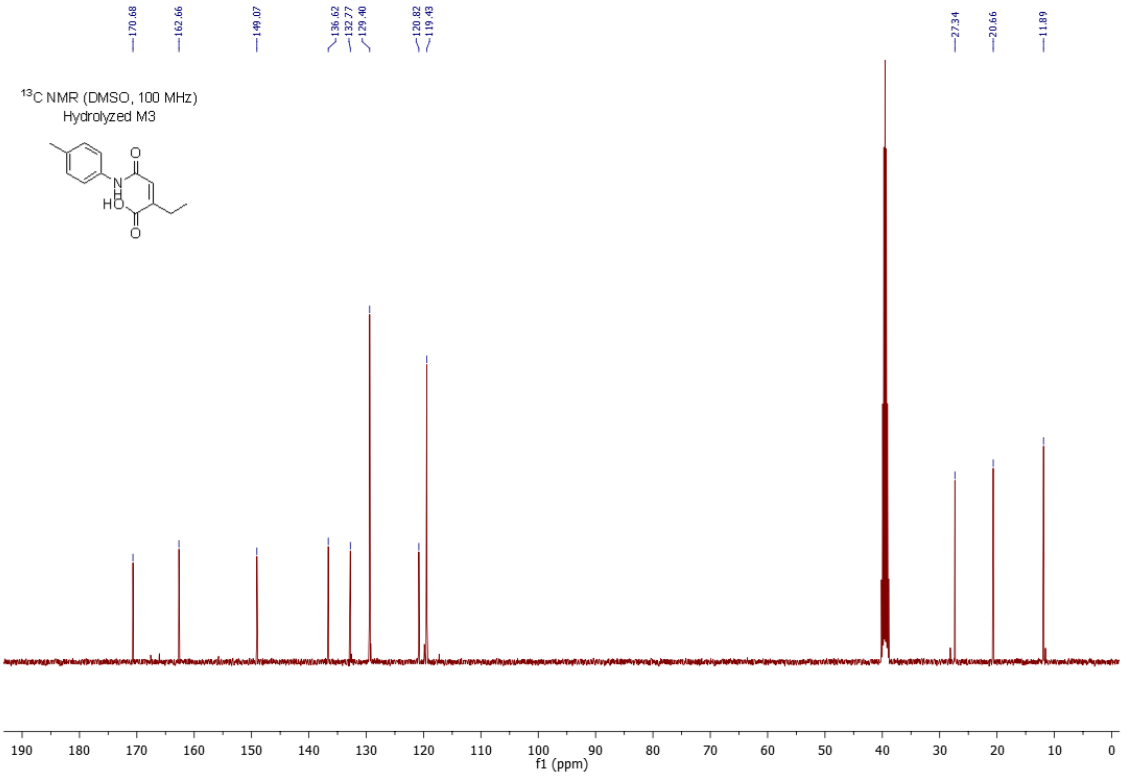

169.54
169.21
163.13
161.41
157.98
156.04
153.07
148.93
139.47
132.64
131.58
124.98
118.30
116.54
110.36
109.83
108.71
96.79
45.32
26.10
21.03
12.57
13C NMR (CDCl3, 100 MHz)
KT12
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

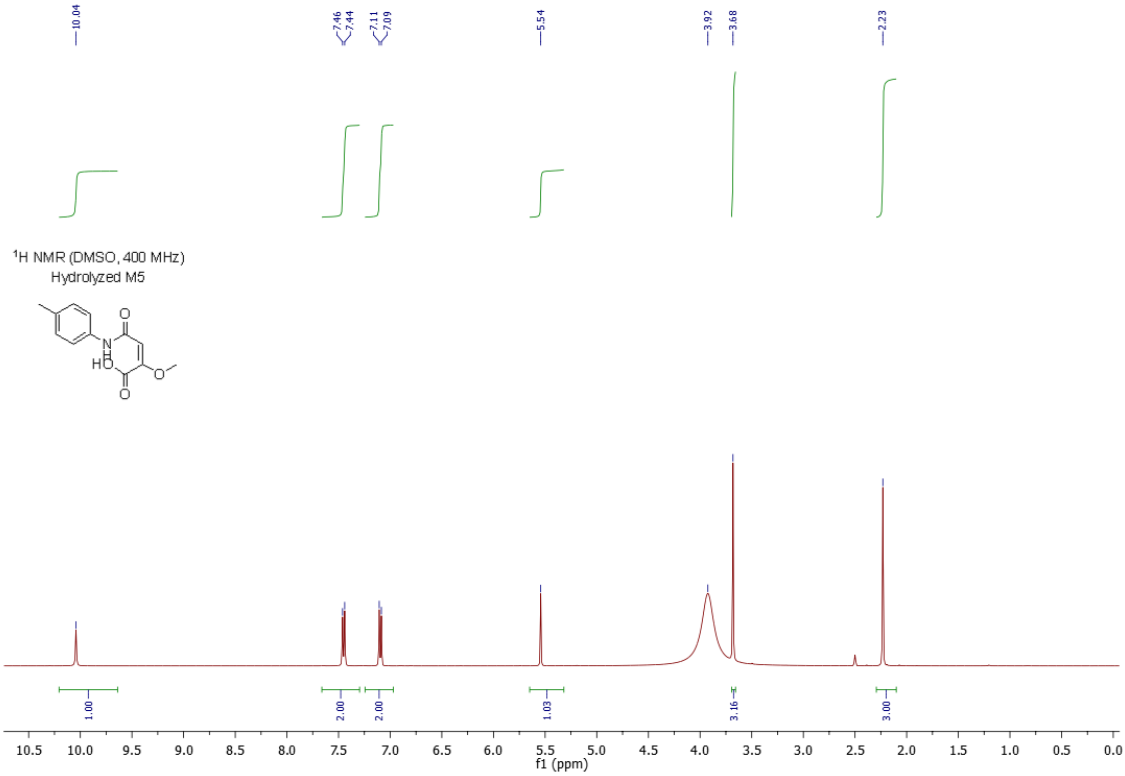

11.04
8.77
7.84
7.83
7.48
7.45
7.26
7.23
6.69
6.66
6.54
5.57
4.01
3.49
3.46
1.28
1.25
1.23
^{1}H NMR (CDCl$_3$, 400 MHz)
YC20
1.00
0.91
1.79
1.06
1.17
0.90
0.92
2.08
6.03
4.32
5.96
f1 (ppm)

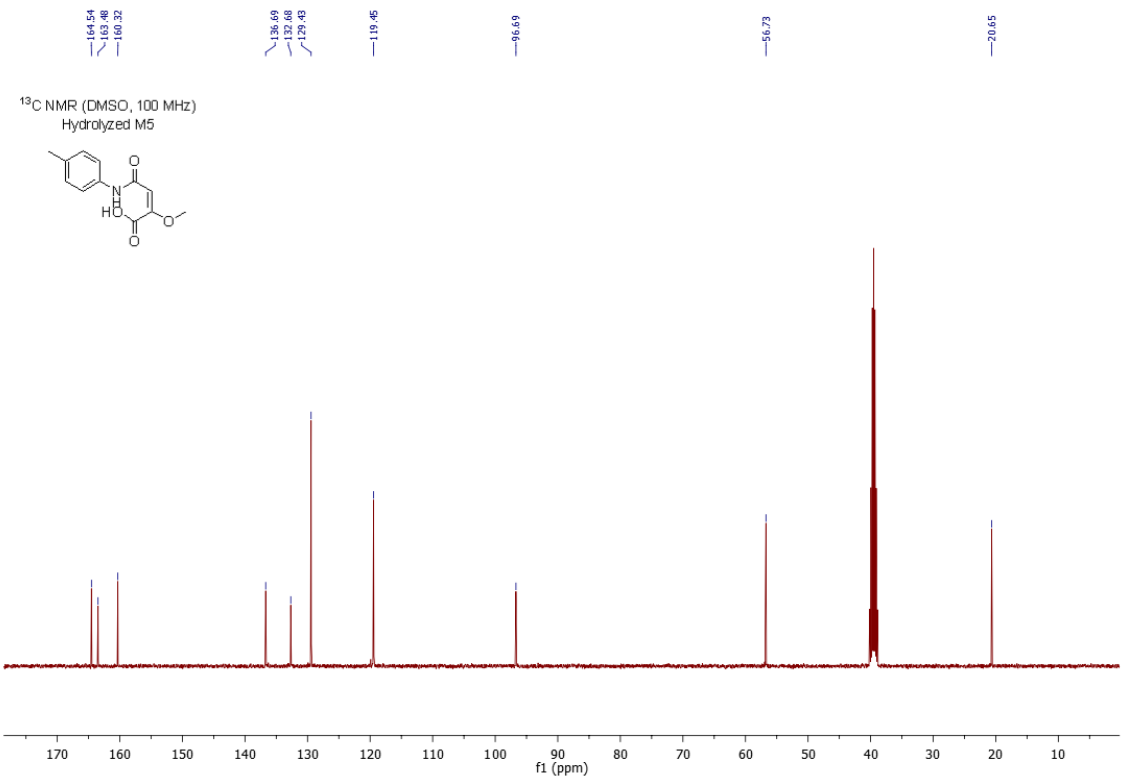

12.93
10.10
7.51
7.48
7.11
7.09
6.10
6.09
2.50
2.50
2.50
2.24
1.98
1.97
0.01
^{1}H NMR (DMSO, 400 MHz)
Hydrolyzed M2
0.81
1.00
1.98
2.01
0.99
2.97
3.00
f1 (ppm)


13C NMR (DMSO, 100 MHz)
Hydrolyzed M2
170.14
162.54
142.55
136.42
132.53
129.17
123.30
119.30
20.63
20.49
f1 (ppm)

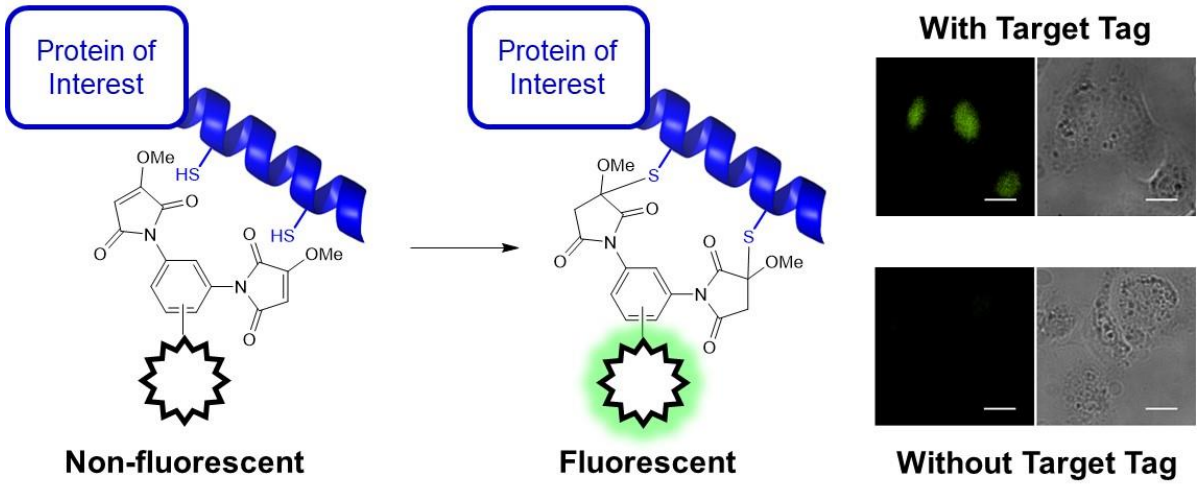

10.08
7.49
7.47
7.11
7.09
5.99
3.77
2.50
2.32
2.30
2.29
2.27
2.23
1.07
1.05
1.03
1H NMR (DMSO, 400 MHz)
Hydrolyzed M3
1.00
2.00
2.00
0.99
2.00
2.96
3.04
f1 (ppm)


170.68
162.66
149.07
136.62
132.77
129.40
120.82
119.43
27.34
20.66
11.89
13C NMR (DMSO, 100 MHz)
Hydrolyzed M3
f1 (ppm)


10.04
7.46
7.44
7.11
7.09
5.54
3.92
3.68
2.23
1H NMR (DMSO, 400 MHz)
Hydrolyzed M5
1.00
2.00
2.00
1.03
3.16
3.00
f1 (ppm)


164.54
163.48
160.32
136.69
132.68
129.43
119.45
96.69
56.73
20.65
13C NMR (DMSO, 100 MHz)
Hydrolyzed M5
O
N
H
HO
O
O
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)

Chapter 3: Imaging New Elements of the FlARe Labelling Method in Cells



3.1. Statement of Contributions

Figure 3.5 and the conclusions drawn from this figure were adapted by permission of The Royal Society of Chemistry (Strmiskova, M.; Tsao, K.; Keillor, J. W. (2018) Rational Design of a Highly Reactive Dicysteine Peptide Tag for Fluorogenic Protein Labelling. *Org. Biomol. Chem.* 16, 6332–6340).[1] Figure 3.9 and the conclusions drawn from this figure were adapted from Chen, Y.; Tsao, K.; Acton, S. L.; Keillor, J. W. (2018) A Green BODIPY-Based, Super-Fluorogenic, Protein-Specific Labelling Agent. *Angew. Chem. Int. Ed. Engl.* 57, 12390–12394 (Copyright (2018) John Wiley & Sons).[2] See Appendix for license.

YC23 was first synthesized by Dr. Yingche Chen and then re-synthesized and optimized by Sydney Acton. **YC20** was first synthesized by Dr. Chen and Kelvin Tsao. The dC10* tag was designed and developed by Miroslava Strmiskova.

3.2. Background

After establishing that the methoxymaleimide labelling agents offered the highest selectivity for the dC10 target peptide, we focused on implementing two new developments of FlARe in cells. This constitutes the final step towards establishing a protein labelling method, or aspects of that method, towards the goal of studying cellular dynamics. Intracellular labelling with the FlARe probes is difficult due to the numerous uncontrollable factors that regulate many cellular processes, which may interfere with the probe reacting with the target peptide. For example, the protein of interest will likely only be present at low concentrations while many of the cell's other components such as glutathione or other accessible thiols will be present in considerably higher concentrations. Furthermore, the probe concentrations within the cell may be limited by its ability to cross the cell membrane and stability within the cell. On top of this, the tag's accessibility when appended to the protein of interest (POI) will also influence labelling efficiency.

3.2.1. Fixed Cell Imaging versus Live Cell Imaging

Before fluorescence microscopy, samples can be fixed so that decay caused by autolysis or putrefaction can be prevented.[3] Fixation terminates ongoing biochemical reactions and takes a "snap-shot" of the current cellular state, allowing for multiple treatments to be performed on fixed samples and for the sample to be observed for longer periods of time. This preservation can be achieved through physical methods such as desiccation, heat-treatment or freezing. Alternatively, chemical methods can be performed using crosslinking or denaturing fixatives like formaldehyde or methanol, respectively. Although these kinds of sample preparation make the samples easy to work with, they alter the state of the cell and

render its processes static. As such, the observations and data collected from imaging may not be as reliable or relevant.

Live-cell imaging offers an alternative to fixation, which allows cells to be observed in real time and for dynamic changes to be monitored. In doing so, cellular processes can be studied in the cells' native environment, which offers more insight into cell biology. Additionally, samples do not need to be prepared through physical or chemical processes before imaging. However, for live-cell imaging to work, cells must be kept alive, which entails tightly controlling factors such as temperature, pH, and growth environment while also compromising on factors that cannot be easily maintained such as atmospheric CO_2 levels. Therefore, live-cell imaging is highly time sensitive as unhealthy cells can produce artifactual results due to changes in biological processes that make up a cell's survival response.[4] Given the relevance of the information offered by live-cell imaging, we opted to demonstrate the intracellular labelling of proteins carrying the dC10 tag with the FlARe probes in live HeLa cells.

3.2.2. Fluorescence Microscopy

The simplest form of fluorescence microscopy is **w**idefield **f**luorescence microscopy (WF), in which the entire sample is illuminated by the excitation light. To do this, excitation light from a source is passed through an excitation filter that selects for the correct wavelength in order to excite the fluorophores within the sample. The light is directed onto the sample upon reaching a dichroic mirror, which reflects excitation wavelength light and allows emission wavelength light to pass through. Once the sample is excited, the longer wavelength fluorescence passes through the dichroic mirror and is filtered through an emission filter before reaching the detector. In doing so, none of the light coming from the excitation source

reaches the detector (Figure 3.1).[5] In more modern microscopes, excitation lasers of specific wavelengths are used, negating the need for an excitation filter. Careful selection of both the excitation lasers and the filter sets to control observed emitted light is important for both obtaining maximal fluorescence and avoiding spectral bleeding or "cross-talk". Spectral bleeding occurs when the excitation and emission spectra of multiple fluorophores overlap. This can complicate cellular imaging when more than one fluorophore is used to label different features in the cell. In WF microscopy, the entire sample is excited and so signals from every point in the sample are detected, leading to lower contrast in thicker samples. Thus, widefield fluorescence works best for thin samples, such as a single layer of adherent cells. Even then, cells are not perfectly flat and so differences in fluorescence signal can be observed throughout the cell due to the vertical stacking of fluorescent molecules that may differ throughout the cell. Despite this, WF microscopy is commonly used as it is simple, fast, sensitive and provides images in high resolution.

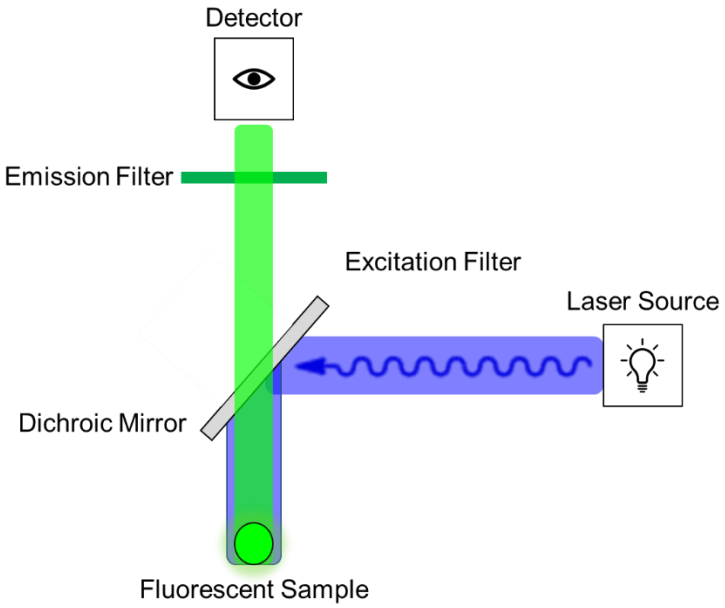


Figure 3.1. Skeleton components of a microscope used to excite and image a fluorescent sample using widefield fluorescence microscopy. Here, the light source is not a laser and so a filter is required to select the correct excitation wavelength. The sample consists of a green emitting fluorophore, which is excited using shorter wavelength blue light.

Confocal laser scanning microscopy (CLSM) follows the same principles as WF microscopy, but instead uses a pinhole to focus the initial excitation light onto only a small targeted part of the sample.[6] The emitted light from this target point is also filtered by the dichroic mirror and a second filter and then passed through a pinhole. This pinhole rejects non-focussed and scattered light and as a result the light is both focussed onto a diffraction limited spot illumination and spatially filtered through a pinhole emission, such that only light from a target point in the sample is observed (Figure 3.2). The microscope then scans the sample at a specific depth to construct a 2D image at that depth. Repeating this process at different depths to obtain multiple 2D images allows for the 3-dimensional reconstruction of the whole sample. The optical sectioning achieved by CLSM allows for an increased signal

to noise ratio, especially in thicker samples. However, the process of scanning the entire sample is time consuming, which can considerably lengthen imaging.

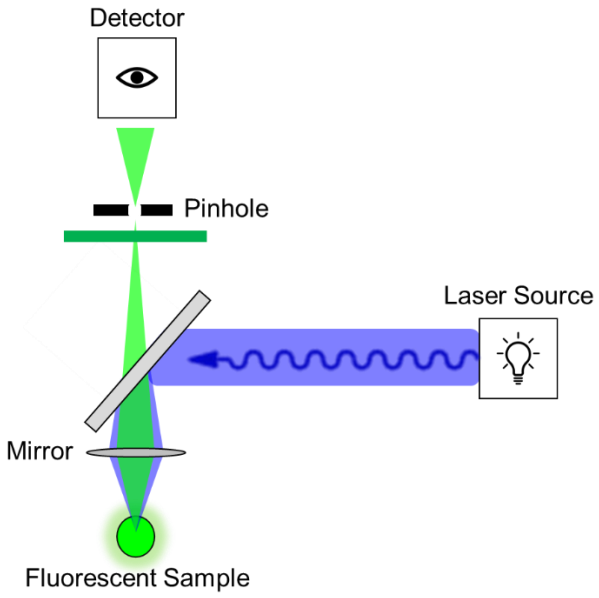


Figure 3.2. Skeletal components of a confocal laser scanning microscope. Here, point illumination is achieved by reflecting the excitation light to a specific point using a mirror. Afterwards, the light from the targeted point is filtered through a pinhole to reject unfocussed and scattered light before reaching the detector.

As previously mentioned, we opted to demonstrate the capabilities of the FlARe labelling method through live-cell imaging. Since cell survival is of utmost importance during live-cell imaging, we chose to perform WF microscopy using the Nikon Ni-U Ratiometric Fluorescence Microscope at the University of Ottawa. This is because the simpler, more sensitive and faster imaging capability offered by WF microscopy limits the time the cells spend outside optimal growth conditions. Although contrast is compromised in thick samples with WF microscopy, this is avoided by using samples consisting only of a monolayer of cells.

3.2.3. A Novel Target Peptide and FlARe Probe

Before this thesis chapter, Miroslava Strmiskova developed a new dC10* target peptide that reacted faster with FlARe probes[1,7] and Dr. Yingche Chen and Sydney Acton

developed a new green fluorogenic FlARe probe, **YC23**.[2,8] These developments showed the potential to both improve the FlARe labelling reaction and expand the method's colour palette. However, both projects had only been validated up to the labelling of a purified test protein in solution, so we aimed to implement these methods in a cellular context as the final step towards their validation. Prior work in the Keillor group has demonstrated the cellular labelling of different dC10 tagged proteins using numerous different FlARe probes in HEK293 cells.[7,9,10] Herein, we apply the FlARe method to label dC10 tagged histone H2B protein in HeLa cells given the ease at which this cell line is grown[11] and HeLa's morphologically distinct nucleus. This makes it easier to assess successful labelling of the nuclear-localized H2B protein while evaluating the novel dC10* target peptide and green fluorogenic **YC23** FlARe probe.

The new dC10* tag was rationally designed by Miroslava Strmiskova to improve its reactivity towards the FlARe probes.[7] This was accomplished through the introduction of positively charged residues around the two cysteine amino acids. In doing so, the pK_a values of the cysteine thiols were lowered, which greatly increased the amount of thiolate present, at the expense of slightly diminishing the nucleophilicity of the thiolate.[12] These additions constructed a dicysteine peptide tag that reacted with the dansyl-based FlARe probe, **dM10-dansyl**, and the coumarin-based FlARe probe, **YC9**, nine times faster than the original dC10 tag (Figure 3.3). The kinetic advantage offered by augmenting the tag's reactivity would ideally improve the probe's selectivity over other free but ideally less reactive thiols in the cellular environment. Additionally, this may also shorten labelling times; however, this can also be influenced by other factors such as probe permeability or the rate at which the probe is taken up by the cell. In this work, we evaluate the dC10* tag's reactivity with the coumarin-

based FlARe probe, **YC20** (Figure 3.3), by labelling histone H2B in order to validate the novel tag in a cellular context.

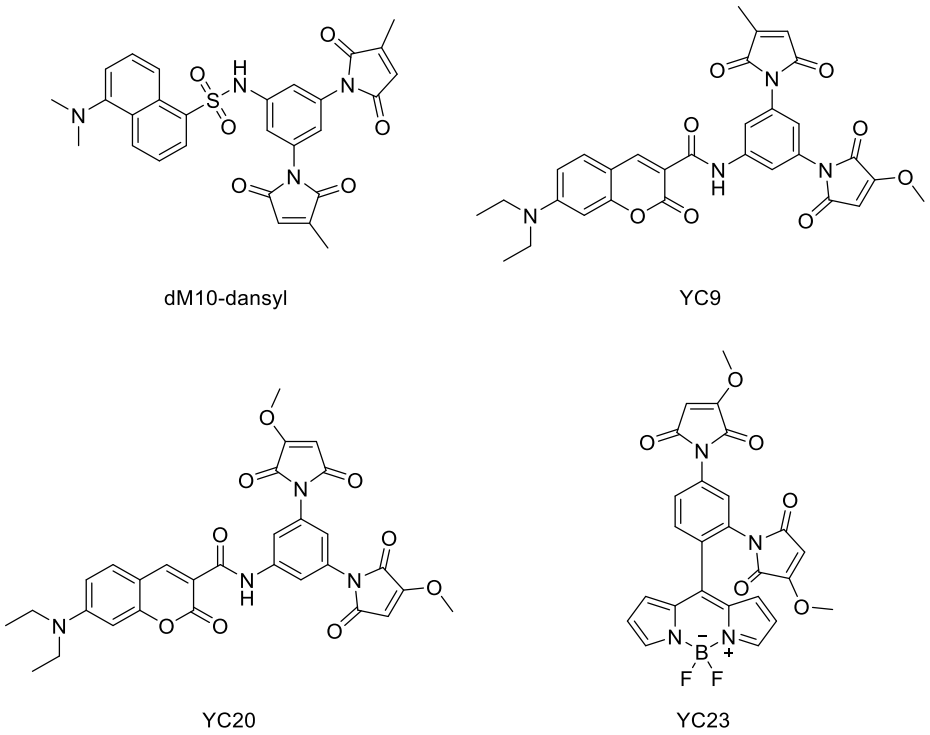


Figure 3.3. Structures of fluorogens **dM10-dansyl**, **YC9**, **YC20** and **YC23**.

In addition to dC10*, a novel green fluorogenic FlARe probe called **YC23** was designed, synthesized and evaluated by Dr. Yingche Chen.[2] Later, Sydney Acton (M.Sc.) optimized the synthesis and re-evaluated **YC23**.[8] **YC23** has an excitation maximum at 480 nm and an emission maximum at 535 nm, which aligns it optimally with the 485 nm excitation laser and 535/45 nm emission filter on the Nikon Ni-U Ratiometric Fluorescence Microscope. Not only does this new FlARe probe expand the colour range of the FlARe labelling method, the longer excitation wavelength of **YC23** allows phototoxic, short-wavelength excitation light to be avoided. The longer emission wavelength of **YC23** also allows higher penetrative ability through thicker tissue. Additionally, the probe's in vitro characteristics such as its selectivity for the dC10 tag, photostability, and quantum yield ($\Phi = 0.27$) make it a promising

protein imaging agent.[8] Thus, we labelled H2B-dC10 in HeLa cells using **YC23** in order to evaluate it in cells.

3.3. Results and Discussion

3.3.1. Establishing a method for live-cell labelling with FlARe probes

We commenced with the live-cell no-wash imaging of histone-H2B-dC10*, expressed in HeLa cells and labelled with the coumarin-based **YC20**. This particular FlARe probe had been previously established as a successful fluorogenic protein labelling agent in HEK293 cells expressing histone H2B-dC10.[10] Initial attempts at live-cell imaging showed a disruptive number of free-floating dead cells, which fluoresce much brighter than live and adhering cells. This cell death was also observed in the control cells transfected with an empty vector (pcDNA 3.1 (+)), suggesting that H2B-dC10* was not the cause. The elevated brightness found in dead cells is likely due to autofluorescence from cell debris and increased uptake of the probe through the cell's compromised membrane integrity.[4] The axial stacking of multiple fluorescent molecules in a spherical suspended cell would also contribute to increased fluorescence intensity. Therefore, it is integral that cell death is limited during live-cell imaging in order to obtain usable protein labelling information. As such, three factors in our labelling method were identified that may potentially contribute to cell death.

The first factor that could lead to cell death is the use of Lipofectamine®2000 in high concentrations and over prolonged periods during the transient transfection of vectors into HeLa cells.[13] Lipofectamine®2000, from ThermoFisher Scientific, is a cationic liposome formulation that complexes with the negatively charged nucleic acid biomolecules, which are naturally repelled by the negatively charged cell membrane.[14] Along with a helper neutral co-lipid, Lipofectamine®2000 helps nucleic acid molecules penetrate the cell and reach the

nucleus, where the cell's transcriptional machinery is located, so that the coded protein can be expressed. The initial conditions used to transfect HeLa cells with a plasmid coding for H2B-dC10* involved the use of 24 µg of lipofectamine and 7 µg of plasmid for 8.5×10^5 cells seeded (see Experimental). However, these conditions were found to result in a large amount of cell death, which impeded the imaging of labelled proteins in HeLa cells. Thus, in light of an optimization study,[14] we decreased the amount of transfection agent used to 10 µg to limit cell death. Dilution and mixing times of 30 min were also included for sufficient liposome formation and DNA-liposome complexation, respectively (see Experimental). The number of dead cells was observed to decrease, although transfection efficiency could not be evaluated without labelling transfected cells with **YC20**. However, since numerous variables could influence labelling, including cell viability during labelling, labelling efficiency, transfection efficiency, using **YC20** labelling to compare transfection efficiency is unreliable. Alternatively, a vector expressing an independently observable model protein such histone H2B-mNeptune could be used for a more rigorous comparison of transfection efficiency. Despite this, vector size also influences transfection efficiency, and so the conditions used to transfect one vector may not be optimal for another.[15]

The second factor that contributes to cell death is insufficient nutrients in the medium for cell growth and studies have shown that limiting cell exposure to serum-free medium can significantly aid cell survivability.[13] The initial labelling procedure involved incubating cells for 1 h in 10 µM YC20 in *serum-free* DPBS ((+) Ca^{2+}, (+) Mg^{2+}). As such, the PBS buffer was substituted with the Live Cell Imaging Solution from Molecule Probes® (20 mM HEPES, pH 7.4, 140 mM NaCl, 2.5 mM KCl, 1.8 mM $CaCl_2$, 1.0 mM $MgCl_2$), which was formulated

to provide improved signal-to-noise ratio and cell viability during imaging. This led to a further decrease in cell death, allowing for more reliable cell imaging to be performed.

The third factor potentially contributing to cell death is the use of the pluronic surfactant F127, which can disrupt the lipid bilayer of cell membranes when used in high concentrations.[16] This non-ionic surfactant is used during labelling to improve the solubility and cell permeability of hydrophobic small molecule probes through micelle formation. However, given its capability to compromise cellular membranes, we evaluated the necessity and cytotoxicity of the F127 surfactant during labelling by incubating H2B-dC10 expressing HeLa cells with 10 μM of the BODIPY-based **YC23** with and without the F127 surfactant (Figure 3.4).

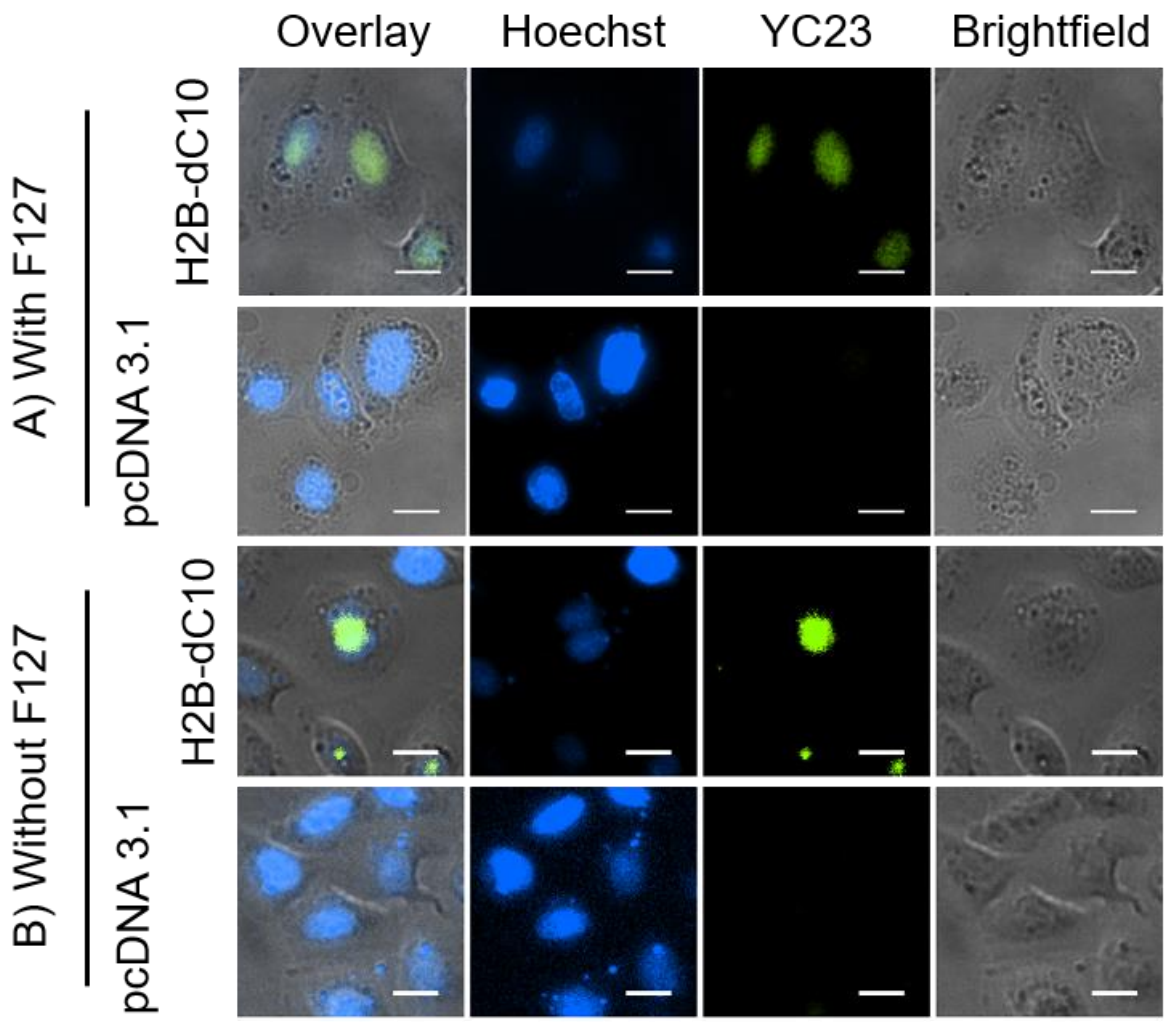


Figure 3.4. Protein labeling images of HeLa cells transfected with histone H2B-dC10 or pcDNA 3.1(+) and then labeled with a 10 μM **YC23** with or without F127 surfactant, followed by 8 μM Hoechst 33342. Cells were observed using a 20X/0.50 NA objective lens. Hoechst was excited at 390 nm and observed in the 485/50 nm channel. **YC23** was excited at 485 nm and observed in the 535/45 nm channel. Scale bars: 20 μm.

After incubation with the labelling solution, only HeLa cells expressing histone H2B-dC10 and labelled with the solution containing the F127 surfactant showed green fluorescence from **YC23** (Figure 3.4A) and limited cell death. This green fluorescence colocalized well with the blue fluorescence from the fluorogenic DNA binding Hoechst stain, which highlights the cell's nucleus. Cells expressing H2B-dC10 and labelled without using the surfactant demonstrated green fluorescence that did not fully colocalize with the Hoechst stain's blue

fluorescence, indicating incomplete labelling that did not permeate through the entire nucleus. Therefore, F127 does not significantly contribute to cell death and may help ensure that labelling nuclear proteins, such as histone H2B, occur throughout the entire nucleus and not in punctuated subsections of the organelle.

3.3.2. Demonstrating dC10 as a target peptide*

Once reliable transfection, labelling, and imaging conditions were determined, the evaluation and implementation of novel FlARe probes and tags could commence. As mentioned, Miroslava Strmiskova improved the reactivity of the previously established dC10 target peptide by introducing positively charged residues proximal to both cysteines on the tag, resulting in the novel dC10* tag. Ideally, this faster reactivity between the FlARe probe and the tag would translate to an improved selectivity and background signal as well as potentially shortening labelling times. Therefore, to evaluate dC10* as a viable peptide tag, HeLa cells expressing the histone H2B-dC10* were incubated with **YC20** and evaluated for successful protein labelling.

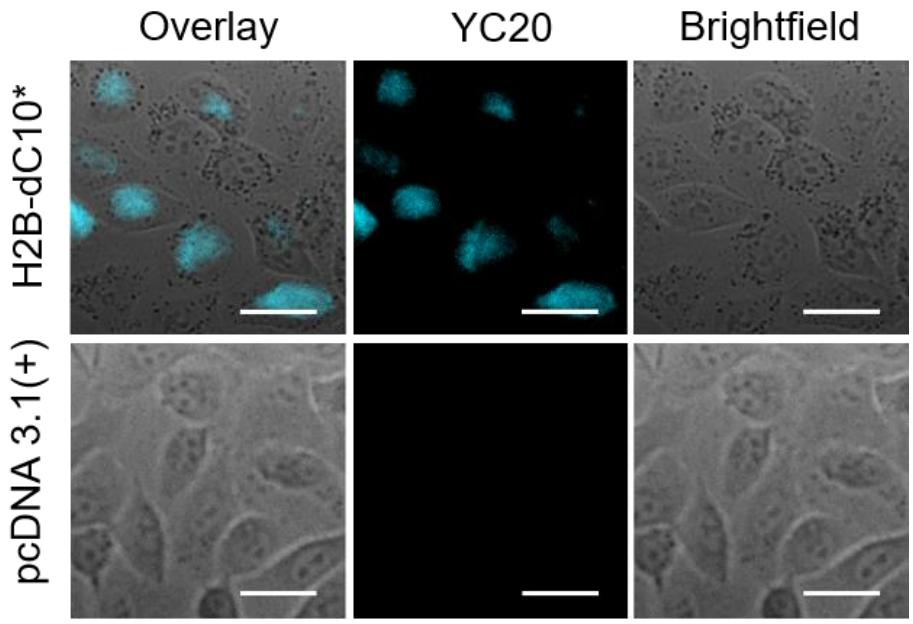


Figure 3.5. Protein labeling images of HeLa cells transfected with histone H2B-dC10* (top) or pcDNA 3.1(+) (bottom) and then labeled with YC20 (10 μM) for 45 min. Cells were observed using a 20X/0.50 NA objective lens. YC20 was excited at 438 nm and observed in the 485/25 nm channel. Scale bars: 20 μm.

HeLa cells transfected with an empty expression vector, and not expressing the dC10* tagged protein did not exhibit fluorescence, confirming that without the di-cysteine tag, the FlARe probe remains quenched (Figure 3.5). Cells expressing H2B-dC10* showed visible cyan fluorescence in only the nuclei where the histone protein is localized. These results are consistent with previous studies where histone H2B-dC10 was labelled using the same **YC20** probe in HEK293 cells.[10] As such, neither the cell line nor the modifications done to the target peptide had a significant effect on the FlARe labelling method. Although the nuclear-localized cyan fluorescence in cells expressing the dC10* tagged protein clearly shows the FlARe probe's selectivity for the tag, it is difficult to ascertain if there is an increase in selectivity due to the tag's faster reactivity. Nonetheless, the successful live-cell and no-wash labelling of a dC10* tagged protein with the **YC20** FlARe probe in HeLa cells establishes the tag's capability as a target peptide for the FlARe labelling method.[1]

After confirming the dC10* tag as a target peptide for FlARe, the rate of cellular labelling in HeLa cells was studied. HeLa cells expressing H2B-dC10* were incubated with **YC20** and imaged at discrete time intervals. Nuclear-localized, cyan fluorescence could be observed in healthy cells as early as 30 minutes but increased significantly after 45 minutes (Figure 3.6). This rapid increase in fluorescence intensity between 30 and 45 minutes suggests a 30-minute lag time before the FlARe reaction begins. The observed lag in labelling may be due to the movement of **YC20** across not only the cell membrane but the nuclear membrane. Given this, labelling proteins localized to other compartments of the cell will likely influence labelling times. Once the FlARe reaction commences, the fluorescence of **YC20** can be observed for up to 60 min. Notably, cells had to be removed from the microscope and returned to the incubator between time points to prevent excessive cell death. As such, the same area of the plate could not be monitored, which introduced variability into the images between each time point. Regardless, the above observations support previous FlARe labelling studies, which used labelling times of 1 h before imaging.[10]

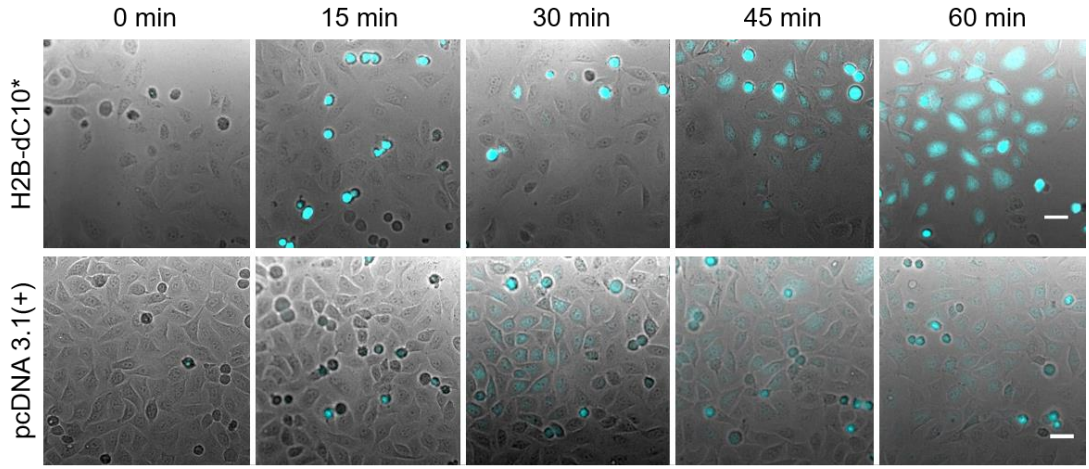


Figure 3.6. Protein labeling images of HeLa cells transfected with histone H2B-dC10* (top) or pcDNA 3.1(+) (bottom) and then labeled with YC20 (10 μM) for 1 h. Cells were observed using a 20X/0.50 NA objective lens. YC20 was excited at 438 nm and observed in the 485/25 nm channel at 0, 15, 30, 45, and 60 min intervals. Scale bars: 20 μm.

It is also instructive to compare the FlARe labelling method to other cysteine-based protein labelling methods. For example, the FlAsH method involves a thiol-reactive FlAsH probe that reacts with a tetracysteine CCPGCC tag (Figure 3.7A, see Section 1.2.3).[17] With this method, labelled protein can be observed as early as 15 min after incubating cells in the labelling solution, with maximal fluorescence being reached after 90 min.[18] Despite this, many FlAsH protocols suggest labelling times of around 1 h.[19,20] Another thiol-dependant method involves BODIPY-diacrylate probes and a dicysteine RC[2] tag (Figure 3.7B). In this method, labelled proteins were observed after 30 min.[21] Although FlARe's fluorescence may take longer to observe, the methods mentioned above both require washing before imaging, and FlARe's labelling procedure is much easier than the protocol for FlAsH.

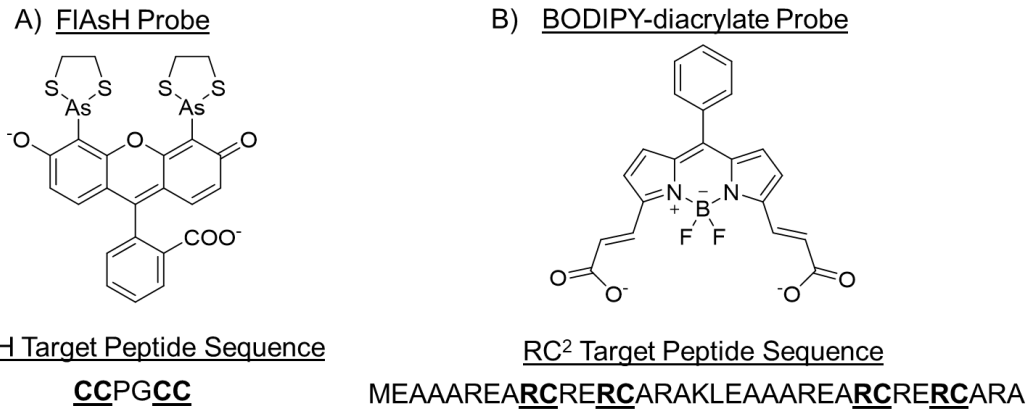


Figure 3.7. Structure and target peptide sequence for A) the FlAsH labelling method and B) the BODIPY-diacrylate probes. Reactive cysteines and important residues are highlighted with **bold underline**.

HeLa cells transfected with the empty pcDNA 3.1(+) vector were also incubated with **YC20** and monitored over time. Cyan fluorescence throughout the entire cell was observed in healthy cells starting at 30 minutes (Figure 3.6), but no significant increase in fluorescence was observed between 30 and 45 minutes. Instead, this basal background fluorescence slowly increased until the end of the experiment. Previous reports showed that **YC20** is remarkably

stable in HEPES buffer at pH 7.4 and resistant to 1-10 mM of GSH.[10] However, the background fluorescence observed here is both contained within the cell but also not co-localized to a specific compartment. Therefore, the FlARe probe may be reacting with a thiol within the cell under conditions not replicated by the prior *in vitro* experiments, leading to a basal level of background fluorescence that increases over time. Regardless, when cells without dC10* tagged proteins are incubated with **YC20** and compared to cells that express a dC10* tagged protein, this background fluorescence is negligible.

As a final experiment, labelling the dC10* and dC10 tag in HeLa cells were compared side-by-side to ascertain if the increased dC10* reactivity improved labelling times. As such, HeLa cells expressing H2B-dC10 and HeLa cells expressing H2B-dC10* were both incubated with **YC20** and imaged at discrete time points. *In vitro* studies from previous works demonstrated that the dC10* tag reacted two-times faster with **YC20** than the dC10 tag when appended to Maltose Binding Protein (MBP).[7] Interestingly, the FlARe reaction could be observed as early as 15 minutes in the cells expressing H2B-dC10, while this reaction was not observed until later in cells expressing H2B-dC10* (Figure 3.8). The earlier start may be due to variability in membrane permeabilization between samples. Alternatively, the dC10 tag may vary in accessibility compared to the dC10* tag, which would influence reactivity. Previous work performed a similar experiment where H2B-dC10 and H2B-dC10* was expressed in HEK293 cells and labelled with **YC20**.[7] Here, protein labelling was observed as early as 10 minutes, further demonstrating that factors other than the probe's reaction rate influences labelling times. In conclusion, although the dC10* tag reacted with the **YC20** FlARe probe two times faster than the dC10 tag *in vitro*, due to uncontrollable factors in a cellular environment, this did not translate to shorter labelling times in observed cells.

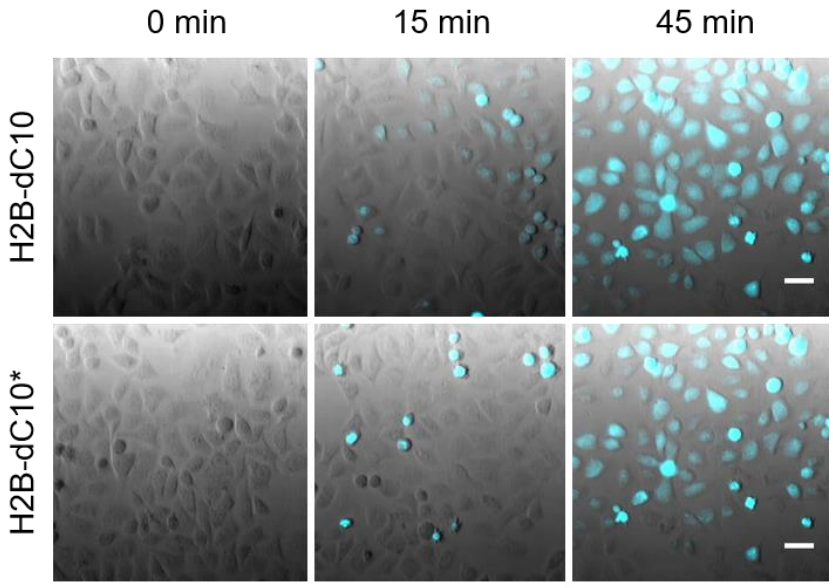


Figure 3.8. Protein labeling images of HeLa cells transfected with histone H2B-dC10 (top) or histone H2B-dC10* (bottom) and then labeled with YC20 (10 μM) for 45 min. Cells were observed using a 20X/0.50 NA objective lens. YC20 was excited at 438 nm and observed in the 485/25 nm channel at 0, 15, and 45 min intervals. Scale bars: 20 μm.

3.3.3. Implementation of YC23 as a live-cell no-wash protein imaging agent

In addition to evaluating the new dC10* tag inside HeLa cells, we also assessed the utility of a new green 3,5-dimethyl-BODIPY based FlARe probe, **YC23**. Here we aimed to label a histone H2B-dC10 construct expressed in HeLa cells. Following cell imaging, green fluorescence was observed in the nuclei of multiple cells (Figure 3.9) expressing H2B-dC10. Furthermore, even though **YC23's** maximal fluorescence could not be achieved in the span of 1 h,[2] the fluorescence achieved in this shorted time is still clearly visible with fluorescence microscopy. Counterstain with the DNA binding blue fluorogenic Hoechst stain further confirmed this nuclear colocalization due to the overlap between the blue and green fluorescent areas. Given the nuclear localization of histone H2B, the location of this fluorescence supports that it is a nuclear protein that is being labelled. The lack of fluorescence

observed from cells transfected with an empty vector demonstrates **YC23's** selectivity for the dC10 tag. These results show that **YC23** is a successful green fluorogenic FlARe probe.

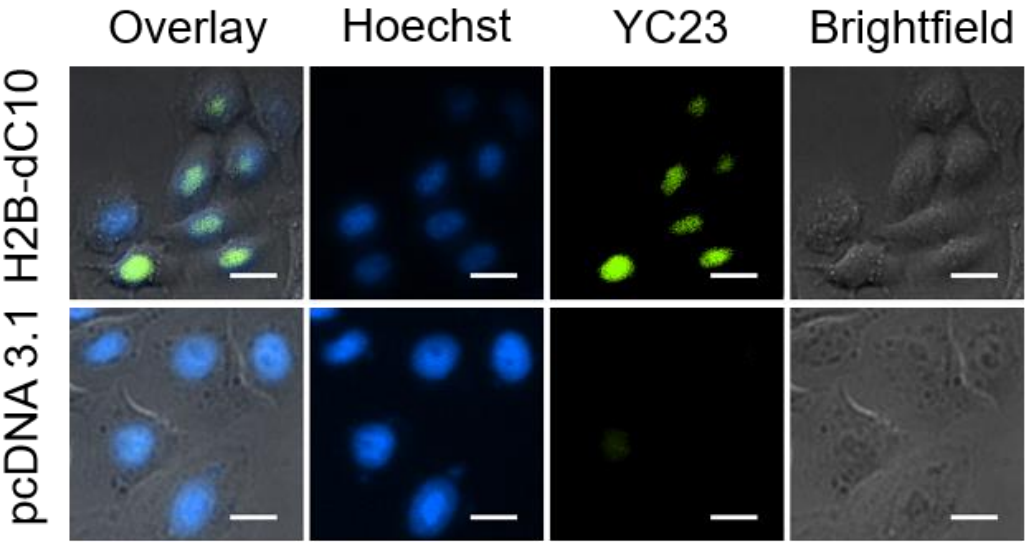


Figure 3.9. Protein labeling images of HeLa cells transfected with histone H2B-dC10 (top) or pcDNA 3.1(+) (bottom) and then labeled with YC23 (10 μM) for 1 h, followed by Hoechst 33342 (8 μM) for 15 min. Cells were observed using a 20X/0.50 NA objective lens. Hoechst was excited at 390 nm and observed in the 485/50 nm channel. YC23 was excited at 485 nm and observed in the 535/45 nm channel. Scale bars: 20 μm.

One potential concern is the incomplete overlap between the green fluorescence from **YC23** and the blue fluorescence of Hoechst in one of the cells (Figure 3.10, white arrow). Here, the brightness observed from YC23 in the area without Hoechst does not necessarily indicate non-specific labelling but may instead indicate that the H2B proteins in this area are not bound to DNA. We hypothesize that this cell may be in the S phase of interphase where DNA is being replicated. During this phase, the DNA, which would interact with the Hoechst dye and exhibit the blue pseudo-coloured fluorescence, is not bound to the histone H2B protein, which would exhibit the green fluorescence due to being labelled by **YC23**. This dissociation from the H2B protein would explain the lack of colocalization in the area.[22] Additionally, incomplete co-localization was also observed in which the green fluorescence

of activated **YC23** overlapped with a smaller portion of the blue fluorescence of the Hoechst stain highlighting the cell nucleus (Figure 3.10, magenta arrow). Additionally, the green fluorescence is weaker in intensity relative to the nearby labelled cell (white arrow). We hypothesize that perhaps in these cells, not all the histone H2B-dC10 was labelled, perhaps due in part to the tag being occluded, insufficient labelling times or a greater concentration of endogenous H2B due to low transfection efficiencies.

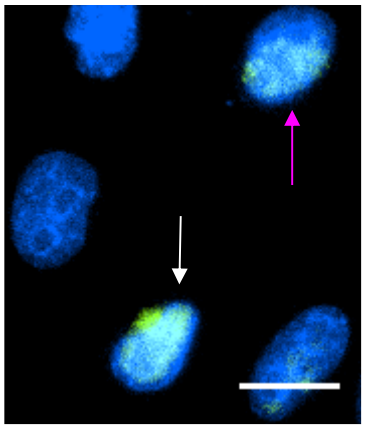


Figure 3.10. Close up of two cells with uneven overlap between the green fluorescence from activated YC23 and the blue fluorescence of Hoechst. Cells were observed using a 40X/0.8 NA objective lens. Scale bar: 20 µm.

Using **YC23**, we also opted to label another dC10 tagged protein, to demonstrate FlARe's ability to label other compartmentalized proteins. To this end, HeLa cells expressing the dC10 tagged red fluorescent protein mNeptune-dC10 were labelled using **YC23**. As mentioned above, a fluorescent protein would allow transfection efficiency to be assessed (see section 3.2.1.). With the addition of the dC10 tag, labelling efficiency could also be assessed independently. Imaging results showed red fluorescence located throughout the cell and localized green fluorescence (Figure 3.11). Given that mNeptune is expressed as a cytoplasmic protein, we expected the red fluorescence to be localized throughout the cell. However, we were surprised to see that the green fluorescence was not equally spread but

instead was more centralized. To confirm this centralization was not localized to the nucleus, a Hoechst counterstain was used to highlight the cell's nucleus. We observed that the green fluorescence did not colocalize with the blue fluorescence and thus **YC23** labelling of mNeptune-dC10 was not localized to the nucleus. Studies that have expressed fluorescent proteins in cells show a concentrated area of fluorescence around the nucleus.[23,24] Therefore, the non-nuclear localized fluorescence observed in the case of mNeptune-dC10 labelling with **YC23** may be due to incomplete labelling, leading to a concentrated region of observed labelling where more of the protein is located. As such, longer labelling times and/or higher concentrations of probe may be necessary for complete labelling throughout the rest of the cell.

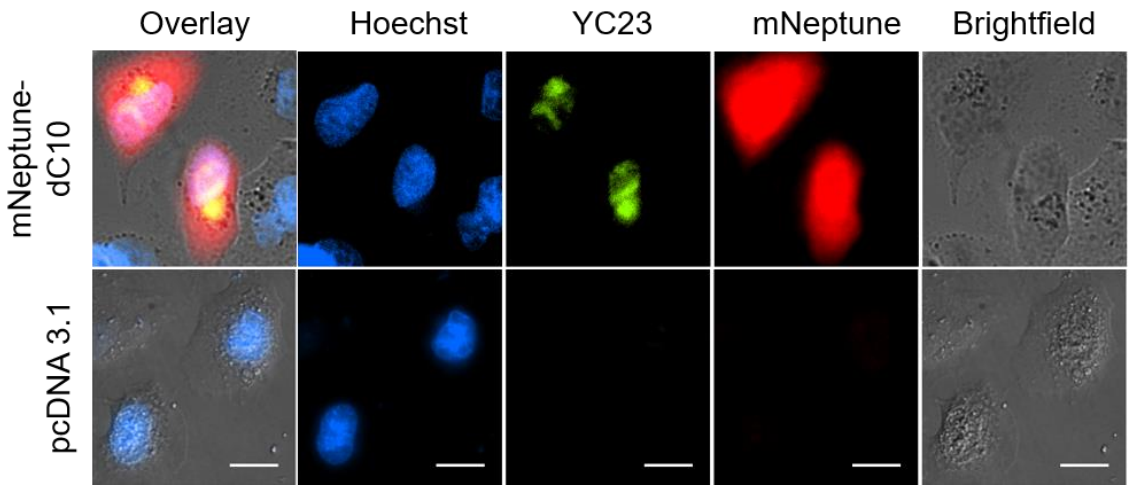


Figure 3.11. HeLa cells transfected with mNeptune-dC10 (top) or pcDNA 3.1(+) (bottom) and then labeled with **YC23** (10 μM) for 1 h, followed by Hoechst 33342 (8 μM) for 15 min. Cells were observed using a 40X/0.8 NA objective lens. Hoechst was excited at 390 nm and observed in the 485/50 nm channel. **YC23** was excited at 485 nm and observed in the 535/45 nm channel, while mNeptune was excited at 575 nm and observed in the 6250/55 nm channel. Scale bars: 20 μm.

Alternatively, the localization of the green fluorescence near the nucleus but not overlapping with the Hoechst stain suggests that the probe may be in the endoplasmic reticulum (ER). We hypothesize that the hydrophobic nature of the probe may have caused it to become trapped in the folds of the ER. Therefore, the selective labelling being observed

may be due to cytoplasmic mNeptune-dC10 encountering the **YC23** probe trapped around the ER. However, since the mNeptune protein is not localized to the ER with the probe, its red fluorescence is localized throughout the entire cell. As such, the observed green fluorescence may represent the *selective* labelling of only a small subset of the available mNeptune-dC10 expressed by the HeLa cells.

3.4. Conclusion

In summary, we first established conditions for the live-cell and no-wash imaging of dC10 tagged proteins being labelled with FlARe probes in HeLa cells. In doing so, the viability of a new dC10* tag was demonstrated as a target peptide for the FlARe labelling technique. Although the dC10* tag showed faster reactivity with FlARe probes *in vitro*, this did not translate to shorter labelling times but offered at least the same level of selectivity as the previous dC10 tag if not better. The studies herein showed that the FlARe reaction commenced after 30 minutes of incubation, and labelling was evident up to 1 h after. Despite being slightly slower than other methods, the labelling times are still on par with the incubation periods suggested by these other methods. Furthermore, FlARe offers a much easier labelling procedure that does not require a washing step, contrary to these other methods.

A new green fluorogenic FlARe probe, **YC23**, was also established. Not only does this expand the colour palette of the current FlARe probes, but the simple substitution of **YC20's** coumarin fluorophore with BODIPY also highlights the synthetic versatility of the FlARe probe design. The implementation of these two novel developments mark the final step towards their establishment in a protein imaging procedure. This is significant for continued

progress for the FlARe labelling method, which offers a non-disruptive and non-toxic fluorogenic protein tagging technique that is amenable to live-cell imaging.

3.5. Experimental

3.5.1. Synthesis of YC20 and YC23

The synthesis of YC20 was previously established in Dr. Yingche Chen[10] and repeated in section 2.4.1.5, scheme 2.11.

YC23 was synthesized by Dr. Yingche Chen and optimized by Sydney L. Acton (M. Sc.).[2,8]

3.5.2. Plasmid Purification

The histone H2B-dC10*, H2B-dC10, and mNeptune-dC10 expression vectors were prepared according to published procedures.[9,10]

E. coli DH5α was first transformed by incubating 50 µL of the bacterial cells with 200 ng of the vector on ice for 30 min. The cells were then heat shocked at 42 °C for 90 s before being placed immediately on ice for 2 min. 500 µL of room temperature LB medium was added to the cells and the solutions were stirred at 37 °C and 200 rpm for 1 h before being plated to screen for ampicillin resistance.

Colonies demonstrating ampicillin resistance were chosen and used to grow a pre-culture by incubating one colony in 5 mL LB Media with 100 µg/mL ampicillin overnight at 37 °C and 200 rpm. The 5 mL preculture was transferred to a larger culture of 500 mL LB medium with 100 µg/mL ampicillin and grown overnight. The vector was then isolated and purified using the PureLink™ HiPure Plasmid Maxiprep Kit from ThermoFisher Scientific by following the manufacturer protocols. DNA pellets were dissolved in 100 µL of the kit's TE buffer for characterization using an Eppendorf BioSpectrometer® and then diluted to 1.0 µg/µL (0.5 µg/µL for pcDNA 3.1(+)) with TE buffer (Table 3.1).

Table 3.1. Concentration and A260/280 of expression vectors isolated and purified using the PureLink™ HiPure Plasmid Maxiprep Kit from ThermoFisher Scientific. The final DNA pellet was dissolved in 100 µL of the kit's TE buffer for characterization.

Expression Vector	Concentration (µg/µL)	$A_{260/280}$
Histone H2B-dC10	1.6918	2.05
Histone H2B-dC10*	1.5964	1.99
mNeptune-dC10	2.2054	2.02
pcDNA 3.1(+)	0.5874	1.87

3.5.3. Original Transfection Conditions

The initial transfection protocol was based on the manufacturer's protocol provided by Thermofisher Scientific for Lipofectamine®2000. However, we aimed to label cells cultured in 60 mm culture dishes, while the manufacturer protocol only provides conditions for up to 6-well culture plates.[25] To address this, we scaled up the conditions suggested in the manufacturer's protocol with the help of the table "Useful Numbers for Cell Culture" by ThermoFisher Scientific regarding the parameters of different cell culture dishes and flasks.[26] The surface area of the 60-mm culture dish (21.5 cm^2) is roughly double that of the 6-well culture plate (9.6 cm^2). Ideally, for a monolayer of cells at 70-90% confluency, most of the surface area will be covered by cells. Therefore, we reasoned that the volumes of DNA and Lipofectamine®2000 should be roughly doubled to account for the increase in surface area from the 6-well culture plates to the 60 mm culture dishes.

As such, the initial transfection protocol is as follows. HeLa cells were grown in DMEM (Life Technologies) supplemented with 10% (v/v) fetal bovine serum (FBS) and 100 IU/mL (1% v/v) penicillin (P) and 100 IU/mL (1% v/v) streptomycin (S). 5.25×10^5 cells in DMEM medium (10% FBS, 1% P/S) were seeded onto a 60-mm culture dish and grown for

24 h until 70-90% confluence. Before transient transfection, both Lipofectamine (24 µg, 1 µg/µL) and the expression vector (7 µg) were diluted in separate serum-free Opti-MEM® Medium (150 µL final volume) solutions. The manufacturer's protocol recommends a range for lipofectamine used, and the higher range was taken to ensure maximal transfection.[25] The diluted DNA was added to the diluted Lipofectamine®2000 reagent and incubated for 5 min at room temperature. Cells were transiently transfected by dropping the combined solution evenly over the cells and gently mixed into the cell medium using a north-south-east-west motion. Cells were then grown for 24 h before being labelled and imaged (*vida infra*).

3.5.4. Improved Transfection Conditions

The improved transfection procedure is as follows, with changes highlighted in **bold.** HeLa cells were grown in DMEM (Life Technologies) supplemented with 10% (v/v) fetal bovine serum (FBS) and 100 IU/mL (1% v/v) penicillin and 100 IU/mL (1% v/v) streptomycin. 5.25×10^5 cells in DMEM medium (10% FBS, 1% P/S) were seeded onto a 60-mm culture dish and grown for 24 h until 70-90% confluence (2 mL medium volume). Before transfection, both Lipofectamine (**10 µg**, 1 µg/µL) and the expression vector (7 µg) were diluted in separate serum-free Opti-MEM® Medium (150 µL final volume) solutions. **The diluted Lipofectamine solution was incubated at room temperature for 30 min to allow liposome formation** before the diluted DNA was added to the diluted Lipofectamine®2000 reagent and incubated for **30 min** at room temperature. Cells were transiently transfected by dropping the combined solution evenly over the cells and gently mixed into the cell medium using a north-south-east-west motion. Cells were then grown for 24 h before being labelled and imaged (*vida infra*).

3.5.5. Cellular Labelling

Before labelling, transfected cells were rinsed with serum-free Opti-MEM® Medium (3×2 mL) to remove residual transfection agent. A 10 µM solution of the FlARe probe diluted in DPBS ((+) Ca^{2+}, (+) Mg^{2+}) or Live Cell Imaging Solution from ThermoFisher Scientific and supplemented with 0.1% (v/v) F127 surfactant and DMSO (final concentration of 0.4% (v/v)) regardless of solution. Cells were incubated with the labelling solution for 45 min at 37 °C and 5% CO_2. Before microscopy, the labelling solution was exchanged for fresh DPBS ((+) Ca^{2+}, (+) Mg^{2+}) or Live Cell Imaging Solution from ThermoFisher Scientific. Regarding Figure 3.6 and Figure 3.8, the labelling solution **was not** exchanged for fresh buffer and was imaged with the labelling solution present.

3.5.6. Fluorescence Microscopy

After labelling HeLa cells, the cells were observed with a Nikon Ni-U Ratiometric Fluorescence Microscopy using a 20X/0.5 NA objective or a 40/0.8 NA objective (specified in the figure captions). When **YC20** was used to label cells, an excitation laser of 438 nm and emission filter of 485/25 nm was used. In the case of **YC23**, an excitation laser of 485 nm and emission filter of 535/45 nm was used. Hoechst 33342 was excited with a 390 nm laser and observed using the 485/25 nm emission filter. Images were obtained using an Ultra-sensitive Andor iXon Ultra 897 cooled EMCCD camera and processed using NIS-Elements software or ImageJ.

3.6. References in Chapter 3

(1) Strmiskova, M.; Tsao, K.; Keillor, J. W. (2018) Rational Design of a Highly Reactive Dicysteine Peptide Tag for Fluorogenic Protein Labelling. *Org. Biomol. Chem.* 16, 6332–6340.

(2) Chen, Y.; Tsao, K.; Acton, S. L.; Keillor, J. W. (2018) A Green BODIPY-Based, Super-Fluorogenic, Protein-Specific Labelling Agent. *Angew. Chem. Int. Ed. Engl.* 57, 12390–12394.

(3) Hopwood, D. (1969) Fixatives and Fixation: A Review. *Histochem. J.* 1, 323–360.

(4) North, A. J. (2006) Seeing Is Believing? A Beginners' Guide to Practical Pitfalls in Image Acquisition. *J. Cell Biol.* 172, 9–18.

(5) Lichtman, J. W.; Conchello, J. A. (2005) Fluorescence Microscopy. *Nat. Methods* 2, 910–919.

(6) James B. Pawley. *Handbook of Biological Confocal Microscopy*, 3rd ed.; (2006).

(7) Strmiskova, M. Development of a Dimaleimide-Based Protein Labelling Technique for Fluorogenic, X-Ray Crystallography and NMR Applications, University of Ottawa, (2016).

(8) Acton, S. L. Design, Synthesis, and Evaluation of Fluorogenic, BODIPY-Based Probes for Specific Protein Labelling in Live Cells, University of Ottawa, (2019).

(9) Guy, J.; Castonguay, R.; Campos-Reales Pineda, N. B.; Jacquier, V.; Caron, K.; Michnick, S. W.; Keillor, J. W. (2010) De Novo Helical Peptides as Target Sequences for a Specific, Fluorogenic Protein Labelling Strategy. *Mol. Biosyst.* 6, 976–987.

(10) Chen, Y.; Clouthier, C. M.; Tsao, K.; Strmiskova, M.; Lachance, H.; Keillor, J. W. (2014) Coumarin-Based Fluorogenic Probes for No-Wash Protein Labeling. *Angew. Chem. Int. Ed. Engl.* 53, 13785–13788.

(11) Asgharian, A.; Banan, M.; Najmabadi, H. (2014) Optimizing a Lipocomplex-Based Gene Transfer Method into HeLa Cell Line. *Cell J.* 15, 372–377.

(12) Raycroft, M. A. R.; Racine, K.; Rowley, C. N.; Keillor, J. W. (2018) Mechanisms of Alkyl and Aryl Thiol Addition to N-Methylmaleimide. *J. Org. Chem.* 83, 11674–11685.

(13) Dodds, E.; Dunckley, M. G.; Naujoks, K.; Michaelis, U.; Dickson, G. (1998) Lipofection of Cultured Mouse Muscle Cells: A Direct Comparison of Lipofectamine and DOSPER. *Gene Ther.* 5, 542–551.

(14) Dalby, B.; Cates, S.; Harris, A.; Ohki, E. C.; Tilkins, M. L.; Price, P. J.; Ciccarone, V. C. (2004) Advanced Transfection with Lipofectamine 2000 Reagent: Primary Neurons, SiRNA, and High-Throughput Applications. *Methods* 33, 95–103.

(15) Yin, W.; Xiang, P.; Li, Q. (2005) Investigations of the Effect of DNA Size in Transient Transfection Assay Using Dual Luciferase System. *Anal. Biochem.* 346, 289–294.

(16) Partearroyo, M. A.; Ostolaza, H.; Goñi, F. M.; Barberá-Guillem, E. (1990) Surfactant-Induced Cell Toxicity and Cell Lysis. A Study Using B16 Melanoma Cells. *Biochem. Pharmacol.* 40, 1323–1328.

(17) Griffin, B. A., Adams, S. R., Tsien, R. Y. (1998) Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* 281, 269–272.

(18) Invitrogen. *TC-FlAsH™ TC-ReAsH™ II In- Cell Tetracysteine Tag Detection Kits*; (2011).

(19) Hoffmann, C.; Gaietta, G.; Zürn, A.; Adams, S. R.; Terrillon, S.; Ellisman, M. H.; Tsien, R. Y.; Lohse, M. J. (2010) Fluorescent Labeling of Tetracysteine-Tagged Proteins in Intact Cells. *Nat. Protoc.* 5, 1666–1677.

(20) Albert Griffin, B.; Adams, S. R.; Jones, J.; Tsien, R. Y. (2000) Fluorescent Labeling of Recombinant Proteins in Living Cells with FlAsH. *Methods Enzymol.* 327, 565–578.

(21) Lee, J. J.; Lee, S. C.; Zhai, D.; Ahn, Y. H.; Yeo, H. Y.; Tan, Y. L.; Chang, Y. T. (2011) Bodipy-Diacrylate Imaging Probes for Targeted Proteins inside Live Cells. *Chem. Commun.* 47, 4508–4510.

(22) Bell, S. P.; Dutta, A. (2002) DNA Replication in Eukaryotic Cells. *Annu. Rev. Biochem.* 71, 333–374.

(23) Llopis, J.; McCaffery, J. M.; Miyawaki, A.; Farquhar, M. G.; Tsien, R. Y. (1998) Measurement of Cytosolic, Mitochondrial, and Golgi PH in Single Living Cells with Green Fluorescent Proteins. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6803–6808.

(24) Masuda, H.; Takenaka, Y.; Yamaguchi, A.; Nishikawa, S.; Mizuno, H. (2006) A Novel Yellowish-Green Fluorescent Protein from the Marine Copepod, Chiridius Poppei, and Its Use as a Reporter Protein in HeLa Cells. *Gene* 372, 18–25.

(25) Invitrogen. Lipofectamine® 2000 Reagent. *Protocol Pub. No. MAN0007824 Rev.1.0*. 2013.

(26) Thermofisher Scientific. Useful Numbers for Cell Culture

https://www.thermofisher.com/ca/en/home/references/gibco-cell-culture-basics/cell-culture-protocols/cell-culture-useful-numbers.html.

3.7. Chapter 3 Appendices

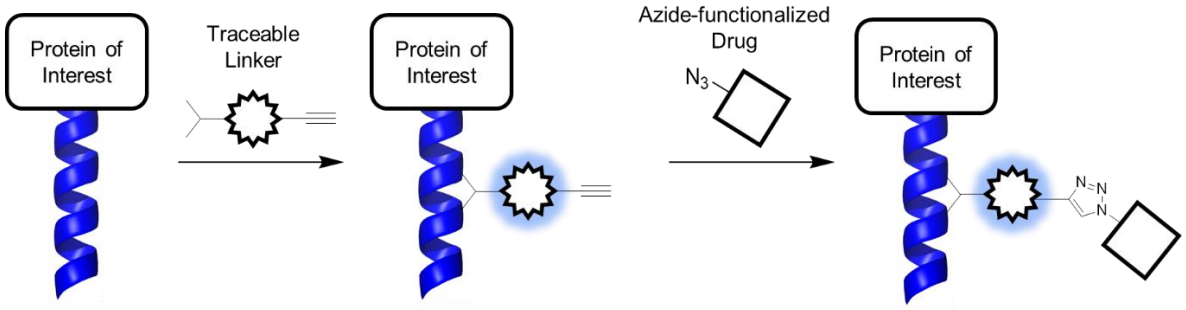
JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Apr 15, 2020

This Agreement between 101 Knudson Drive ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4810291007414
License date	Apr 15, 2020
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Angewandte Chemie International Edition
Licensed Content Title	A Green BODIPY-Based, Super-Fluorogenic, Protein-Specific Labelling Agent
Licensed Content Author	Jeffrey W. Keillor, Sydney L. Acton, Kelvin Tsao, et al

Chapter 4: A Fluorogenic and Site-specific FlARe Bioconjugating Linker



4.1. Statement of Contribution

The contents of this chapter have been adapted from Tsao, K. K., Lee, A. C., Racine, K. É., Keillor, J. W. (2020) Site-Specific Fluorogenic Protein Labelling Agent for Bioconjugation. *Biomolecules*. 10, 369-385.[1] No license was required to reproduce the contents of the article by Kelvin Tsao as Biomolecules is a peer-reviewed open-access journal, and only proper citation is necessary. The work described herein involved the combined efforts of Kelvin Tsao, Ann Lee, and Karl Racine, who worked together under the principal investigator, Dr. Jeffrey Keillor.

The synthesis compound **4.1** and **4.6** was first started by Karl Racine, under the guidance and supervision of Kelvin Tsao. Compound **4.10** and **4.15** was first started by Ann Lee, also under the guidance and supervision of Kelvin Tsao. The final synthesis of these compounds was completed by Kelvin Tsao who also performed the kinetic evaluation and proof-of-principle demonstration of the bioconjugation by SDS-PAGE. Kelvin Tsao also wrote the original manuscript. Dr. Jeffrey Keillor confirmed the evaluation of these compounds, provided valuable insight and feedback during the project, and aided greatly in the final editing and submission of the manuscript. The Keillor group would like to thank the Pezacki Research Group for providing the references necessary for the synthesis of **RhB-N_3**.

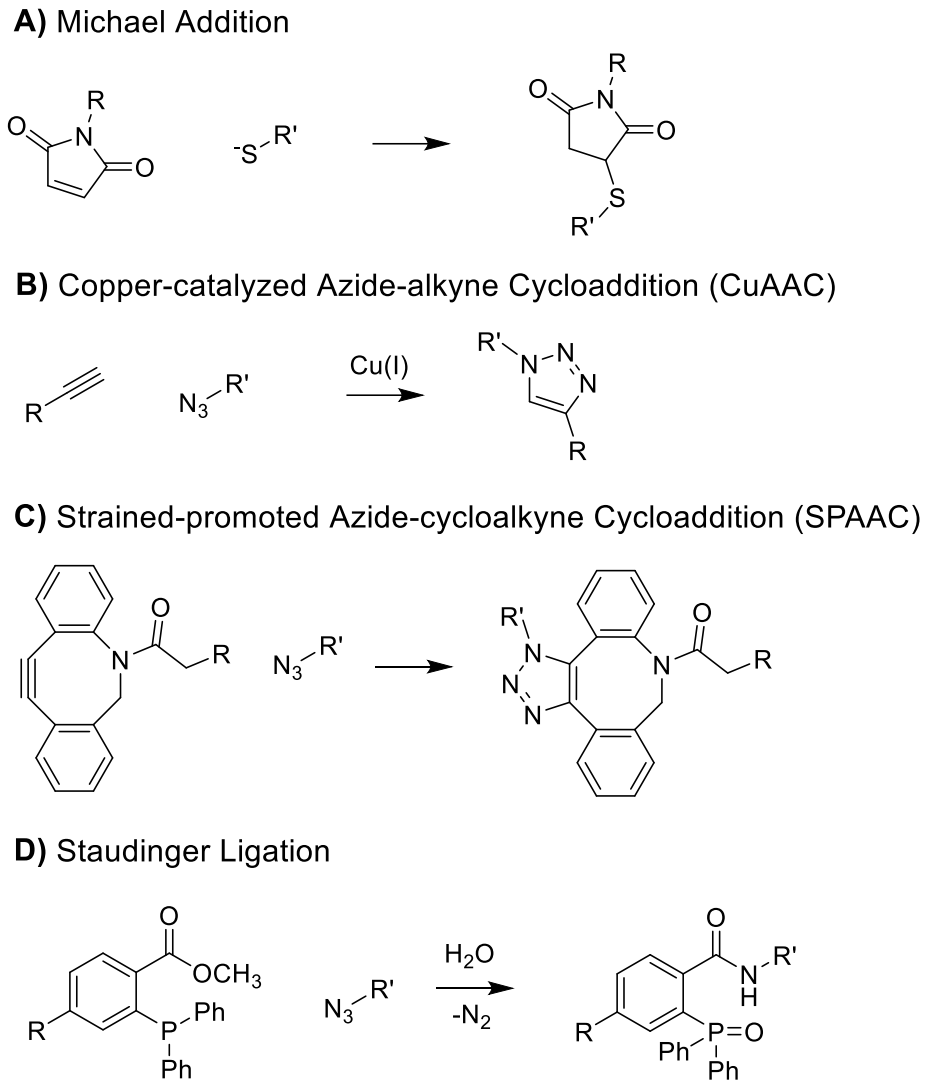
4.2. Background

Bioconjugation involves the conjugation of molecular cargo to a biomolecule, such as a protein, and has contributed to numerous advancements in medical therapeutics and diagnostics. For example, bioconjugates incorporating polyethylene glycol improve the pharmacokinetic properties of biomolecules,[2–5] and conjugating small molecule drugs to antibodies allows for the targeted delivery of therapeutics.[6,7] Additionally, bioconjugation permits the specific and easy attachment of fluorescent probes to proteins or antibodies for imaging.[8] Thus, these bioconjugates can consist of various different biomolecules and cargo.[9,10] Although the strategies involved in making bioconjugates can vary, these methods usually introduce a linker on a protein or antibody that serves as an attachment site for the intended cargo. To ensure that cargo is efficiently and selectively attached to the biomolecules, the linkers typically rely on "click" type reactions.

4.2.1. "Click" Type Reactions

Click reactions are one-pot synthetic reactions that can occur under aqueous conditions and generate minimal unreactive byproducts.[8] These reactions also have a high thermodynamic driving force, and their reactants are bioorthogonal, which facilitates the rapid and irreversible formation of a single reaction product in high yield. The defining characteristics of click reactions are relevant as it allows for specific synthesis to occur within the complex milieu of the cell. Outside a biological context, click reactions are used for the facile and modular formation of synthetic molecules or bioconjugates. Examples of "click" type reactions include the Michael addition, copper-catalyzed azide-alkyne cycloaddition (CuAAC), strained-promoted azide-cycloalkyne cycloaddition (SPAAC) and the Staudinger

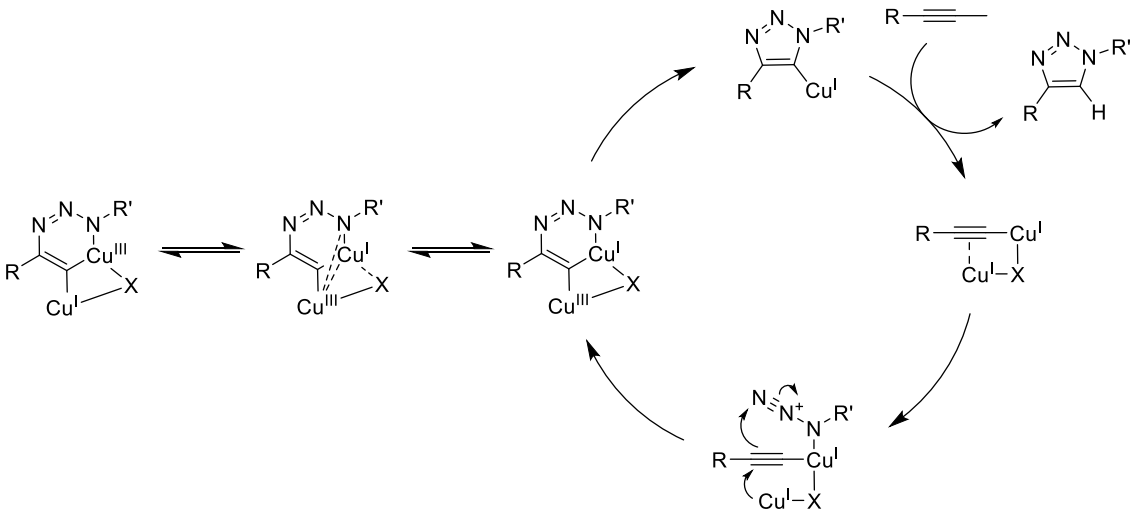
ligation (Scheme 4.1).[11–14] This project uses the CuAAC due to the small size of a simple terminal alkyne, which helps synthetically integrate it into linker design.



Scheme 4.1. Reactions of commonly used "click" type reactions in bioconjugation.

One of the most recently proposed mechanisms for the CuAAC reaction suggests that the reaction starts with the formation of a σ,π-di(Cu^{I}) acetylide, in which two copper(I) ions coordinate to the terminal alkyne through both σ and π bonding (Scheme 4.2).[15] Afterwards, an azide/alkyne/Cu(I) ternary complex is formed, followed by its conversion to a metallacycle intermediate and the oxidization of one Cu(I) to Cu(III). Subsequent reductive ring contraction

results in a Cu(I) triazolide intermediate, which then takes a proton from an alkyne to produce the triazole product and regenerate the Cu(I) catalyst.



Scheme 4.2. Current proposed mechanism of the CuAAC reaction. Figure adapted from reference 15.

4.2.2. Bifunctional Linkers for Bioconjugation

Traditionally, bioconjugation is monitored discontinuously (or *ex situ*), with procedures such as sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), high throughput liquid chromatography (HPLC), or mass spectrometry. However, modern fluorogenic linkers allow the bioconjugation event to be monitored continuously *in situ* by fluorescence. Additionally, conjugation strategies can range from the non-site-specific attachment of cargo, through the modification of natural amino acids, to the highly site-specific attachment of cargo, such as by labelling genetically fused tags with linkers or introducing bioorthogonal unnatural amino acids.[16,17]

One example of a non-specific linker consists of a coumarin fluorophore capable of undergoing both CuAAC and Michael addition reactions through the terminal alkyne and acrylate functional groups, respectively.[18] This linker allowed the initial attachment of azide-

functional cargo followed by reaction with a target protein's native Cys thiol group (Figure 4.1A). Fluorescence was only generated after both the alkyne and acrylate groups had reacted. However, a weakness of this linker is its dependence on the reactivity and accessibility of an intrinsic cysteine residue. Additionally, the non-site-specific manner of attachment (to any accessible cysteine residue) makes it difficult to control the location of the cargo on the protein. In another case, azide-functionalized PEG was conjugated to salmon calcitonin (sCT) through a fluorogenic *N*-propargyl-2,3-(dibromo)maleimide linker (Figure 4.1B).[19] Like the example cited above, only the complete bioconjugate produced fluorescence. However, the reversible nature of the conjugation may deter the efficient formation of bioconjugates. Lastly, a 3-azidocoumarin linker has been used to conjugate PEG-NH_2 to alkyne functionalized bovine serum albumin (BSA) (Figure 4.1C).[20] The initial attachment of 3-azidocoumarin to BSA can be monitored since fluorescence is released upon the reaction of the fluorescence quenching azide group. The weakness of this strategy is inherent to the specific techniques required to introduce unnatural amino acids. Numerous other strategies to synthesize bioconjugates have been developed and reviewed.[21] Given the drawbacks of the above strategies, complementary methods need to be developed to expand the scope of biomolecules that are easily and efficiently conjugated.

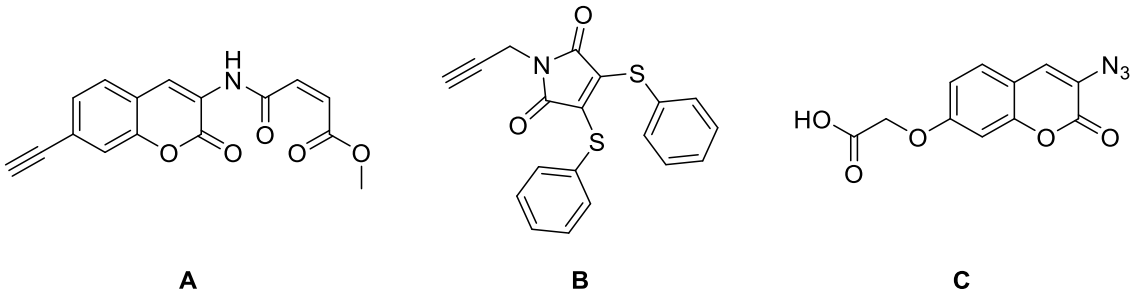


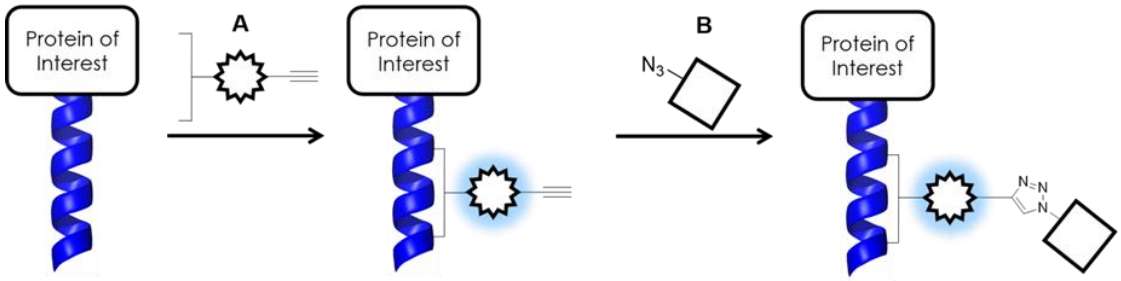
Figure 4.1. Examples of different fluorogenic linkers in previous bioconjugation strategies.[21]

4.2.3. A Bifuctional FlARe Probe for Bioconjugation

A new linker that can be easily integrated into any protein is required for the modular design of novel protein-cargo conjugates. This linker should allow for the site-specific attachment of cargo, so the native function of the protein itself is not disrupted. Additionally, the linker must be fluorogenic for the conjugation to be monitored *in situ*. Therefore, we opted for a strategy involving a fluorogenic linker that reacts site-specifically with a small peptide tag that is genetically fused to either terminus of a target protein. In doing so, the necessity of incorporating unnatural amino acid residues into the protein of interest (POI) is eliminated. Since the linker reacts site-specifically with the target peptide tag, modification of the POI is facile and will minimally disrupt the protein's native function.

The linker proposed herein introduces an alkyne functional group onto a specific protein in both a site-specific and fluorogenic manner (Scheme 4.3). The design consists of a coumarin fluorophore core, functionalized on one end with a dimaleimide moiety to take advantage of the FlARe protein labelling technology. Like previous coumarin-based FlARe probes,[22] the maleimide functional groups serve to quench the coumarin's fluorescence and react with the dC10* tag that is genetically fused to the POI. As a result, the linker is attached to the POI in a site-specific manner. The coumarin is also functionalized with a terminal alkyne on the other end. Therefore, after the heterobifunctional linker reacts with a dC10* tagged protein, an alkyne group is introduced onto the exposed terminal target peptide sequence. Since labelling the dC10* tagged protein with the linker is fluorogenic, the increase in fluorescence signal can be used to monitor the initial conjugation *in situ*. After the initial labelling reaction is complete, a broad array of azide-functionalized cargo can then be conjugated to the modified protein bearing the alkyne functional group through CuAAC.

Therefore, this novel linker provides a strategy for the facile and effective *in vitro* synthesis of fluorescent protein-conjugates, which can be easily adapted to engineered and therapeutically significant proteins.

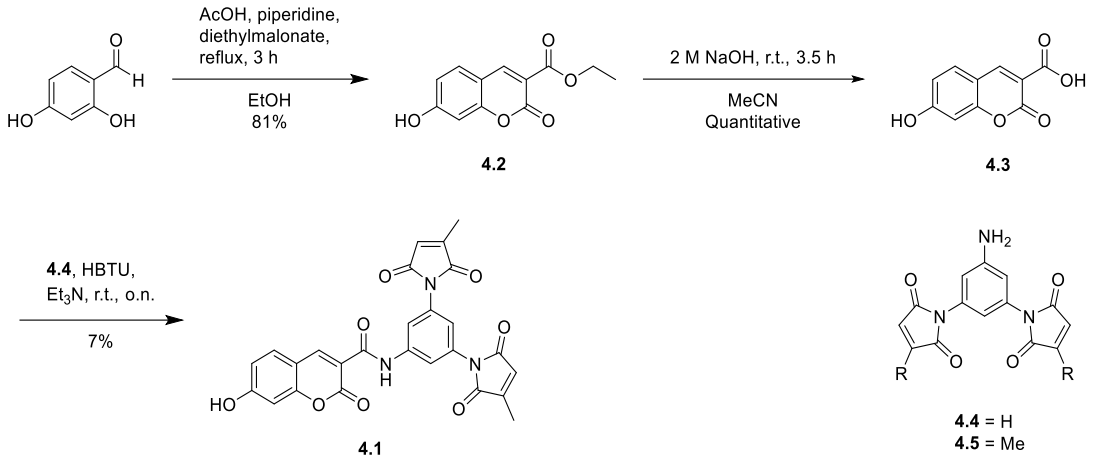


Scheme 4.3. A protein-cargo conjugate is formed by first labelling a protein with a heterobifunctional fluorogenic coumarin linker (A). Azide-functionalized cargo (B) is then attached to the protein through the alkyne moiety on the fluorogenic probe. Thus, fluorogenic and site-specific bioconjugation is achieved.[1]

4.3. Results and Discussion

4.3.1. Initial Design

Our first design consisted of a coumarin fluorophore with the dimaleimide quencher moiety at the C_3 position and a hexyne tail at the C_7 position connected through an ether linkage. The coumarin FlARe probe without the alkyne tail (**4.1**) was first synthesized to check that the fluorophore could be quenched by the dimaleimide quencher (Scheme 4.4). This approach also helps ascertain how adding the terminal alkyne tail would influence the probe's photophysical characteristics.



Scheme 4.4. Synthesis of compound **4.1**.

A Knoevenagel condensation between 4-formylresorcinol and diethylmalonate was performed to provide the coumarin core **4.2** in good yield (Scheme 4.4). The ester was then hydrolyzed to obtained carboxylic acid **4.3** in quantitative yield. At first, we attempted to couple acid **4.3** with aniline **4.4**, which was synthesized using the same strategy to make the methylmaleimide variant **4.5** as established in the development of past coumarin-based FlARe probes (See Experimental, Scheme 4.8).[22] However, the instability of **4.4**'s unsubstituted maleimide groups complicated the synthesis and rendered it impractical. Thus, aniline **4.5** was used instead as its methyl-substituted maleimide groups are more stable against hydrolysis.[23] The FlARe probe **4.1** was synthesized by coupling acid **4.3** to aniline **4.5** using HBTU but this was achieved in low yield, likely due to the low nucleophilicity of the substituted aniline **4.5**.

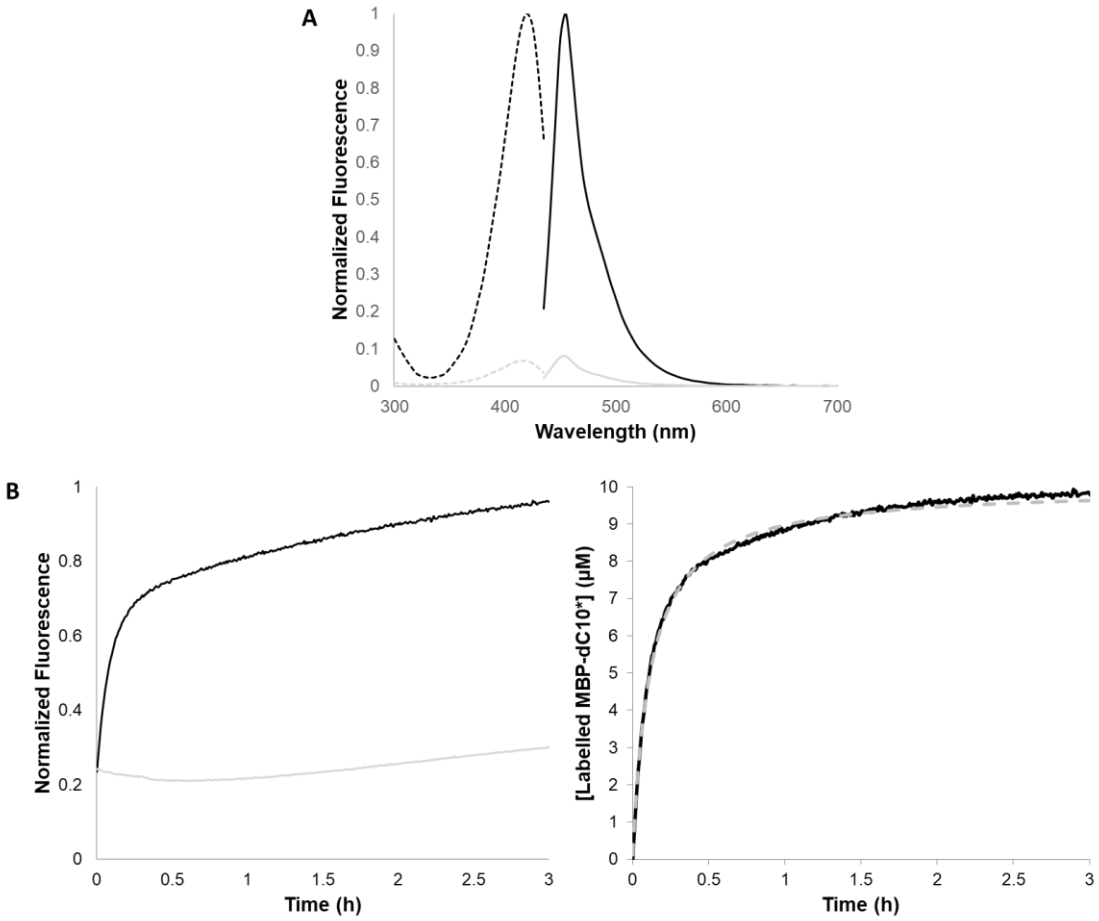


Figure 4.2. (A) Compound **4.1**'s excitation (dotted line) and emission (solid line) spectra of the reaction before (grey line) and after reacting with MBP-dC10* (black line). (B) (Left) Reaction kinetics for the fluorogenic reaction between 10 μM of **4.1** with 10 μM of MBP-dC10* (black), relative to the stability of **4.1** in the absence of MBP-dC10* (grey). (Right) Reaction kinetics (solid black line) fitted to the second-order rate equation (dashed grey line) to calculate the second-order rate constants. Rate constant k_2 = 245 ± 8 $M^{-1}s^{-1}$.[1] Reaction conditions: 50 μM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), 10% DMSO, 25 °C.

In order to validate the activation of the probe **4.1** by FlARe, **4.1** was reacted with an equimolar concentration of maltose binding protein genetically fused to the dC10* target peptide (MBP-dC10*), which served as a representative target protein (Figure 4.2A).[24] **4.1** was observed to react with MBP-dC10* in a fluorogenic manner and showed a fluorescence enhancement (FE) of 12.7. This enhancement is the ratio between the maximal fluorescence exhibited by the FlARe probe in its activated and unreacted forms. The fluorescence enhancement of **4.1** is considerably lower than the FE value of 69 manifested by the similar

FlARe probe that comprised the 7-(diethylamino)coumarin (**YC5**) instead of the 7-hydroxycoumarin fluorophore core (Figure 4.3).[22] However, the overall fluorescence intensities of both activated and quenched **4.1** are over 10-fold higher than those reported by **YC5**.

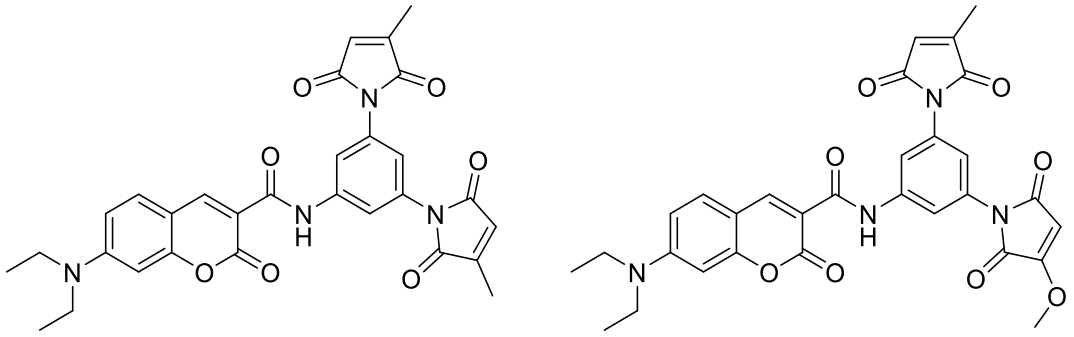
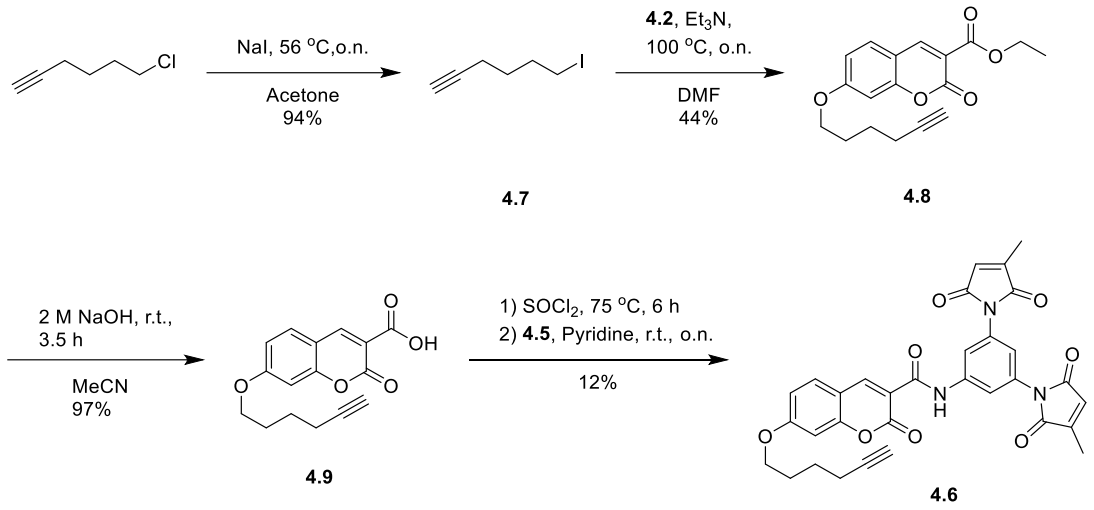


Figure 4.3. Structures of past FlARe probes **YC5** and **YC9**.

The majority of 10 µM of **4.1** reacted with µM dC10* tag within 40 minutes, and the reaction was essentially complete after 2 h (Figure 4.2B). From this reaction's kinetic time course, the second-order rate constant was calculated as $k_2 = 245 \pm 8$ $M^{-1}s^{-1}$. This calculated rate constant can be compared to the similarly structured **YC5** FlARe probe, carries a larger diethylamino substituent compared to the hydroxyl group of **4.1**.[22] A second-order rate constant of $k_2 = 22.2$ $M^{-1}s^{-1}$ can be calculated for the reaction between **YC5** and MBP carrying the slower reacting dC10 tag.[22] However, the dC10* tag has been shown to react nine times faster than the dC10 tag when reacting under similar conditions with **YC9**, which is an asymmetric variant of **YC5** (Figure 4.3).[24] Based on this, we hypothesize that **YC5** would have reacted with MBP-dC10* with a second-order rate constant of $k_2 = 200$ $M^{-1}s^{-1}$. It is interesting to note that the estimated rate constant of **YC5** is lower than that of **4.1**, despite being originally measured at 37 °C, as opposed to 25 °C. Therefore, the second-order rate

constant calculated for the reaction between **4.1** and MBP-dC10* in equimolar concentrations may be greater than the rate constant estimated for **YC5** under similar conditions. This suggests that the changes in the coumarin's C_7 substituent may influence reactivity between the probe and the dC10* tag and allude to the steric influences we observed later. Despite this, encouraged by the successful dC10* tag labelling with **4.1**, we moved forward with attaching the terminal hexyne tail at the C_7 hydroxyl group through an ether linkage using the synthetic strategy outlined for compound **4.6** (Scheme 4.5).



Scheme 4.5. Synthesis of compound **4.6**.

To synthesize **4.6,** we first attached the hexyne group to the coumarin core before appending the dimaleimide quenching intermediate **4.5** (Scheme 4.5). Therefore, **4.7** was synthesized from 6-chlorohexyne, after which **4.8** was obtained in decent yield by displacing the iodide in **4.7** using **4.2**. Like the synthesis of **4.1**, ester **4.8** was hydrolyzed to give **4.9** in near quantitative yield, which was then reacted with $SOCl_2$ to provide the acid chloride. Aniline **4.5** was then reacted with the acid chloride to form **4.6** in low yield, although slightly better than the HBTU-mediated reaction in Scheme 4.4.

As before, the completed FlARe linker and MBP-dC10* were incubated in equimolar concentrations (10 µM), but no fluorescence increase was observed, even after 24 h of incubation. Compared to **4.1**, **4.6** only differs by the presence of the terminal hexyne tail, and so we suspected this functional group of disrupting the coumarin's intrinsic fluorescence. It is known that maleimides quench the fluorescence of nearby coumarin fluorophores through photoinduced electron transfer (PeT), and this has been supported using Density Functional Theory (DFT) calculations (Figure 4.4).[22] Alkynes have also been found to quench the fluorescence of nearby coumarin fluorophores, although the mechanism is not confirmed to be PeT.[25,26] Here, DFT calculations of **4.6** showed that neither the alkyne group's LUMO or HOMO had the correct energy to quench the linker's fluorescence through PeT (Figure 4.4). Therefore, the absence of fluorescence from the coumarin in **4.6** was not likely due to PeT quenching.

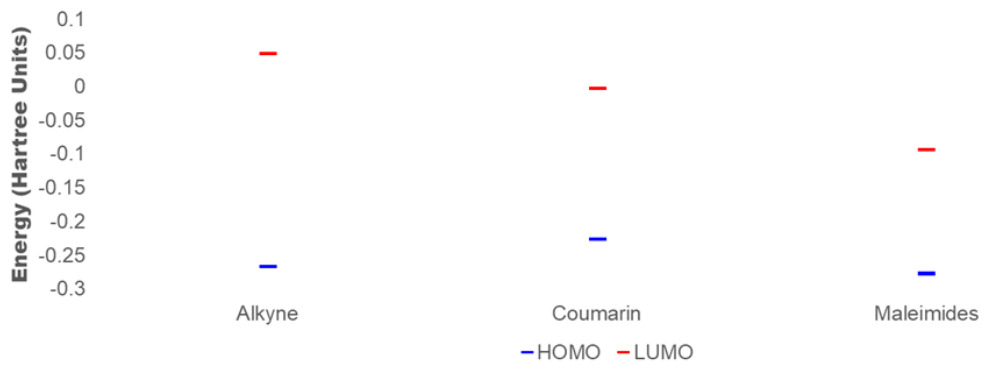
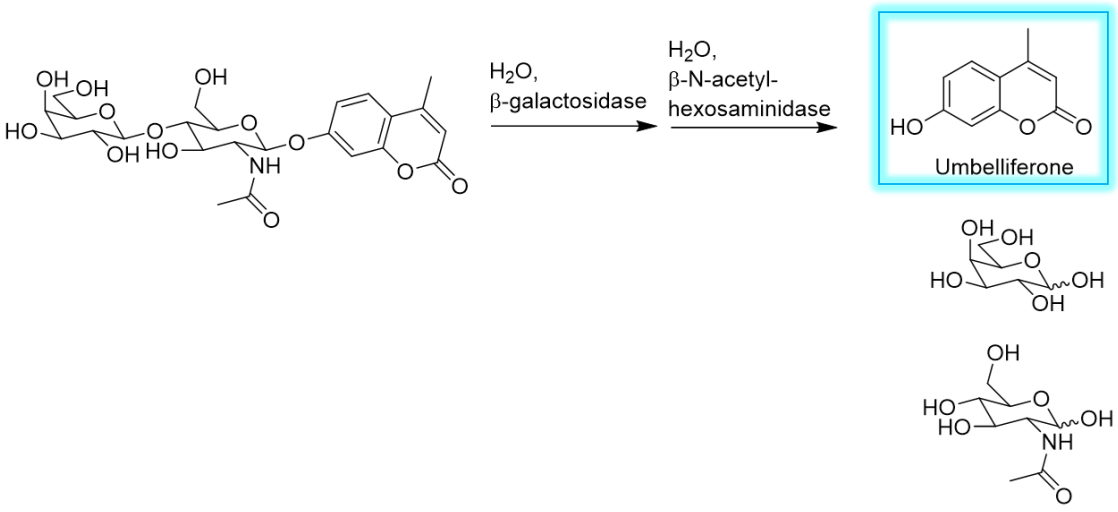


Figure 4.4. HOMO and LUMO energies and their relative positions of the alkyne, coumarin, and maleimides in **4.6**. The geometry of ground-state structures was optimized using density functional theory (DFT) with the B3LYP functional and 6-31G(d) set with water as the solvent. Following this, time-dependent DFT (TD-DFT) calculations at the B3LYP/6-31G(d) level were performed based on the optimized ground state structure. Gaussian 09 was used to perform all calculations.[1]

Alternatively, the coumarin fluorophore's emission has been shown to be very sensitive to the functional group's rigidity at the C_7 position in other studies.[27] Enzyme activity probes involving carbohydrates linked to umbelliferone at the C_7 position exploit the

sensitivity mentioned above as enzymatic cleavage of the carbohydrate releases a highly fluorescent 7-hydroxylcoumarin fluorophore (Scheme 4.6) from a non-fluorescent precursor.[28,29] Similarly, compound **4.1** also demonstrated high fluorescence intensity upon reacting with the dC10 tag. Therefore, converting the alkoxide substituent to an ether functional group likely disrupted the fluorescence of **4.6.** Additionally, the C_7 terminal hexyne tail may further diminish the probe's fluorescence as it offers a means for rotational relaxation. Both changes combined likely led to the negligible fluorescence observed upon reaction of **4.6** with the dC10* tag. Strikingly, Liu's 7-propargyl coumarin probe, which bears a slightly shorter alkyl chain, is fluorescent (Figure 4.5) despite also implementing an ether linkage.[18] This suggests that extending the alkyl chain beyond three carbons on 7-alkoxycoumarins is detrimental to the fluorophore's emission.



Scheme 4.6. Schematic of glycoside hydrolase enzyme activity assay.[29]

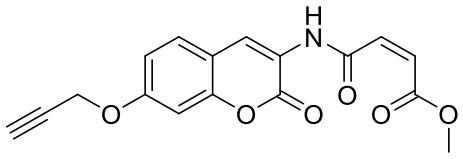


Figure 4.5. Structure of the 7-propargylcoumarin linker developed by Liu.[18]

Although **4.6** did not demonstrate fluorogenic activation by FlARe, the linker was still tested for whether it could successfully introduce an alkyne group to a target protein site specifically, albeit non-fluorogenically. As such, we tested the utility of **4.6** to conjugate rhodamine B azide (RhB-N_3, Figure 4.6) to MBP-dC10* using the CuAAC conjugating reaction. RhB-N_3 was synthesized according a protocol established by Yang, *et al.* (See experimental, Scheme 4.9).[30] Thus, MBP-dC10* was first reacted with **4.6** for 24 h to ensure complete labelling. Next, the CuAAC with RhB-N_3 was performed, and the bioconjugation was evaluated by fluorescent SDS-PAGE (Figure 4.7). A protein band exhibiting only red fluorescence was observed on the gel (Figure 4.7B, lane 3), indicating the formation of the MBP-dC10*-RhB-N_3 conjugate using compound **4.6** as the linker. As expected, no blue fluorescence was observed from the band, which is consistent with the non-fluorogenic nature observed for **4.6 (**Figure 4.7C). Furthermore, no red or blue fluorescence was observed from the protein bands in lane 5 of Figure 4.7. This confirms that without first using **4.6** to introduce the alkyne handle, RhB-N_3 could not be conjugated to MBP-dC10*. Thus, **4.6** still introduced an alkyne group onto MBP-dC10*, which RhB-N_3 could then be conjugated through CuAAC, even though **4.6** was not observed to be fluorogenic as intended.

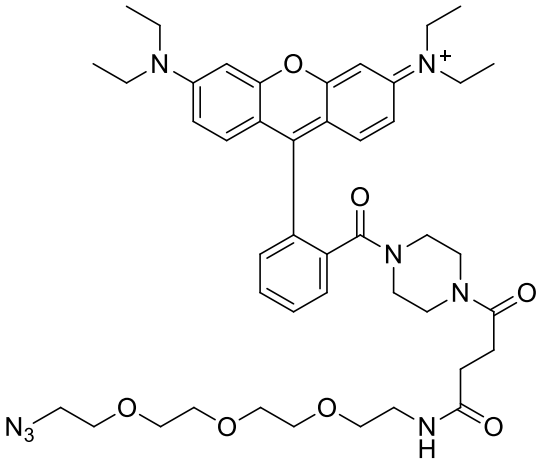


Figure 4.6. Structure of RhB-N_3. The fluorescent azide-functionalized cargo used in the bioconjugation studies shown herein.

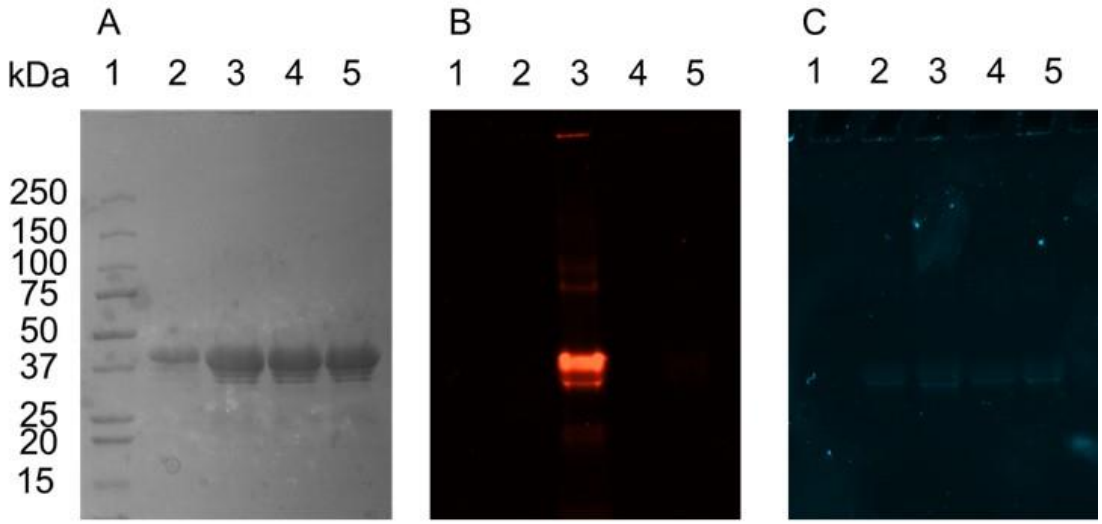
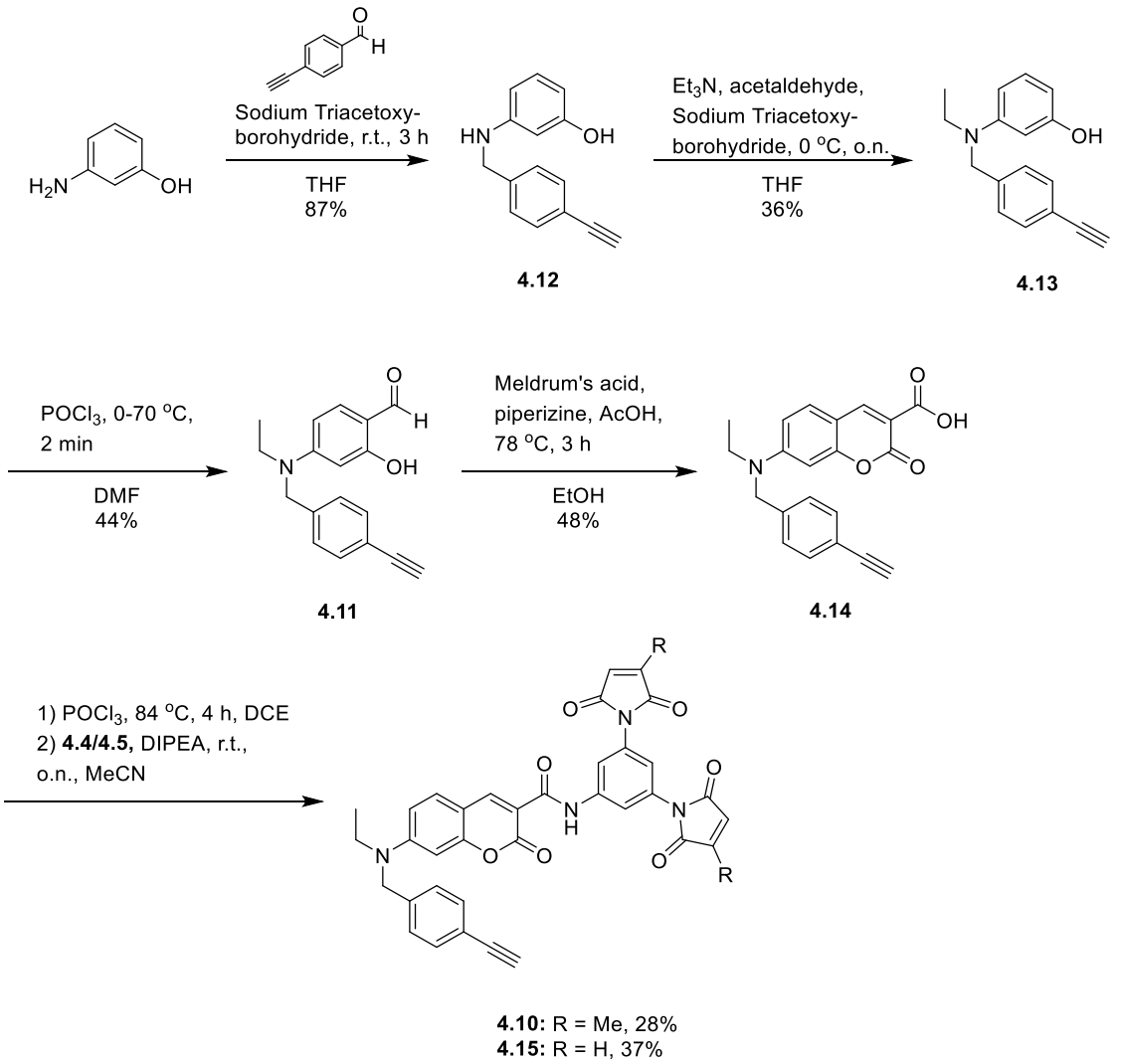


Figure 4.7. SDS PAGE gel of the protein conjugates obtained by the reaction **4.6** with MBP-dC10* using FlARe and then conjugating rhodamine B azide (RhB-N_3) to MBP-dC10* through **4.6** by CuAAC. A) Coomassie stained gel showing MBP-dC10* bands. B) Unstained gel excited with Green Epi Illumination and imaged using a 605/50 filter. C) Unstained gel excited with Blue Epi Illumination and imaged with a 530/28 filter. Lanes: 1) Prestained protein standards 2) MBP-dC10* 3) MBP-dC10* reacted with **4.6** and then conjugated to RhB-N_3 4) MBP-dC10* reacted with **4.6** 5) MBP-dC10* incubated with RhB-N_3 under CuAAC conditions.[1]

4.3.2. Improved Design

Although **4.6** was a viable conjugating linker, its non-fluorogenic nature made it difficult to evaluate how rapidly and efficiently the alkyne group was introduced. Thus, a second fluorogenic bioconjugation linker was designed in which the alkyne group was

attached to the coumarin core's C_7 position as a 4-ethynylbenzylamine substituent. We envisioned this functional group to exhibit less rotational flexibility and maintain the coumarin's intrinsic fluorescence. The design of the fluorogenic linker **4.10** integrates these modifications (Scheme 4.7).



Scheme 4.7. Synthesis of compound **4.10** and **4.15**.

Before **4.10** was formed, the alkyne group was first introduced by synthesizing the salicylaldehyde **4.11**. Thus, 3-aminophenol was reacted with *p*-ethynylbenzaldehyde in a reductive amination using sodium triacetoxyborohydride (STAB) to afford the secondary

amine **4.12** in good yield (Scheme 4.7). This secondary amine was then capped with acetaldehyde using STAB in a second reductive amination to yield the tertiary amine **4.13** in moderate yield. To obtain the functionalized salicylaldehyde **4.11**, a Vilsmeier-Haack reaction was performed. This afforded **4.11** in good yield, which was then followed by a Knoevenagel condensation with Meldrum's acid to obtain coumarin **4.14.** The carboxylic acid **4.14** was converted to the acid chloride with $POCl_3$ before **4.14** was reacted with the di(methylmaleimide) aniline intermediate **4.5** to form the amide bond and afford **4.10** in good yield.

4.10 was reacted with MBP-dC10* in equimolar amounts of 10 µM, and a 37-fold fluorescence enhancement was observed after complete reaction with the tag (Figure 4.8A). In addition to indicating the linker's successful fluorescence activation by FlARe, this result also represents a significant improvement over the FE observed with **4.1**. However, the complete reaction of **4.10** with MBP-dC10* took nearly 10 h, providing a rate constant of k_2 = 8.4 ± 0.7 $M^{-1}s^{-1}$. This rate constant is more than an order of magnitude lower than the rate constant measured for **4.1** (Figure 4.8C), suggesting that the labelling reaction is hindered by the modified alkyne tail of **4.6**. On a practical level, this means that the bioconjugation of micromolar concentrations of a protein conjugate would require at least two days to complete, and so a quicker alternative was designed. As mentioned before, unsubstituted maleimides react more quickly than methylmaleimides.[23] Therefore, the dimaleimide aniline was used to replace the di(methylmaleimide) aniline moiety leading to the design of the linker **4.15** (Scheme 4.7).

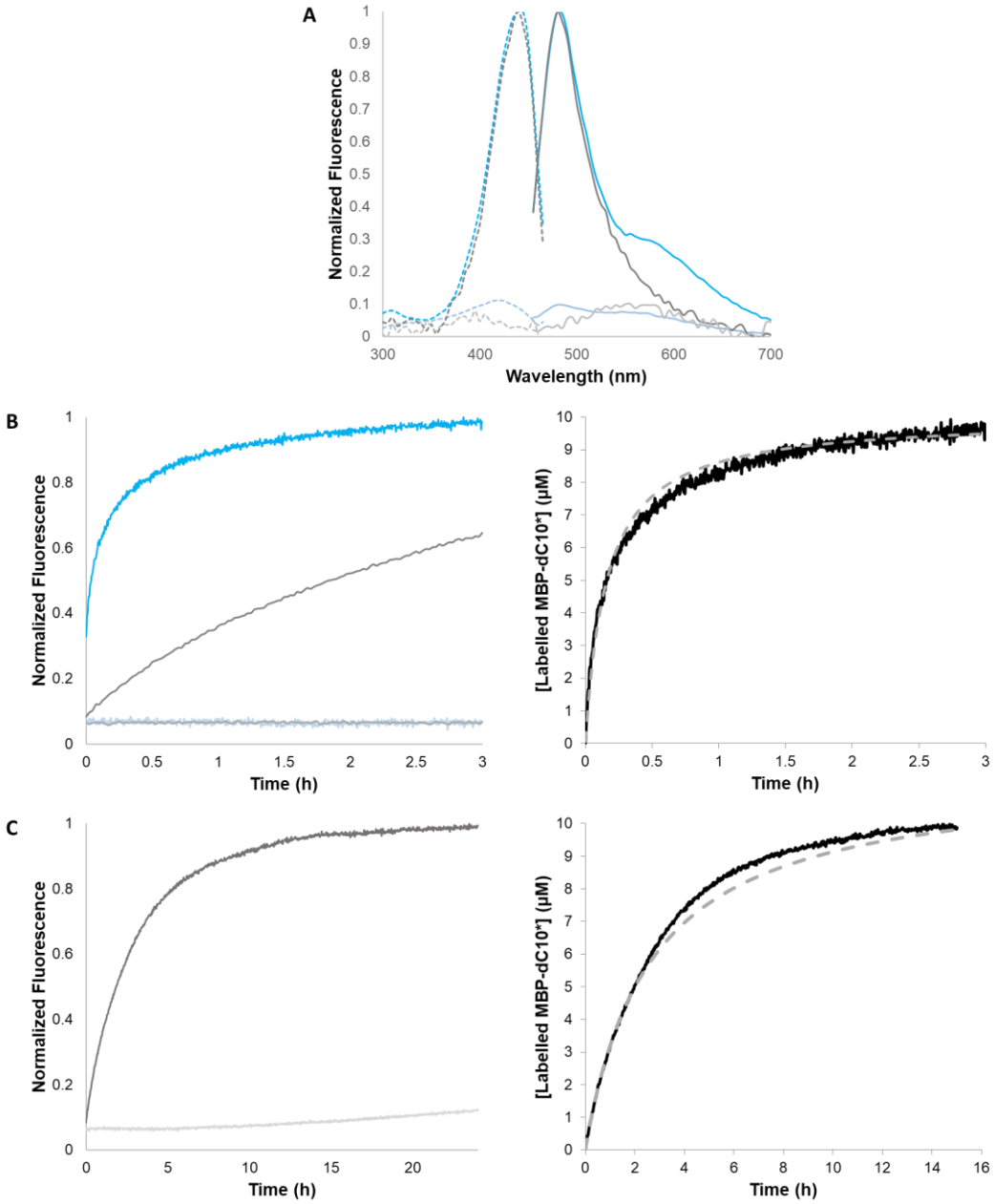


Figure 4.8. (A) Excitation (dashed line) and emission (solid line) spectra of 10 μM **4.10** (grey lines) and **4.15** (blue lines) before (light) and after (dark) the reaction with 10 μM MBP-dC10*. (B) (Left) Kinetic time course for the fluorogenic reaction between 10 μM of **4.10** (grey lines) and **4.15** (blue lines) with 10 μM of MBP-dC10* (dark lines), relative to the stability of **4.10** and **4.15** in the absence of MBP-dC10* (light lines). (Right) Kinetic time course for the reaction between 10 μM of **4.15** with 10 μM of MBP-dC10* (solid black line) fitted to the second-order rate equation (dashed grey line). Rate constant (k_2) = 180 ± 10 $M^{-1}s^{-1}$. (C) (Left) Complete kinetic time course for the fluorogenic reaction between 10 μM of **4.10** with 10 μM MBP-dC10* (dark grey line), relative to the stability of **4.10** in the absence of MBP-dC10* (light grey line). (Right) Kinetic time course for the reaction between 10 μM of **4.10** with 10 μM of MBP-dC10* (dark solid line) fitted to the second-order rate equation (dashed grey line). Rate constant (k_2) = 8.4 ± 0.7 $M^{-1}s^{-1}$.[1] Reaction conditions: 50 μM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), 10% DMSO, 25 °C.

The novel fluorogenic linker **4.15** was synthesized using the same strategy outlined in Scheme 4.7. Specifically, **4.14** was first converted to the acyl chloride using $POCl_3$ and then reacted with aniline **4.4** to obtain **4.15** in decent yield. After reacting **4.15** with equimolar amounts of MBP-dC10* at 10 µM, a 10-fold FE was observed. To account for the concentration of the activated adduct, a quantum yield of $\Phi = 0.093$ was measured for the final MBP-dC10*-**4.15** adduct, which is 57 times higher than the quantum yield of unreacted **4.15** (Figure 4.9). The FE of **4.15** was three-fold lower than the FE measured for **4.10** but still comparable to the FE measured for **4.1** (Figure 4.8A). The reaction between **4.15** and MBP-dC10* at 10 µM was complete after only 30 minutes, and a rate constant of $k_2 = 180 \pm 10\ M^{-1}s^{-1}$ was measured (Figure 4.8B). This rate constant is comparable to the one measured for the reaction between **4.1** and MBP-dC10*, suggesting that the accelerating effects of switching to the unsubstituted maleimide counteracted the inhibiting effects of the alkyne tail.

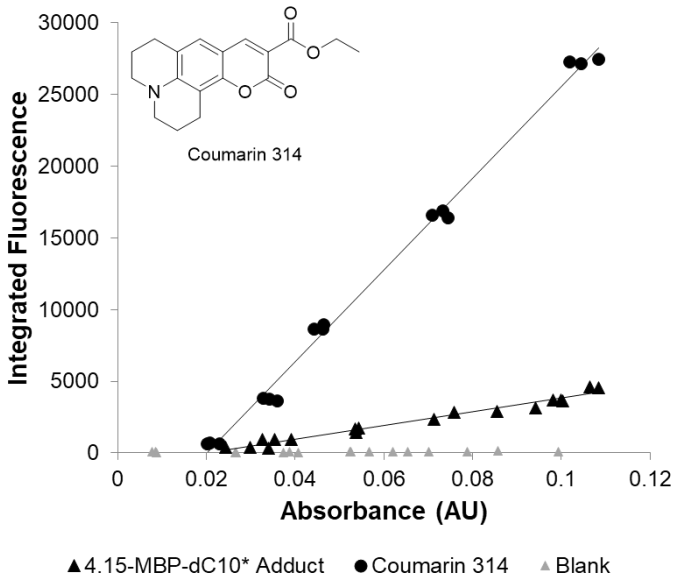


Figure 4.9. Graph of the integrated fluorescence vs. absorbance of the coumarin 314 standard (black circles, $m_{std} = 32.0 \pm 0.7 \times 10^4$, $R^2 = 0.9946$) and the **4.15**-MBP-dC10* adduct (black triangles, $m_{sample} = 4.8 \pm 0.2 \times 10^4$, $R^2 = 0.9704$) for determining the quantum yield of activated **4.15**. The blank (grey triangles, $m_{sample} = 0.08 \times 10^4$, $R^2 = 0.6032$) consists of unreacted **4.15** in the absence of MBP-dC10*. Structure of the coumarin 314 standard is shown in the top left corner of the graph.

On a practical level, switching to the unsubstituted maleimide decreased the total time required to prepare a protein conjugate to one day. Encouraged by both the labelling speed and fluorogenic nature of **4.15**, the FlARe linker's utility for bioconjugation was evaluated similarly to **4.6**. Therefore, **4.15** and MBP-dC10* was incubated together at 10 µM concentrations for 1 h, after which RhB-N_3 was then conjugated to MBP-dC10* through the alkyne handle introduced by **4.15** using CuAAC. The bioconjugation was analyzed by SDS-PAGE, which showed a protein band that was both red *and* blue fluorescence (Figure 4.10B and 4C, lane 4). This indicated that the red fluorescence rhodamine B was successfully conjugated to MBP-dC10* through the blue fluorescent **4.15** coumarin linker. The blue fluorescence of this protein band was notably dimmer than the protein band containing MBP-dC10* labelled with only **4.15** (Figure 4.10C, lane 2). We hypothesized that this attenuation

of fluorescence is due to a FRET interaction between the rhodamine B group of the model cargo and the coumarin fluorophore of **4.15**. The FRET interaction between these two fluorophores has been identified previously in literature.[31–33] Furthermore, a comparison between the coumarin emission spectrum of the MBP-dC10*-**4.15** adduct and the excitation spectrum of rhodamine B (Figure 4.11) reveals the significant spectral overlap necessary for FRET. Molecular modelling of the triazole adduct formed after CuAAC coupling suggests that the fluorophores would be seperated by less than 30 Å, which also suggests that efficient resonance energy transfer could occur. Lastly, the protein band in lane 5 exhibits neither red nor blue fluorescence, which confirms that successful conjugation of RhB-N_3 to MBP-dC10* requires the prior labelling of the target protein with linker **4.15** (Figure 4.10B, lane 5). These results show that **4.15** was successful in site-specifically introducing an exposed alkyne group in a fluorogenic manner, which allowed for the attachment of azide-functionalized cargo, RhB-N_3.

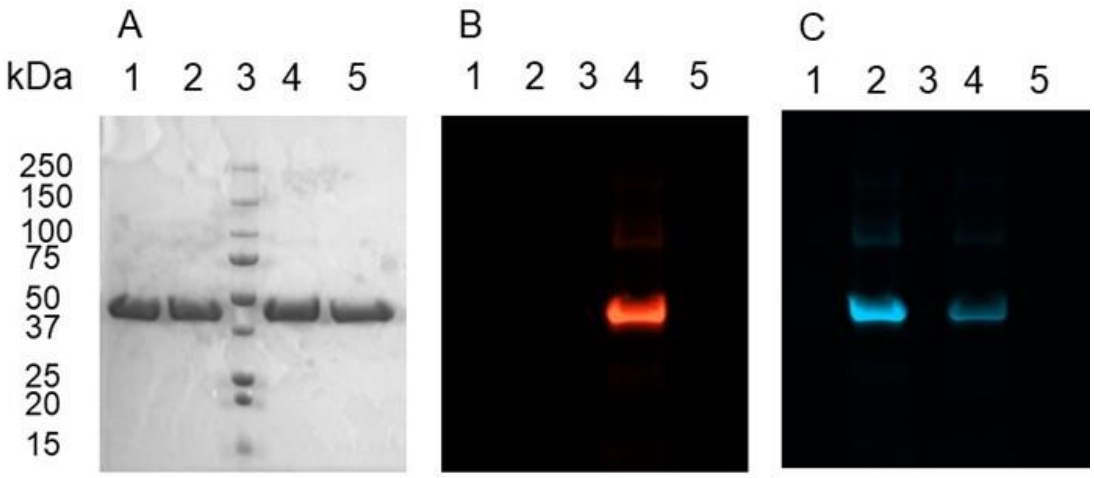


Figure 4.10. SDS PAGE gel of the protein conjugates obtained by the reaction of **4.15** with MBP-dC10* using FlARe and then conjugating rhodamine B azide (RhB-N_3) to MBP-dC10* through **4.15** by CuAAC. A) Coomassie stained gel showing MBP-dC10* bands. B) Unstained gel imaged using a 605/50 filter. C) Unstained gel excited with Blue Epi Illumination and imaged with a 530/28 filter. Lanes: 1) MBP-dC10* 2) MBP-dC10*-**4.15** reaction product 3) pre-stained protein molecular weight markers 4) MBP-dC10* reacted with **4.15** followed by conjugation to RhB-N_3 5) MBP-dC10* incubated with RhB-N_3 under CuAAC conditions.[1]

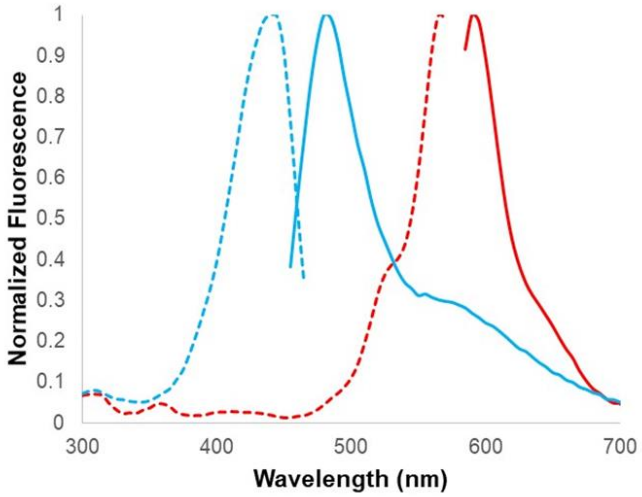


Figure 4.11. Normalized excitation (dashed line) and emission (solid line) spectra for RhB-N_3 (red) and the **4.15**-MBP-dC10* adduct (blue). Note the significant overlap between the coumarin emission spectrum of the adduct (solid blue line) and the excitation spectrum of rhodamine B (dashed red line). The **4.15**-MBP-dC10* adduct was formed by reacting 10 µM of **4.15** with 10 µM MBP-dC10* for 4 h, as outlined in the experimental, and spectra were recorded thereafter (Reaction conditions: 50 µM TCEP, 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4), 10% DMSO, 25 °C.). The spectra recorded for RhB-N_3 were obtained at 25 µM under the same reaction conditions (λ_{ex} = 565 nm, λ_{em} = 588 nm).[1]

In addition to its rapid kinetics, **4.15** also demonstrated notable resistance to hydrolysis under assay conditions over the course of 3 h. Thus, we performed the CuAAC between **4.15**

and RhB-N_3 *before* conjugation of the triazole product to MBP-dC10* since we hypothesized that the maleimides would remain intact after the CuAAC. After the initial CuAAC between **4.15** and RhB-N_3, MBP-dC10* was added in equimolar concentrations to **4.15** and the **4.15**-RhB triazole product was allowed to react with the dC10* tag for 1 h. As before, SDS-PAGE was used to analyze the bioconjugation (Figure 4.12). A protein band that was red and blue fluorescent ((Figure 4.12B and C, lane 2) indicated the successful formation of the MBP-dC10*-rhodamine conjugate. Like the previous experiment, which involved conjugating the RhB-N_3 cargo to MBP-dC10* *after* the protein was labelled with **4.15**, the blue fluorescence of this band is dimmer than the protein band consisting of MBP-dC10* labelled with only **4.15** in lane 3. This reflects the FRET interaction occurring between the fluorophores, which had been observed above. Furthermore, the band from lane 4 showed no red fluorescence, which demonstrates the importance of **4.15** for successful bioconjugation.

Notably, the fluorescence intensity of *all* the protein bands shown in Figure 4.12 appears to be lower than those from the previous experiment, seen in Figure 4.10. However, the density of the Coomassie stained protein bands of both gels are similar (Figure 4.10A vs. Figure 4.12A), suggesting that this observation is not due to a difference in gel loading. Thus, the labelling reaction of the **4.15**-RhB triazole adduct with MBP-dC10* (Figure 4.12) is hypothesized to be less efficient than the reaction of **4.15** with MBP-dC10* (Figure 4.10), resulting in a lower proportion of protein being labelled after 1 hour of incubation, and thus a less intense signal being observed. The steric bulk of the adduct pre-formed between the fluorogen **4.15** and rhodamine B may slow down the subsequent labelling reaction of MBP-dC10* like what was observed when comparing **4.1** and **4.10** (*vida supra)*. This would decrease the efficiency of labelling MBP-dC10* with the triazole product, resulting in less

observable fluorescence. Thus, results show that FlARe protein-labelling reaction can be performed before or after the CuAAC reaction. However, it appears that higher efficiency is attained when the protein is labelled before attaching the cargo to the labelled protein through the alkyne group.

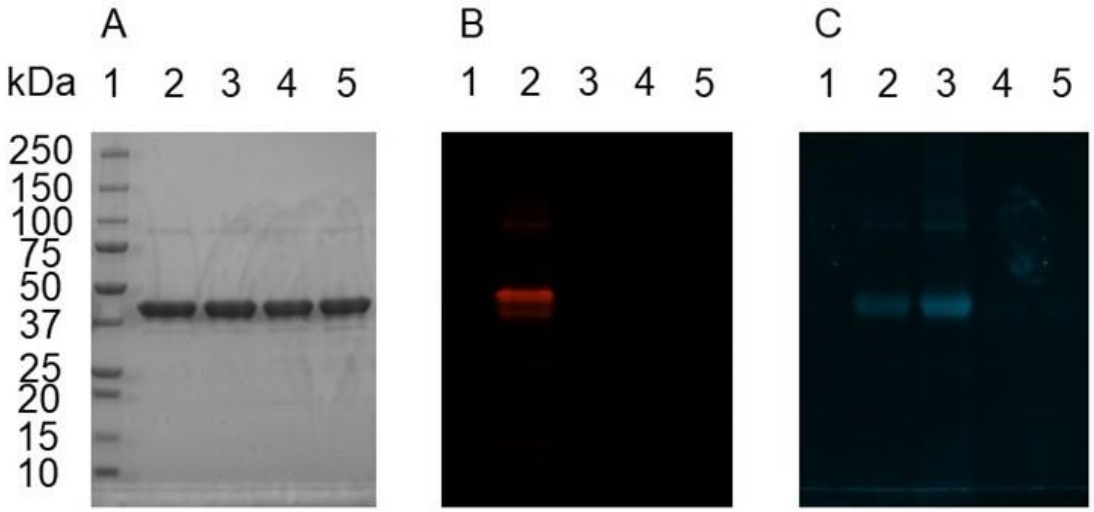


Figure 4.12. SDS PAGE analysis of protein conjugates obtained after first reacting **4.15** with rhodamine B azide (RhB-N_3) using CuAAC, after which the adduct was conjugated to MBP-dC10* by FlARe. A) Coomassie stained gel showing MBP-dC10* bands. B) Unstained gel excited with Green Epi Illumination and imaged using a 605/50 filter. C) Unstained gel excited with Blue Epi Illumination and imaged with a 530/28 filter. Lanes: 1) Prestained protein standards 2) MBP-dC10* conjugated the **4.15**-RhB-N_3 product of the CuAAC. 3) MBP-dC10* reacted with **4.15** that was first subjected to a mock CuAAC without RhB-N_3. 4) MBP-dC10* incubated with RhB-N_3 under CuAAC conditions. 5) MBP-dC10*.[1]

4.4. Conclusions

In summary, a new bioconjugation method was developed that comprises the initial introduction of an alkyne group onto a target protein in a site-specific, rapid, and fluorogenic manner. In doing so, the subsequent site-specific attachment of a broad range of azide-functionalized cargo can be performed. The design of the first bifunctionalized conjugating agent (**4.6**) was non-fluorescent, likely due to its long, flexible hexyne tail. Despite its non-fluorogenic behaviour, **4.6** was still able to introduce an alkyne group to the test protein MBP-

dC10*, through which the model cargo RhB-N_3 could be attached. In the second design, a shorter and more rigid alkynyl amine substituent replaced the long alkyne tail of **4.6,** while dimaleimide aniline replaced the di(methylmaleimide) aniline moiety, resulting in the heterobifunctional fluorogenic linker **4.15**. As a result, the new fluorogenic linker **4.15** demonstrated a 10-fold increase in fluorescence upon reacting with the dC10* target peptide ($\Phi = 0.093$). Therefore, the fluorescence increase was used to monitor the introduction of the exposed alkyne group *in situ*. Model cargo RhB-N_3 was subsequently attached to MBP-dC10* through the site-specifically installed alkyne handle presented by **4.15**. We also demonstrated that azide-functionalized cargo could also be reacted with **4.15** by CuAAC *before* reacting with the target protein. Overall, this novel fluorogenic linker allows the site-specific introduction of an exposed alkyne group into a dC10* tagged protein of interest. This approach is easily adaptable for the synthesis of new therapeutically relevant protein-cargo conjugates and the initial protein modification step can be easily monitored *in situ* by fluorescence due to the fluorogenic design of the bioconjugating linker. Thus, our novel FlARe-based linker provides an adaptable, modular, and facile method for the *in vitro* synthesis of future protein-cargo conjugates.

4.5. Experimental

All protocols in the experimental can be found from reference 1.

*4.5.1. Expression and Purification of Test Protein MBP-dC10**

The expression and purification of MBP-dC10* have been outlined in Chapter 2, section 2.4.3.

4.5.2. Determination of Fluorescence Enhancement Ratios

A Synergy H4 Hybrid Multi-Mode Microplate Reader with excitation and emission monochromators set at 9-nm bandpass was used to record fluorescence spectra and intensity at 25 ºC. The initial excitation and emission spectra were recorded using a solution of 10 µM labelling reagent (compounds **4.1**, **4.6**, **4.10**, or **4.15**) or 25 µM RhB-N_3 in 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4) with 10% DMSO and 50 µM TCEP. The final excitation and emission spectra were recorded after a solution of 10 µM labelling reagent in 50 mM HEPES (I = 100 mM adjusted with NaCl, pH 7.4) with 10% DMSO and 50 µM TCEP with the addition of 10 µM MBP-dC10* was allowed to react for 4 h (in the case of **4.1** or **4.15**) or 24 h (in the case of **4.6** or **4.10**). Fluorescence enhancement (FE) was calculated by taking the ratio of the fluorescence intensity at the maximum emission between the initial and final spectra.

4.5.3. Kinetic Studies

A Synergy H4 Hybrid Multi-Mode Microplate Reader with excitation and emission monochromators set at a 20-nm bandpass was used to carry out kinetic experiments at 25 ºC. A 96-well plate was prepared with solutions containing 10 µM MBP-dC10* in 50 mM HEPES buffer (I = 100 mM adjusted with NaCl, pH 7.4) with 10% DMSO and 50 µM TCEP.

Labelling reagent was added to a final concentration of 10 µM immediately before recording. Samples were excited at 435 nm for **4.10** and **4.15** or 415 nm for **4.1** and emission intensity was followed at 485 nm for **4.10** and **4.15** or 460 nm for **4.1** as a function of time. Second-order rate constants (k_2) for the reaction between the label and MBP-dC10* were obtained by fitting all time-dependent fluorescence curves to a second-order rate equation (Equation 4.1).

$$\text{Equation } \mathbf{4.1}\text{: } [Labelled\ protein] = P_f - \frac{1}{k_2 \times t + \frac{1}{P_f}}$$

where P_f is the expected final concentration of labelled protein at the end of the reaction.

4.5.4. Bioconjugation of rhodamine B azide to MBP-dC10-**4.15** or **4.6** adduct*

A protocol from the procedures developed by Nasheri and Yang was adapted to conjugate rhodamine B azide (RhB-N_3) to MBP-dC10* labelled with **4.6** or **4.15.**[30,34] A solution of 2.5 µM MBP-dC10* in 50 µM TCEP in HEPES buffer (50 mM HEPES (I = 100 mM NaCl), pH 7.4) at 25 °C was incubated with an aliquot of 5 µM **4.15** in a total volume of 200 µL for 1 h. The reaction was monitored using a Synergy H4 Hybrid Multi-Mode Microplate Reader with excitation (435 nm) and emission (485 nm) monochromators and filters set at a 20-nm bandpass. Upon reaction completion, the mixture was transferred to a 500-µL 10-kDa Molecular Weight Cut Off (MWCO) Amicon filter and buffer exchanged into fresh HEPES buffer three times and adjusted to 100 µL after the third time. A 500-µL click mixture (1.0 mM $CuSO_4$, 100 µM TBTA, 1.0 mM TCEP-HCl, 50 µM RhB-N_3, PBS pH 7.4, 0.25% DMSO) was prepared, and 100 µL of this click mixture was added to the buffer exchanged solution containing the MBP-dC10*-**4.15** adduct. After allowing the click reaction to progress for 2 h in the dark with gentle shaking at 25 °C, 500 µL of acetone, chilled to -80 °C, was added and the protein was precipitated overnight at -80 °C. To pellet the protein, the mixture was spun

at 14 000 rpm for 15 min at 4 °C. The supernatant was then removed, and any excess acetone was dried off under a gentle stream of air. The pellet was then washed with 100 µL MeOH chilled to -80 °C before pelleting again at 14 000 rpm for 15 min at 4 °C. The pellet was taken up in 35 µL SDS PAGE loading buffer (50 mM TCEP) and shaken for 15 mins at room temperature. The samples were spun at 14 000 rpm for 5 min and then boiled at 90 °C for 5 min before 10 µL was loaded onto a 10%-acrylamide SDS-PAGE gel and run at 120 V for 1.5 h.

In the case of **4.6,** the reaction between MBP-dC10* and **4.6** was allowed to react for 24 h and the reaction was not monitored.

*4.5.5. Bioconjugation of Rh-B-N_3-**4.15** adduct to MBP-dC10**

The protocol for performing the CuAAC between Rhodamine B azide (RhB-N_3) and **4.15** was adapted from procedures developed by Nasheri and Yang.[30,34] A 500-µL click mixture (1.0 mM $CuSO_4$, 100 µM TBTA, 1.0 mM TCEP-HCl, 50 µM RhB-N_3, PBS pH 7.4, 0.25% DMSO) was first prepared and 90 µL of this click mixture was incubated with 10 µL of a 200 µM **4.15** stock (final concentration 20 µM). The click reaction was allowed to progress in the dark at r.t. for 2 h with gentle shaking. Next, 100 µL of a 10-µM MBP-dC10* solution was added to the mixture before allowing the entire solution to react for 1 h at room temperature. After, 500 µL of acetone chilled to -80 °C was added, and the protein was precipitated overnight at -80 °C. The mixture was spun at 14 000 rpm for 15 min at 4 °C to pellet precipitated protein before removing the supernatant and drying off any excess acetone under a gentle stream of air. The pellet was washed with 100 µL MeOH chilled to -80 °C before pelleting again at 14 000 rpm for 15 min at 4 °C. The pellet was dissolved in 35 µL of SDS PAGE loading buffer (50 mM TCEP) and sonicated for 3 minutes before being shaken

for 15 mins at room temperature. The samples were spun at 14 000 rpm for 5 min and then boiled at 90 °C for 5 min before 10 µL was loaded onto a 10% acrylamide SDS-PAGE gel and ran at 120 V for 1.5 h.

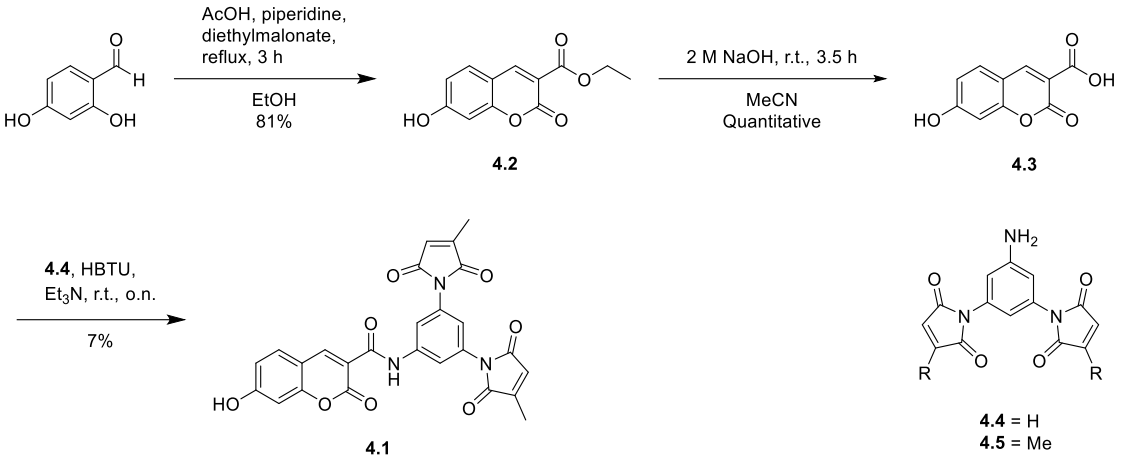
4.5.6. Synthesis

All reagents and solvents for reactions were used as received unless otherwise stated. Dichloromethane, dimethylformamide, methanol and acetonitrile were dried using a solvent purification system from LC Technology Solution Inc.

Reactions were monitored by thin layer chromatography (TLC) using E. Merck silica gel 60F254 precoated aluminum plates and visualized by illumination with a short-wavelength ultraviolet light or long-wavelength visible light. Flash column chromatography was performed on ZEOCHEM® silica gel 60 (ECO 40-63 µm) using ethyl acetate/n-hexane or methanol/dichloromethane as eluting solvents.

Nuclear magnetic resonance (NMR) spectra were recorded in deuterated DMSO or deuterated chloroform at ambient temperature. The experiments were performed either on a Bruker Avance 400 Fourier Transform Spectrometer operating at 400 MHz for ^{1}H and at 100 MHz for ^{13}C or on a Bruker Avance 600 Fourier Transform Spectrometer operating at 600 MHz for ^{1}H and at 150 MHz for ^{13}C. EI-MS spectra were recorded on a Kratos concept mass spectrometer for both low resolution and high-resolution mass spectra. ESI-MS spectra were recorded on a Waters Micromass Q-TOF mass spectrometer. Melting points were measured on an EZ-Melt automated melting point apparatus and are uncorrected.

4.5.6.1. Synthesis of Compound 4.1 (Scheme 4.4)



Compounds **4.2** and **4.3** were synthesized according to the procedures laid out by Abdizadeh *et al*.[35]

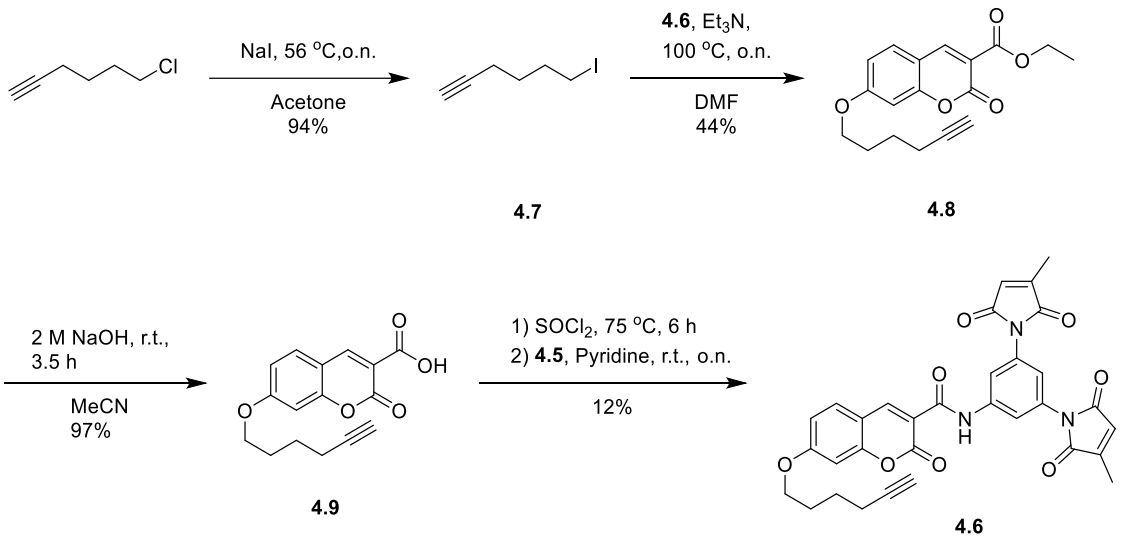
Compound 4.2: Piperidine (0.02 eq, 19.8 µL, 0.2 mmol) and acetic acid (0.02 eq, 11.4 µL, 0.2 mmol) was added to a solution of 4-hydroxysalicylic acid (1.38 g, 10.0 mmol) and diethylmalonate (1.0 eq, 1.53 mL, 10.0 mmol) in EtOH (5 mL). The solution was stirred at room temperature for 20 min and then heated to 78 °C and stirred for 3 h and then cooled to room temperature. The product was crystallized as an orange solid in 81% yield. ^{1}H NMR (400 MHz, DMSO) δ 11.04 (s, 1H), 8.73 – 8.53 (m, 1H), 7.71 (d, *J* = 8.6 Hz, 1H), 6.80 (dd, *J* = 8.6, 2.3 Hz, 1H), 6.75 – 6.47 (m, 1H), 4.22 (q, *J* = 7.1 Hz, 2H), 1.25 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (100 MHz, DMSO) δ 164.5, 163.4, 157.6, 156.9, 149.9, 132.6, 114.5, 112.6, 110.9, 102.3, 61.3, 14.6. HRMS (ESI): calcd for $C_{12}H_{10}O_5Na$ ([MNa]$^+$): 257.0426, found: 257.0414.

Compound 4.3: A solution of NaOH (10.0 eq, 63.6 mg, 1.59 mmol) dissolved in water (2.0 mL) was added dropwise to a solution of **4.2** (50 mg, 0.159 mmol) dissolved in acetonitrile (3.0 mL). The solution was stirred at room temperature for 3.5 h and then diluted with 10 mL water followed by acidification with 10 mL 1 M HCl. The aqueous solution was

extracted with EtOAc (3 × 30 mL) and the combined organic layers were dried with $MgSO_4$. After removing the solvent on the rotary evaporator, compound **4.3** was obtained as a yellow solid in quantitative yield. ^{1}H NMR (400 MHz, DMSO) δ 12.00 (bs, 1H), 8.64 (s, 1H), 7.73 (d, *J* = 8.4 Hz, 1H), 6.83 (d, *J* = 8.4 Hz, 1H), 6.72 (s, 1H). ^{13}C NMR (100 MHz, DMSO) δ 164.32 (d, *J* = 14.1 Hz), 157.75, 156.99, 148.96, 131.91, 114.17, 112.69, 110.53, 101.83. ^{13}C NMR (100 MHz, DMSO) δ 164.3, 157.8, 157.0, 149.0, 131.9, 114.2, 112.7, 110.5, 101.8. HRMS (ESI): calcd for $C_{10}H_6O_5Na$ ($[MNa]^+$): 229.0113, found: 229.0128.

Compound 4.1: HBTU (1.1 eq, 52.0 mg, 0.138 mmol) was added to a solution of compound **4.3** (26.0 mg, 0.130 mmol) in anhydrous DMF (2.5 mL). After stirring the mixture at room temperature for 10 min, **4.5** (1.5 eq, 59.0 mg, 0.188 mmol) was added and the mixture was stirred for an additional 20 min before Et_3N (3.0 eq, 50 µL, 0.380 mmol) was added. The mixture was then stirred overnight and then followed with diluting the mixture in EtOAc (100 mL) and washing with 1 M HCl (2 × 100 mL). The organic layer was dried with $MgSO_4$ and then the solvent was removed using the rotary evaporator. The crude was purified by flash chromatography using a gradient elution system of 0-5% MeOH in DCM to obtain compound **4.1** as a pale-yellow solid in 7% yield. ^{1}H NMR (600 MHz, DMSO) δ 11.18 (bs, 1H), 10.86 (s, 1H), 8.83 (s, 1H), 7.86 (d, *J* = 8.7 Hz, 1H), 7.78 (d, *J* = 1.7 Hz, 2H), 7.12 (t, *J* = 1.8 Hz, 1H), 6.92 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.85 (d, *J* = 2.0 Hz, 1H), 6.83 (q, *J* = 1.8 Hz, 2H), 2.09 (d, *J* = 1.7 Hz, 6H). ^{13}C NMR (150 MHz, DMSO) δ 170.3, 169.4, 164.1, 161.2, 160.7, 156.5, 149.6, 148.4, 145.9, 138.7, 132.6, 132.3, 127.6, 120.0, 116.8, 114.6, 114.1, 111.2, 102.0, 10.9. HRMS (ESI): calcd for $C_{26}H_{17}N_3O_8Na$ ($[MNa]^+$):522.0913, found: 522.0929. **mp**: decomp. at 267.7-271.1 °C.

4.5.6.2. Synthesis of Compound 4.6 (Scheme 4.5)



Compound 4.7: The synthesis of compound **4.7** was done according to Jiang et al.[36] and spectral data is consistent with what was reported. **4.7** was synthesized using a Finkelstein reaction by adding 6-chloro-1-hexyne (8.30 mmol, 1.00 mL) and NaI (58.1 mmol, 8.71 g, 7 eq.) to dry acetone (100 mL). The solution was brought to reflux and stirred overnight before the acetone was removed using a rotary evaporator. The residue was then partitioned between hexanes (100 mL) and water (100 mL) and the organic layer was washed with water (100 mL) and brine (100 mL). $MgSO_4$ was used to dry the organic layer and the solvent was removed using a rotary evaporator to obtain compound **4.7** as a transparent liquid in 94% yield (7.78 mmol, 1.62 g). ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.21 (t, J = 6.9 Hz, 2H), 2.23 (td, J = 7.0, 2.6 Hz, 2H), 2.03 – 1.86 (m, 3H), 1.70 – 1.57 (m, 2H).

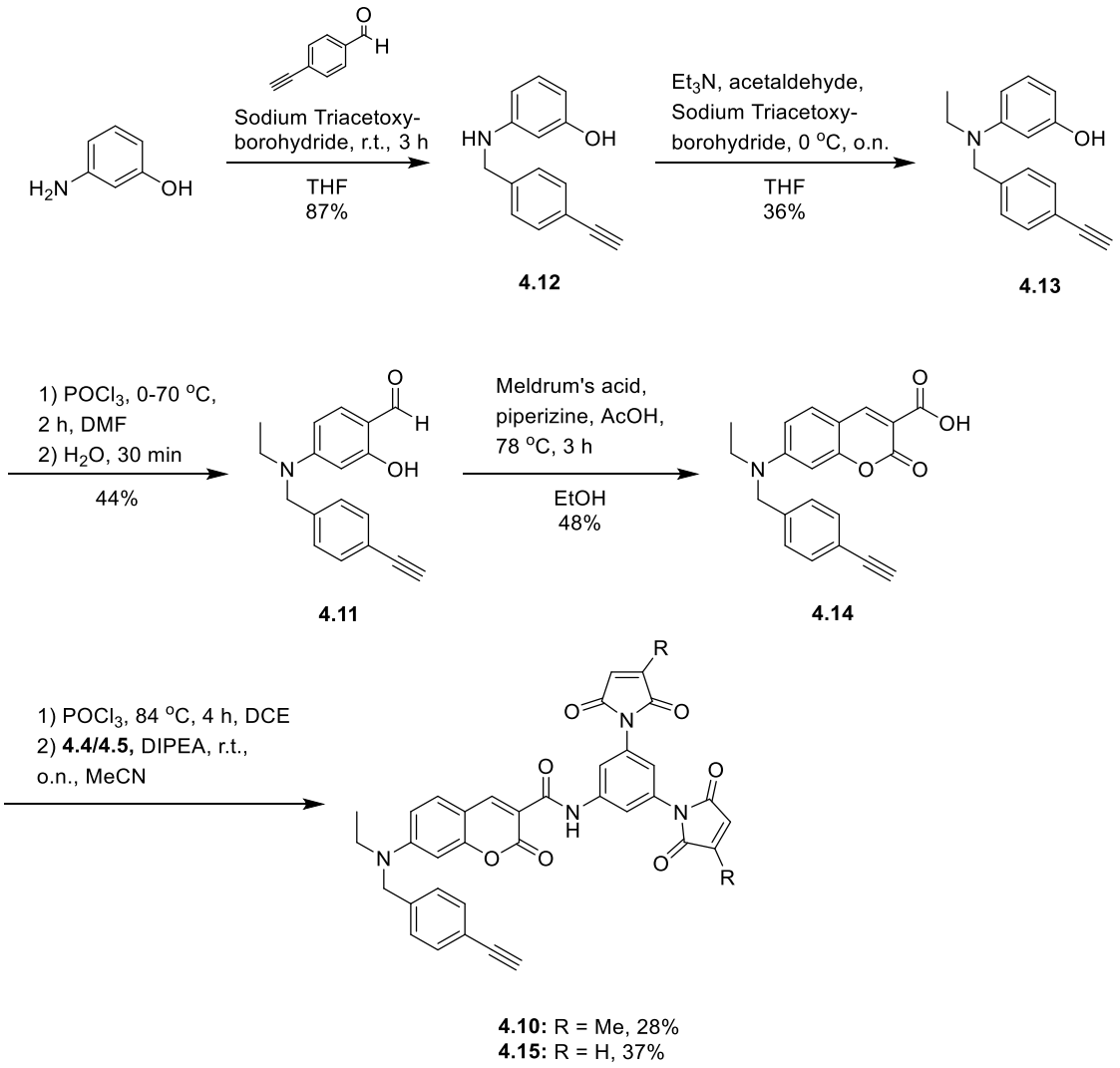
Compound 4.8: A solution of compound **4.8** (1.00 mmol, 234 mg) in acetonitrile was prepared and compound **4.7** was added. After adding K_2CO_3 (1.2 mmol, 165 mg, 1.2 eq.) and a catalytic amount of 18-crown-6 (0.1 mmol, 26.4 mg), the mixture was brought to reflux and stirred overnight. The acetonitrile was removed using a rotary evaporator and the residue was

partitioned between EtOAc (20 mL) and water (20 mL). The aqueous layer was extracted with EtOAc (3 × 30 mL) and $MgSO_4$ was used to the combined organic layers before the solvent was removed. The residue was purified by flash chromatography using a gradient elution system of 10-30% EtOAc in hexanes. Compound **4.8** was obtained as a white solid in 44% yield (0.437 mmol, 138 mg). ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 8.50 (s, 1 H), 7.49 (d, *J* = 8.72 Hz, 1 H), 6.88 (dd, *J* = 8.67, 2.20 Hz, 1 H), 6.80 (d, *J* = 2.06 Hz, 1 H), 4.40 (q, *J* = 7.15 Hz, 2 H), 4.08 (t, *J* = 6.22 Hz, 2 H), 2.30 (td, *J* = 6.91, 2.55 Hz, 2 H), 1.97 (m, 3 H), 1.74 (tt, *J* = 7.30 Hz, 2 H), 1.40 (t, *J* = 7.10 Hz, 3 H). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 164.6, 163.5, 157.6, 157.2, 149.0, 130.7, 114.0, 114.0, 111.6, 100.8, 83.7, 68.9, 68.3, 61.7, 27.9, 24.9, 18.1, 14.3. HRMS (ESI): calcd for $C_{18}H_{18}O_5Na$ ($[MNa]^+$): 337.1052, found: 337.1074. **mp**: 92.6 – 93.6 °C.

Compound 4.9: NaOH (4.93 mmol, 197 mg, 10 eq.) was added as a 2 M solution in water (2.5 mL) to a solution of **4.8** (0.437 mmol, 138 mg) dissolved in acetonitrile. After stirring the reaction for 3.5 h it was partitioned between water (10 mL), 1 M HCl (10 mL) and EtOAc (30 mL). The aqueous phase was extracted with EtOAc (3 × 30 mL) and then $MgSO_4$ was used to dry the combined organic phases. After filtering to remove the drying agent, a rotary evaporator was used to remove the remaining EtOAc to obtain compound **4.9** as a white solid in 97% yield (0.425 mmol, 122 mg). ^{1}H NMR (400 MHz, DMSO) δ ppm 12.97 (br. s., 1 H), 8.71 (s, 1 H), 7.82 (d, *J* = 8.72 Hz, 1 H), 7.01 (m, 2 H), 4.14 (t, *J* = 6.22 Hz, 2 H), 2.79 (t, *J* = 2.55 Hz, 1 H), 2.24 (td, *J* = 6.69, 2.20 Hz, 2 H), 1.83 (tt, *J* = 8.50, 7.70 Hz, 2 H), 1.61 (tt, *J* = 7.30 Hz, 2 H). ^{13}C NMR (100 MHz, DMSO) δ ppm 164.1, 163.9, 157.2, 156.8, 148.9, 131.5, 113.7, 113.5, 111.5, 100.6, 84.1, 71.4, 68.0, 27.4, 24.4, 17.3. HRMS (ESI): calcd for $C_{16}H_{14}O_5Na$ ($[MNa]^+$): 309.0739, found: 309.0727. **mp**: 170.0 – 170.9 °C.

Compound 4.6: Compound **4.9** (0.349 mmol, 100 mg) was dissolved in $SOCl_2$ (6 mL) and then the solution was brought to reflux and stirred for 6 h. After, excess $SOCl_2$ was removed using the rotary evaporator and the residue was dissolved in pyridine (5 mL). Compound **4.5** (0.349 mmol, 149 mg, 1.0 eq) dissolved in pyridine (5 mL) was then added and the mixture was left to stir overnight. After taking the solution up in EtOAc (20 mL) it was washed with 1 M HCl (3 × 20 mL), 5% $CuSO_4$ (20 mL) and brine (20 mL). $MgSO_4$ was used to dry the organic layer and the solvent was removed on the rotary evaporator to obtain the crude residue. **4.6** was obtained as a pale-yellow solid in 12% yield (0.0431 mmol, 25 mg) after purification by flash chromatography using a 5:4:1 hexanes/DCM/EtOAc mixture. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 10.96 (s, 1 H), 8.92 (s, 1 H), 7.85 (d, *J* = 1.86 Hz, 2 H), 7.62 (d, *J* = 8.72 Hz, 1 H), 7.28 (t, *J* = 1.86 Hz, 1 H), 6.96 (dd, *J* = 8.72, 2.35 Hz, 1 H), 6.89 (d, *J* = 2.25 Hz, 1 H), 6.50 (q, *J* = 1.67 Hz, 2 H), 4.11 (t, *J* = 6.27 Hz, 2 H), 2.31 (td, *J* = 6.96, 2.65 Hz, 2 H), 2.18 (d, *J* = 1.86 Hz, 6 H), 1.99 (m, 3 H), 1.75 (m, 2 H). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 170.1, 168.9, 164.7, 162.1, 160.1, 156.9, 149.3, 145.9, 138.9, 132.7, 131.2, 127.6, 118.4, 116.4, 114.7, 114.3, 112.4, 100.8, 83.7, 69.0, 68.4, 27.9, 24.8, 18.1, 11.2. HRMS (ESI): calcd for $C_{32}H_{25}N_3O_8Na$ ($[MNa]^+$): 602.1539, found: 602.1550. **mp**: decomp. at 192.9 – 194.2 °C.

4.5.6.3. Synthesis of Compound 4.10 and Compound 4.15 (Scheme 4.7)



4.10: R = Me, 28%
4.15: R = H, 37%

Compound 4.12: 4-ethynylbenzaldehyde was added to a solution of 3-aminophenol (1 g, 9.2 mmol, 2.0 eq.) in THF (20 mL) and then stirred for 1.5 h at room temperature. STAB (1.36 g, 6.44 mmol, 1.4 eq.) was added and the mixture was stirred for an additional 3 h before the reaction was quenched with sat. $NaHCO_3$ (60 mL) and extracted with ether (3 × 20 mL). After purifying compound **4.12** by flash chromatography using a 2:8 EtOAc/hexanes eluent, the product was obtained as a transparent viscous oil in 75% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.51 – 7.39 (m, 2H), 7.31 (d, *J* = 8.4 Hz, 2H), 7.01 (t, *J* = 8.0 Hz, 1H), 6.20 (dddd,

J = 10.4, 8.0, 2.3, 0.8 Hz, 2H), 6.08 (t, *J* = 2.3 Hz, 1H), 4.32 (s, 2H), 3.06 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 156.8, 149.6, 140.4, 132.6, 130.4, 127.4, 121.1, 106.2, 104.9, 99.8, 83.6, 48.1. HRMS (EI): calcd for $C_{15}H_{13}NO$: 223.09971, found: 223.09831.

Compound 4.13: A solution of compound **4.12** (460 mg, 2.06 mmol) dissolved in MeOH was cooled to 0 °C. After cooling, acetaldehyde (4.0 eq, 0.461 mL, 8.24 mmol) was added and the solution was stirred for 2 h. The solution was removed from ice before adding $NaBH_4$ (6.0 eq., 468 mg, 12.4 mmol) and stirring the mixture at room temperature overnight. Purification by flash chromatography by gradient elution (0-30% EtOAc in hexane) yielded compound **4.13** in 50% yield as a transparent viscous oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 7.43 (d, *J* = 8.23 Hz, 2 H), 7.18 (d, *J* = 7.84 Hz, 2 H), 7.02 (t, *J* = 8.28 Hz, 1 H), 6.25 (m, 1 H), 6.15 (m, 2 H), 4.62 (s, 1 H), 4.48 (s, 2 H), 3.43 (q, *J* = 7.09 Hz, 2 H), 3.04 (s, 1 H), 1.19 (t, *J* = 7.05 Hz, 3 H). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 156.7, 149.9, 140.1, 132.4, 130.2, 126.5, 120.5, 105.1, 103.3, 99.2, 83.6, 53.8, 45.4, 12.2. HRMS (ESI): calcd for $C_{17}H_{18}NO$ ($[MH]^+$): 252.1388, found: 252.1362.

Compound 4.11: The synthesis of compound **4.4** was adapted from a procedure developed by Takakura.[37] $POCl_3$ (0.464 mL, 4.77 mmol, 4.0 eq.) was added dropwise to DMF (1.06 mL, 13.7 mmol, 1..5 eq.) over ice and stirred for 5 min before compound **4.13** was added. The solution was stirred at room temperature for 1 h and then heated to 70 °C and stirred for an additional 1 h before cooling the solution to room temperature. The intermediate was hydrolyzed by adding H_2O (10 mL) to the reaction and stirring for 30 min. The reaction was then neutralized with NaOH and extracted with EtOAc (3 × 50 mL). After drying the combined organic layers with $MgSO_4$, compound **4.11** was purified by flash chromatography using a 2:8 EtOAc/hexanes elution system and was obtained as a white solid in 44% yield. ^{1}H

NMR (400 MHz, $CDCl_3$) δ 11.57 (s, 1H), 9.57 – 9.18 (m, 1H), 7.45 (d, *J* = 8.3 Hz, 2H), 7.26 (t, *J* = 4.4 Hz, 2H), 7.13 (d, *J* = 8.5 Hz, 2H), 6.26 (dd, *J* = 8.9, 2.4 Hz, 1H), 6.12 (d, *J* = 2.4 Hz, 1H), 4.59 (s, 2H), 3.52 (q, *J* = 7.1 Hz, 2H), 3.06 (s, 1H), 1.26 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 192.4, 164.3, 154.9, 138.0, 135.5, 132.6, 126.2, 121.2, 112.1, 104.8, 97.5, 83.3, 53.4, 45.8, 12.2. HRMS (EI): calcd for $C_{18}H_{17}NO_2$: 279.12593, found: 279.12498. **mp**: 97.1 – 97.8 °C.

Compound 4.14: Piperidine (0.02 eq., 1.3 µL 0.0130 mmol) and AcOH (0.02 eq., 0.75 µL, 0.0130 mmol) was added to a solution of compound **4.11** (184 mg, 0.658 mmol) and Meldrum's acid (1.0 eq., 95 mg, 0.658 mmol) in 99% EtOH (5 mL). The solution was stirred at room temperature for 20 min before it was heated to 78 °C and stirred for an additional 3 h. After cooling the mixture to room temperature and then to 0 °C, the product was purified by overnight crystallization and collected as a yellow-green solid in 48% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 12.30 (s, 1 H), 8.67 (s, 1 H), 7.46 (t, *J*=8.52 Hz, 3 H), 7.13 (d, *J*=8.33 Hz, 2 H), 6.72 (dd, *J*=9.01, 2.45 Hz, 1 H), 6.57 (d, *J*=2.25 Hz, 1 H), 4.67 (s, 2 H), 3.62 (q, *J*=7.15 Hz, 2 H), 3.08 (s, 1 H), 1.32 (t, *J*=7.10 Hz, 3 H). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 165.3, 164.1, 157.8, 154.3, 150.4, 136.8, 132.8, 132.0, 126.1, 121.7, 111.3, 109.2, 106.8, 97.7, 83.0, 53.9, 46.4, 12.1. HRMS (ESI): calcd for $C_{21}H_{17}NO_4Na$ ($[MNa]^+$): 370.1055, found: 370.1065. **mp**: 207.3 – 208.8 °C.

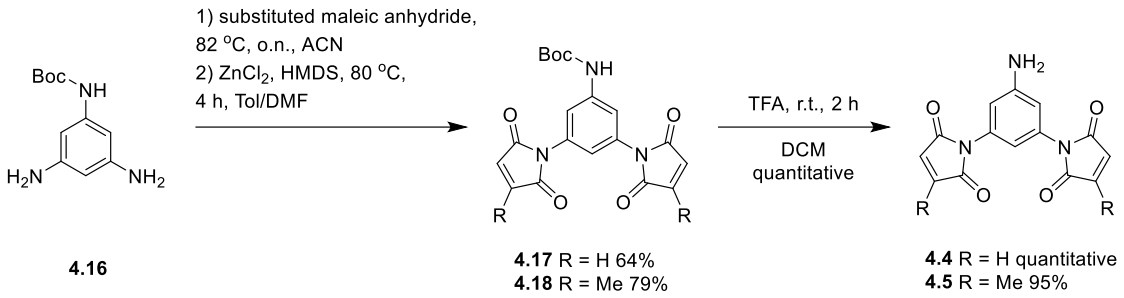
Compound 4.10 and 4.15: $POCl_3$ (3.0 eq., 40.4 µL, 432 µmol) was added to a solution of compound **4.5** (50 mg, 144 µmol) dissolved in DCE (1.4 mL). The solution was then heated to 84 °C and stirred for 4 h to form the acyl chloride. Rotary evaporation was used to remove the solvent and the residue was dissolved in acetonitrile (1.4 mL). A separate solution containing **4.5** or **4.4** (to make **4.10** or **4.15** respectively) (1.2 eq., 49 mg, 173 µmol) and

DIPEA (2.0 eq, 50 µL, 288 µmol) in acetonitrile (1.4 mL) was then added to the solution of the acyl chloride. The combined reaction mixture was stirred at room temperature overnight. After using the rotary evaporator to remove the solvent the product was purified from crude by flash chromatography using a gradient elution system of 0 – 50% EtOAc in hexanes elution system.

Compound 4.10: yellow solid in 28% yield. ^{1}H NMR (600 MHz, $CDCl_3$) δ 11.00 (s, 1H), 8.77 (s, 1H), 7.83 (d, *J* = 1.9 Hz, 2H), 7.51 – 7.41 (m, 3H), 7.23 (t, *J* = 1.8 Hz, 1H), 7.14 (d, *J* = 8.3 Hz, 2H), 6.67 (dd, *J* = 8.9, 2.5 Hz, 1H), 6.56 (d, *J* = 2.3 Hz, 1H), 6.48 (q, *J* = 1.7 Hz, 2H), 4.64 (s, 2H), 3.59 (q, *J* = 7.1 Hz, 2H), 3.07 (s, 1H), 2.17 (d, *J* = 1.8 Hz, 6H), 1.30 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 170.2, 169.1, 162.9, 161.2, 157.7, 153.6, 149.0, 146.0, 139.4, 137.5, 132.9, 132.7, 131.6, 127.7, 126.3, 121.6, 118.3, 116.5, 110.8, 109.4, 97.7, 83.3, 77.7, 54.0, 46.4, 12.3, 11.3. HRMS (ESI): $C_{37}H_{28}N_4O_7$ calcd for: 663.1856, found: 663.1855 ($[MNa]^+$). **mp**: decomp. at 330.0 – 331.2 °C.

Compound 4.15: yellow solid in 37% yield. ^{1}H NMR (600 MHz, DMSO) δ 10.93 (s, 1H), 8.74 (s, 1H), 7.79 (d, *J* = 1.8 Hz, 2H), 7.74 (d, *J* = 9.1 Hz, 1H), 7.46 (d, *J* = 8.3 Hz, 2H), 7.23 (d, *J* = 8.4 Hz, 2H), 7.21 (s, 4H), 7.11 (t, *J* = 1.8 Hz, 1H), 6.86 (d, *J* = 11.2 Hz, 1H), 6.70 (s, 1H), 4.78 (s, 2H), 4.16 (s, 1H), 3.64 (q, *J* = 7.0 Hz, 2H), 1.19 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (150 MHz, DMSO) δ 169.7, 162.0, 161.0, 157.1, 153.4, 148.3, 138.8, 134.8, 132.4, 132.0, 131.9, 127.7, 126.8, 126.0, 120.4, 120.2, 117.1, 110.9, 109.8, 108.4, 96.8, 83.3, 80.8, 52.8, 45.7, 29.8, 12.2. HRMS (ESI): calcd for $C_{35}H_{24}N_4O_7Na$ ($[MNa]^+$): 635.1543, found: 635.1537. **mp**: decomp. at 184.8-184.9 °C.

4.5.6.4. Synthesis of Compound 4.4 (Scheme 4.8)



Scheme 4.8. Synthesis of compound **4.4** and compound **4.5**.

Compound 4.16: A protocol developed by Chen[22] was used to synthesize **4.16**. Synthetic procedure and characterization of compound **4.16** can be found in Chapter 2, section 2.4.1.4, scheme 2.10, compound **2.25**.

Compound 4.17: Maleic anhydride (3.0 eq, 1.05 g, 10.7 mmol) was added to a solution of compound **4.16** (800 mg, 3.58 mmol) dissolved in MeCN (35 mL) after which the solution was heated to reflux and stirred overnight. Next, the rotary evaporator was used to remove the solvent and the residue was triturated with diethyl ether (3 × 35 mL). The triturated residue was dissolved in DMF (25 mL) and diluted with toluene (100 mL). A solution of HMDS (4.5 eq, 3.36 mL, 16.1 mmol) in toluene (20 mL) was then added dropwise, followed by the addition of $ZnCl_2$ (3.0 eq, 1.46 g, 10.7 mmol). The mixture was heated to 80 °C and stirred for 4 h after which it was diluted with EtOAc (250 mL). The solution was washed with 0.1 M HCl (3 × 80 mL), sat. $NaHCO_3$ (80 mL) and brine (80 mL) and then dried with $MgSO_4$ and evaporated on the rotary evaporator. The crude was purified by flash chromatography using a 4:6 EtOAc/hexanes elution system to obtain compound **4.17** as a yellow solid in 64% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 (d, *J* = 1.6 Hz, 2H), 7.15 (t, J = 1.8 Hz, 1H), 6.84 (s, 4H), 6.70 (s, 1H), 1.50 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.1, 152.4, 139.9, 134.4,

132.5, 117.1, 114.7, 81.4, 28.4. HRMS (ESI): calcd for $C_{19}H_{17}N_3O_6Na$ ([MNa]$^+$): 406.1015, found: 406.1012.

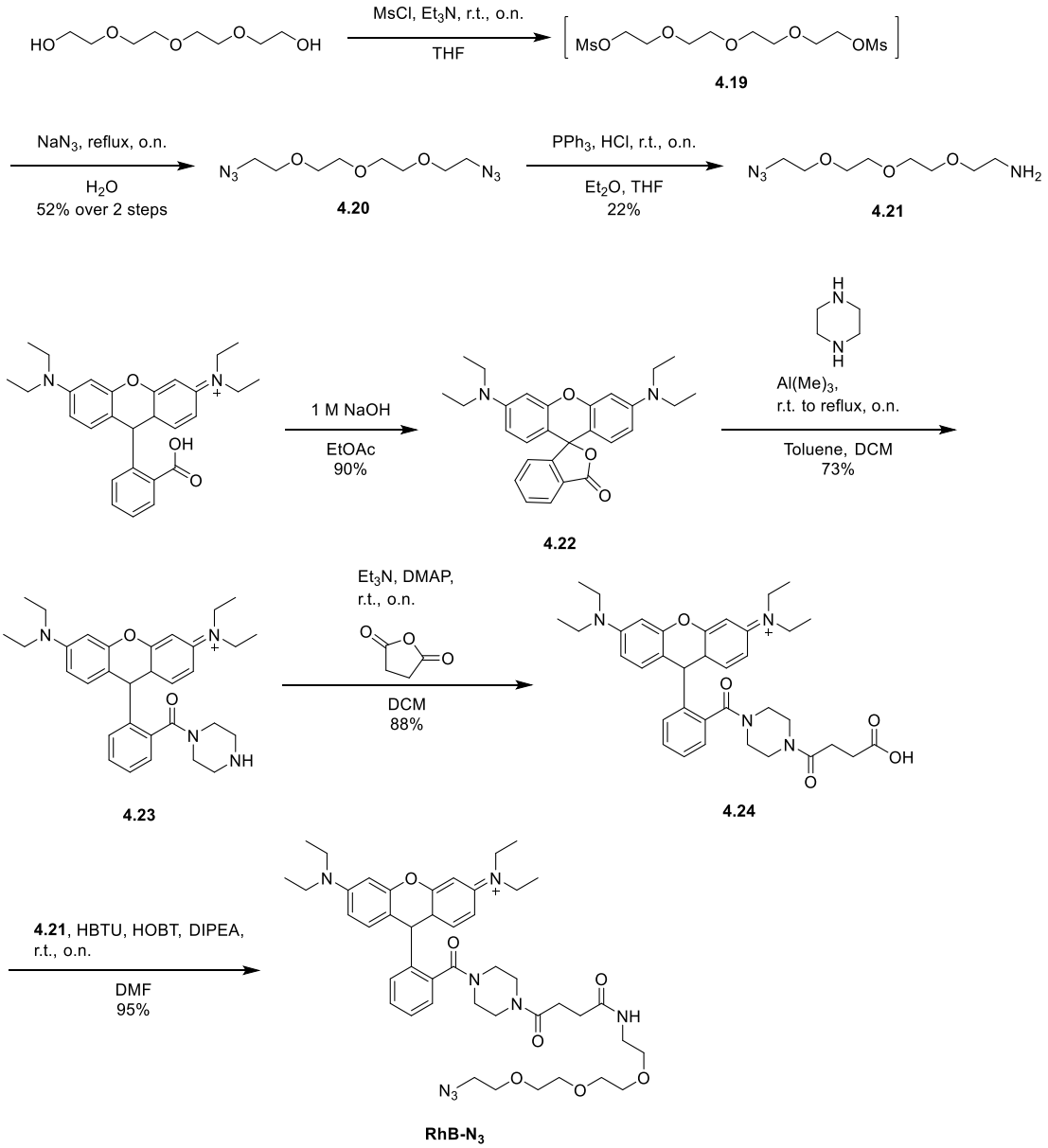
Compound **4.18**: Compound **4.18** was synthesized using the same protocol as compound **4.17** but with citraconic anhydride instead of maleic anhydride. It has been synthesized before by Chen,[22] and spectra are consistent with what was reported. Starting with 1.00 g (4.50 mmol) of compound **4.16**, 1.45 g (3.53 mmol) of compound **4.18** was obtained as a yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 (d, *J* = 1.6 Hz, 2H), 7.16 (t, *J* = 1.8 Hz, 1H), 6.64 (s, 1H), 6.46 (q, *J* = 1.8 Hz, 2H), 2.15 (d, *J* = 1.8 Hz, 6H), 1.50 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.2, 169.1, 152.4, 146.0, 139.7, 132.9, 127.6, 116.9, 114.3, 81.3, 28.4, 11.3.

Compound **4.4**: TFA (57 eq., 10 mL, 130 mmol) was added dropwise to a solution of compound **4.17** (877 mg, 2.29 mmol) dissolved in DCM (10 mL). The solution was stirred at room temperature for 2 h before the liquid reagents and solvents were removed using a rotary evaporator. The crude was triturated with DCM (5 × 50 mL) and compound **4.4** was obtained as a white solid in quantitative yield. ^{1}H NMR (400 MHz, DMSO) δ 7.16 (s, 4H), 6.66 (d, *J* = 1.5 Hz, 2H), 6.57 (s, 1H). ^{13}C NMR (100 MHz, DMSO) δ 169.7, 147.0, 134.6, 132.4, 113.7, 112.6. HRMS (ESI): calcd for $C_{14}H_9N_3O_4Na$ ([MNa]$^+$): 306.0491, found: 306.0499. **mp**: 175.3 – 177.0 °C.

Compound **4.5**: Compound **4.5** was synthesized using the same protocol as compound **4.17** but with compound **4.18** instead of compound **4.17**. It has been synthesized before by Chen,[22] and spectra are consistent with what was reported. From 1.45 g (3.53 mmol) of compound **4.18**, 1.83 g (5.89 mmol) of compound **4.18** was obtained as a yellow solid in quantitative yield. ^{1}H NMR (400 MHz, DMSO) δ 8.14 (bs, 2H), 6.78 (q, *J* = 1.7 Hz, 2H), 6.72

(d, *J* = 1.7 Hz, 2H), 6.64 (s, 1H), 2.06 (d, *J* = 1.8 Hz, 6H). ^{13}C NMR (100 MHz, DMSO) δ 170.3, 169.4, 158.5, 145.7, 132.8, 127.5, 114.3, 113.0, 10.8. HRMS (ESI): calcd for $C_{16}H_{13}N_3O_4Na$ ([MNa]$^+$): 334.0804, found: 334.0818.

4.5.6.5. Synthesis of Rhodamine B Azide (RhB-N_3) (Scheme 4.9)



Scheme 4.9. Synthesis of rhodamine B azide (RhB-N_3).

Compound **4.19** and **4.20**: Compound **4.19** and **4.20** was synthesized according to a protocol by Li, et al.[38] Obtained spectra are consistent with what was reported. Triethylamine (6.1 mL, 43.5 mmol) was added to a solution of tetramethylene glycol (5.0 mL, 28.95 mmol) dissolved in THF (30 mL). MsCl (4.96 mL, 63.73 mmol) was added to this solution over ice, after which the solution was brought to room temperature and stirred overnight. Once compound **2.19** was formed, NaN_3 (5.64 g, 86.9 mmol) and H_2O (0.25 mL) was added to the solution, which was then brought to reflux and stirred overnight. Next, THF was removed by evaporation and the residue was diluted with H_2O (30 mL) and extracted with DCM (3 × 30 mL). The combined organic phases were dried with $MgSO_4$ and filtered, and then DCM was removed by evaporation. The crude was purified by flash chromatography using a gradient elution system from 10-35% EtOAc in hexanes, which afforded compound **2.20** as a yellow liquid in 52% yield over two steps (3.65 g, 15.0 mmol). ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.69 – 3.65 (m, 12H), 3.43 – 3.33 (m, 4H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 70.8, 70.2, 50.8.

Compound **4.21**: Compound **4.21** was synthesized according to a protocol established by Li, et al.[38] Compound **4.20** was dissolved in a mixture of Et_2O (24 mL), THF (6 mL) and 1 M HCl (24 mL). A solution of PPh_3 in Et_2O (3.93 g, 15.0 mmol, 0.5 M solution) was added to the solution formed above. The combined mixture of both solutions was vigorously stirred at room temperature overnight. Following this, the reaction mixture was extracted with 4 M HCl (2 × 25 mL). The combined aqueous extractions were extracted with EtOAc (3 × 50 mL) and then basified to pH 14 with NaOH pellets. The basified aqueous layer was then further extracted with DCM (3 × 50 mL). The combined EtOAc and DCM organic phases were then dried with $MgSO_4$ and filtered before removing the solvent by evaporation. The crude was

obtained in 22% yield (714 mg, 3.27 mmol) and used in the next step without further purification.

Compound **4.22**: Compound **4.22** was synthesized according to a protocol by Nguyen, et al.[39] Obtained spectra are consistent with what was reported. Rhodamine B (3.30 g, 6.90 mmol) was partitioned between 1 M NaOH (30 mL) and EtOAc (30 mL). The organic phase was set aside and the aqueous phase was further extracted with EtOAc (2 × 50 mL) before combining all the organic phases and washing them with 1 M NaOH (100 mL) and brine (100 mL). The organic phase was dried over $MgSO_4$, filtered, and then the solvent was removed by evaporation. Compound **4.22** was obtained as a pink foam in 90% yield (2.75 g, 6.21 mmol). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.03 – 7.97 (m, 1H), 7.60 (dtd, *J* = 23.6, 7.4, 1.1 Hz, 2H), 7.24 – 7.18 (m, 1H), 6.57 (d, *J* = 8.9 Hz, 2H), 6.44 (d, *J* = 2.6 Hz, 2H), 6.34 (dd, *J* = 8.9, 2.6 Hz, 2H), 3.36 (q, *J* = 7.1 Hz, 8H), 1.17 (t, *J* = 7.1 Hz, 12H).

Compound **4.23**: Compound **4.22** was synthesized according to a protocol by Nguyen, et al.[39] Obtained spectra are consistent with what was reported. A solution of $Al(Me)_3$ in toluene (6.20 mL, 12.4 mmol, 2.00 M solution) was added dropwise to a stirring solution of piperazine (2.14 g, 24.8 mmol) in DCM (10 mL). This solution was allowed to stir for 1 h at room temperature before a solution of compound **4.22** (2.75 g, 6.21 mmol) in DCM (6 mL) was added dropwise. The combined solutions were brought to reflux and stirred for 24 h. After this, 0.1 M HCl was added dropwise until no more gas was evolved. The solution was then filtered and rinsed with DCM (100 mL) and a 4:1 DCM/MeOH (100) mixture before being concentrated by evaporation. The residue was partitioned between 5% $NaHCO_3$ (50 mL) and EtOAc (50 mL) and the aqueous layer was washed with EtOAc (3 × 50 mL) before it was saturated with NaCl and acidified with 1 M HCl. The aqueous layer was then extracted with a 2:1 mixture of

isopropylalcohol and DCM (4 × 50 mL) until only a faint pink colour persisted. The organic 2:1 isopropylalcohol/DCM layers were combined and dried with $MgSO_4$. After filtering the organic layer, the solvent was removed by evaporation and the residue was dissolved in minimal MeOH. Compound **4.23** was obtained by the dropwise addition of a large amount of Et_2O and collected as a dark purple precipitate in 73% (2.30 g, 4.51 mmol) yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 – 7.63 (m, 2H), 7.61 – 7.52 (m, 1H), 7.40 – 7.29 (m, 1H), 7.16 (d, *J* = 9.5 Hz, 2H), 6.97 (d, *J* = 8.7 Hz, 2H), 6.73 (d, *J* = 2.0 Hz, 2H), 3.91 (d, *J* = 24.6 Hz, 4H), 3.70 – 3.43 (m, 8H), 3.10 (d, *J* = 27.1 Hz, 4H), 1.32 (t, *J* = 7.1 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.3, 157.8, 155.8, 134.7, 131.7, 131.1, 130.5, 130.4, 130.3, 127.9, 114.6, 113.7, 96.5, 46.4, 12.8. HRMS (ESI): calcd for $C_{32}H_{39}N_4O_2$ ($[M]^+$): 511.3073, found: 511.3057.

Compound **4.24**: Compound **4.22** was synthesized according to a protocol by Nguyen, et al.[39] Obtained spectra are consistent with what was reported. Compound **4.23** (1.02 g, 2.00 mmol), succinic anhydride (200 mg, 2.00 mmol), and DMAP (244 mg, 2.00 mmol) were dissolved in DCM (10 mL) followed by the addition of Et_3N (0.362 mL, 2.60 mmol). The solution was stirred at room temperature overnight and then diluted with EtOAc (20 mL) and 1 M K_2CO_3 (20 mL). The aqueous layer was washed with EtOAc (3 × 20 mL) and then saturated with NaCl. Next, the NaCl saturated aqueous layer was extracted with a 2:1 isopropylalcohol/DCM mixture (3 × 40 mL) and the combined isopropylalcohol/DCM layers were dried with $MgSO_4$ and filtered. After removing the solvent by evaporation, compound **4.24** was obtained as a shiny black solid in 88% yield (1.08 g, 1.76 mmol). ^{1}H NMR (400 MHz, MeOD) δ 7.76 (dt, *J* = 9.6, 3.8 Hz, 2H), 7.72 – 7.66 (m, 1H), 7.53 – 7.47 (m, 1H), 7.27 (d, *J* = 9.5 Hz, 2H), 7.07 (d, *J* = 9.2 Hz, 2H), 6.96 (d, *J* = 2.4 Hz, 2H), 3.68 (q, *J* = 7.1 Hz, 8H), 3.40 (d, *J* = 8.9 Hz,

8H), 2.58 (t, *J* = 6.6 Hz, 2H), 2.43 (t, *J* = 6.6 Hz, 2H), 1.29 (t, *J* = 7.1 Hz, 12H). HRMS (ESI): calcd for $C_{36}H_{43}N_4O5$ ($[M]^+$): 611.3233, found: 611.3218.

RhB-N_3: Compound **4.22** was synthesized according to a protocol by Yang, et al.[30] Only a MS was provided, which was consistent with the MS obtained herein. Compound **4.24** (500 mg, 0.818 mmol), HBTU (620 mg, 1.64 mmol), and HOBT (221 mg, 1.64 mmol) were dissolved in DMF (5 mL) and a solution of compound **2.21** (714 mg, 3.27 mmol) in DMF (5 mL) was added to this solution. The reaction mixture was stirred at room temperature overnight and then diluted with a solution of HCl in sat. NaCl (5% %v/v, 100 mL). The aqueous solution was extracted with DCM (3 × 50 mL) and the combined organic layers were then washed with sat. $NaHCO_3$ (2 × 50 mL), 5% HCl in sat. NaCl (2 × 50 mL), and brine (100 mL). The organic layer was dried with $MgSO_4$, filtered, and evaporated to afford RhB-N_3 as a purple solid in 95% yield (631 mg, 0.777 mmol). HRMS (ESI): calcd for $C_{44}H_{59}N_8O_7$ ($[M]^+$): 811.4507, found: 811.4457.

4.6. Chapter 4 References

(1) Tsao, K. K.; Lee, A. C.; Racine, K. É.; Keillor, J. W. (2020) Site - Specific Fluorogenic Protein Labelling Agent for Bioconjugation. *Biomol* 10, 369.

(2) Kim, J. S.; Sirois, A. R.; Vazquez Cegla, A. J.; Jumai'An, E.; Murata, N.; Buck, M. E.; Moore, S. J. (2019) Protein-Polymer Conjugates Synthesized Using Water-Soluble Azlactone-Functionalized Polymers Enable Receptor-Specific Cellular Uptake toward Targeted Drug Delivery. *Bioconjug. Chem.* 30, 1220–1231.

(3) Klok, H. A. (2009) Peptide/Protein-Synthetic Polymer Conjugates: Quo Vadis. *Macromolecules* 42, 7990–8000.

(4) Duncan, R. (2003) The Dawning Era of Polymer Therapeutics. *Nat. Rev. Drug Discov.* 2, 347–360.

(5) Pelegri-Oday, E. M.; Lin, E. W.; Maynard, H. D. (2014) Therapeutic Protein-Polymer Conjugates: Advancing beyond Pegylation. *J. Am. Chem. Soc.* 136, 14323–14332.

(6) Ducry, L.; Stump, B. (2010) Antibody-Drug Conjugates: Linking Cytotoxic Payloads to Monoclonal Antibodies. *Bioconjug. Chem.* 21, 5–13.

(7) Agarwal, P.; Bertozzi, C. R. (2015) Site-Specific Antibody-Drug Conjugates: The Nexus of Bioorthogonal Chemistry, Protein Engineering, and Drug Development. *Bioconjug. Chem.* 26, 176–192.

(8) Chauhan, D. P.; Saha, T.; Lahiri, M.; Talukdar, P. (2014) BODIPY Based "click on" Fluorogenic Dyes: Application in Live Cell Imaging. *Tetrahedron Lett.* 55, 244–247.

(9) Pasut, G.; Veronese, F. M. (2009) PEG Conjugates in Clinical Development or Use as

Anticancer Agents: An Overview. *Adv. Drug Deliv. Rev.* 61, 1177–1188.

(10) Sievers, E. L.; Senter, P. D. (2013) Antibody-Drug Conjugates in Cancer Therapy. *Annu. Rev. Med.* 64, 15–29.

(11) Canalle, L. A.; Löwik, D. W. P. M.; Van Hest, J. C. M. (2010) Polypeptide-Polymer Bioconjugates. *Chem. Soc. Rev.* 39, 329–353.

(12) Deonarain, M. P.; Yahioglu, G.; Stamati, I.; Marklew, J. (2015) Emerging Formats for Next-Generation Antibody Drug Conjugates. *Expert Opin. Drug Discov.* 10, 463–481.

(13) Chudasama, V.; Maruani, A.; Caddick, S. (2016) Recent Advances in the Construction of Antibody-Drug Conjugates. *Nat. Chem.* 8, 114–119.

(14) Saxon, E.; Bertozzi, C. R. (2000) Cell Surface Engineering by a Modified Staudinger Reaction. *Science* 287, 2007–2010.

(15) Mashayekh, K.; Shiri, P. (2019) An Overview of Recent Advances in the Applications of Click Chemistry in the Synthesis of Bioconjugates with Anticancer Activities. *ChemistrySelect* 4, 13459–13478.

(16) Veronese, F. M. (2001) Peptide and Protein PEGylation: A Review of Problems and Solutions. *Biomaterials* 22, 405–417.

(17) Yamada, K.; Ito, Y. (2019) Recent Chemical Approaches for Site-Specific Conjugation of Native Antibodies: Technologies toward Next-Generation Antibody–Drug Conjugates. *ChemBioChem* 20, 2729–2737.

(18) Liu, G.; Shi, G.; Sheng, H.; Jiang, Y.; Liang, H.; Liu, S. (2017) Doubly Caged Linker for AND-Type Fluorogenic Construction of Protein/Antibody Bioconjugates and In

Situ Quantification. *Angew. Chemie Int. Ed.* 56, 8686–8691.

(19) Robin, M. P.; Wilson, P.; Mabire, A. B.; Kiviaho, J. K.; Raymond, J. E.; Haddleton, D. M.; O'Reilly, R. K. (2013) Conjugation-Induced Fluorescent Labeling of Proteins and Polymers Using Dithiomaleimides. *J. Am. Chem. Soc.* 135, 2875–2878.

(20) Dirks, A. T. J.; Cornelissen, J. J. L. M.; Nolte, R. J. M. (2009) Monitoring Protein-Polymer Conjugation by a Fluorogenic Cu (I)-Catalyzed Azide- Alkyne 1, 3-Dipolar Cycloaddition. *Bioconjug. Chem.* 20, 1129–1138.

(21) Liu, G.; Hu, J.; Liu, S. (2018) Emerging Applications of Fluorogenic and Non-Fluorogenic Bifunctional Linkers. *Chem. Eur.* 22, 565–572.

(22) Chen, Y.; Clouthier, C. M.; Tsao, K.; Strmiskova, M.; Lachance, H.; Keillor, J. W. (2014) Coumarin-Based Fluorogenic Probes for No-Wash Protein Labeling. *Angew. Chem. Int. Ed. Engl.* 53, 13785–13788.

(23) Chen, Y.; Tsao, K.; De Francesco, É.; Keillor, J. W. (2015) Ring Substituent Effects on the Thiol Addition and Hydrolysis Reactions of N-Arylmaleimides. *J. Org. Chem.* 80, 12182–12192.

(24) Strmiskova, M.; Tsao, K.; Keillor, J. W. (2018) Rational Design of a Highly Reactive Dicysteine Peptide Tag for Fluorogenic Protein Labelling. *Org. Biomol. Chem.* 16, 6332–6340.

(25) Jewett, J. C.; Bertozzi, C. R. (2011) Synthesis of a Fluorogenic Cyclooctyne Activated by Cu-Free Click Chemistry. *Org. Lett.* 13, 5937–5939.

(26) Key, J. A.; Koh, S.; Timerghazin, Q. K.; Brown, A.; Cairo, C. W. (2009) Photophysical

Characterization of Triazole-Substituted Coumarin Fluorophores. *Dye. Pigment.* 82, 196–203.

(27) Grimm, J. B.; English, B. P.; Chen, J.; Slaughter, J. P.; Zhang, Z.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; et al. (2015) A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy. *Nat. Methods* 12, 244–250.

(28) Kwan, D. H.; Ernst, S.; Kötzler, M. P.; Withers, S. G. (2015) Chemoenzymatic Synthesis of a Type 2 Blood Group A Tetrasaccharide and Development of High-Throughput Assays Enables a Platform for Screening Blood Group Antigen-Cleaving Enzymes. *Glycobiology* 25, 806–811.

(29) Zhang, X.; Chen, F.; Petrella, A.; Chacón-Huete, F.; Covone, J.; Tsai, T. W.; Yu, C. C.; Forgione, P.; Kwan, D. H. (2019) A High-Throughput Glycosyltransferase Inhibition Assay for Identifying Molecules Targeting Fucosylation in Cancer Cell-Surface Modification. *ACS Chem. Biol.* 14, 715–724.

(30) Yang, P. Y.; Liu, K.; Ngai, M. H.; Lear, M. J.; Wenk, M. R.; Yao, S. Q. (2010) Activity-Based Proteome Profiling of Potential Cellular Targets of Orlistat - An FDA-Approved Drug with Anti-Tumor Activities. *J. Am. Chem. Soc.* 132, 656–666.

(31) Yuan, L.; Lin, W.; Zheng, K.; Zhu, S. (2013) FRET-Based Small-Molecule Fluorescent Probes: Rational Design and Bioimaging Applications. *Acc. Chem. Res.* 46, 1462–1473.

(32) Chen, Y.; Tsao, K.; Keillor, J. W. (2014) Fluorogenic Protein Labelling: A Review of Photophysical Quench Mechanisms and Principles of Fluorogen Design. *Can. J. Chem.*

398, 1–10.

(33) Xiao-Fan Zhang, Tao Zhang, Shi-Li Shen, J.-Y. M. and B.-X. Z. (2015) A Ratiometric Lysosomal PH Probe Based on the Coumarin–Rhodamine FRET System. *RCS Adv.* 5, 49115–49121.

(34) Nasheri, N.; Joyce, M.; Rouleau, Y.; Yang, P.; Yao, S.; Tyrrell, D. L.; Pezacki, J. P. (2013) Modulation of Fatty Acid Synthase Enzyme Activity and Expression during Hepatitis C Virus Replication. *Chem. Biol.* 20, 570–582.

(35) Abdizadeh, T.; Kalani, M. R.; Abnous, K.; Tayarani-Najaran, Z.; Khashyarmanesh, B. Z.; Abdizadeh, R.; Ghodsi, R.; Hadizadeh, F. (2017) Design, Synthesis and Biological Evaluation of Novel Coumarin-Based Benzamides as Potent Histone Deacetylase Inhibitors and Anticancer Agents. *Eur. J. Med. Chem.* 132, 42–62.

(36) Jiang, X. R.; Wang, P.; Smith, C. L.; Zhu, B. T. (2013) Synthesis of Novel Estrogen Receptor Antagonists Using Metal-Catalyzed Coupling Reactions and Characterization of Their Biological Activity. *J. Med. Chem.* 56, 2779–2790.

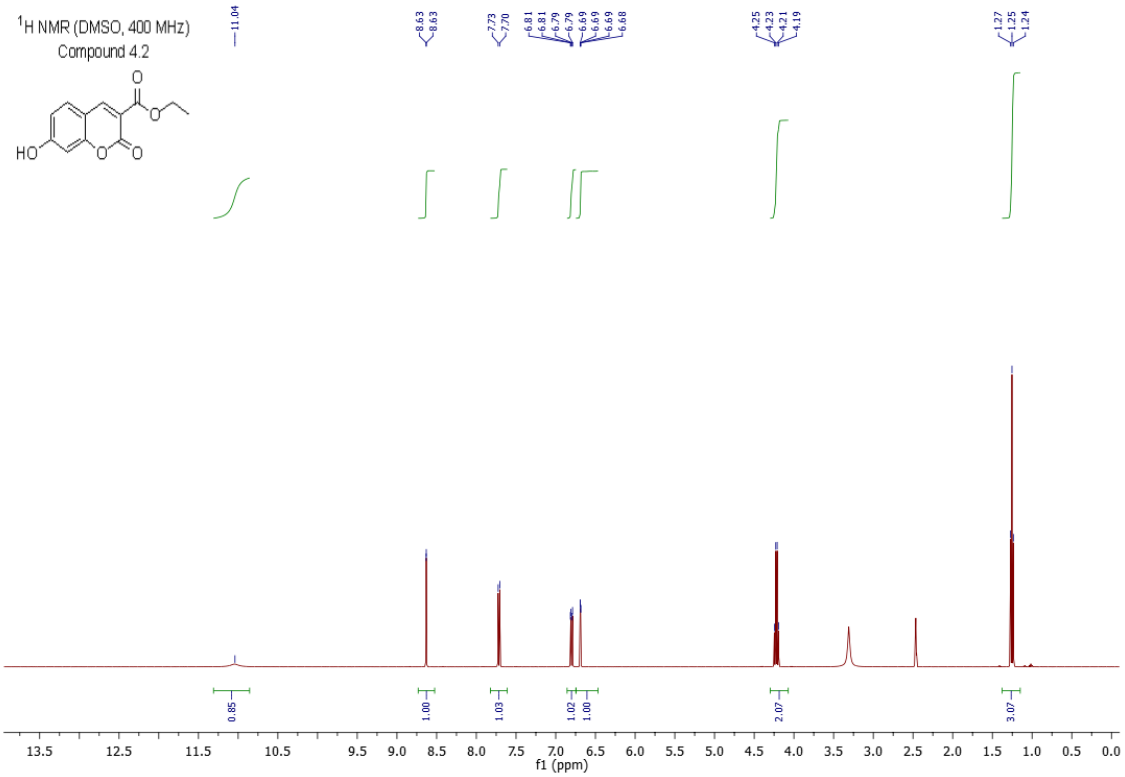
(37) Takakura, H.; Sasakura, K.; Ueno, T.; Urano, Y.; Terai, T.; Hanaoka, K.; Tsuboi, T.; Nagano, T. (2010) Development of Luciferin Analogues Bearing an Amino Group and Their Application as BRET Donors. *Chem-Asian J.* 5, 2053–2061.

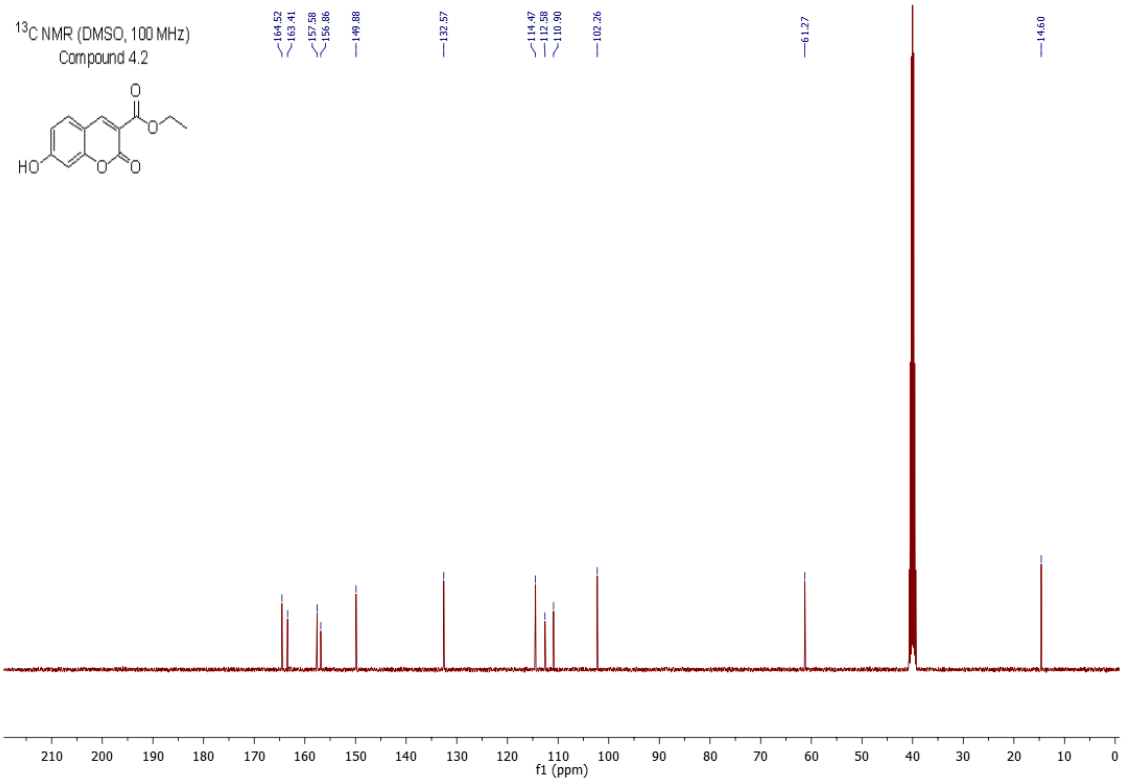
(38) Li, K.; Chen, Y.; Li, S.; Nguyen, H. G.; Niu, Z.; You, S.; Mello, C. M.; Lu, X.; Wang, Q. (2010) Chemical Modification of M13 Bacteriophage and Its Application in Cancer Cell Imaging. *Bioconjug. Chem.* 21, 1369–1377.

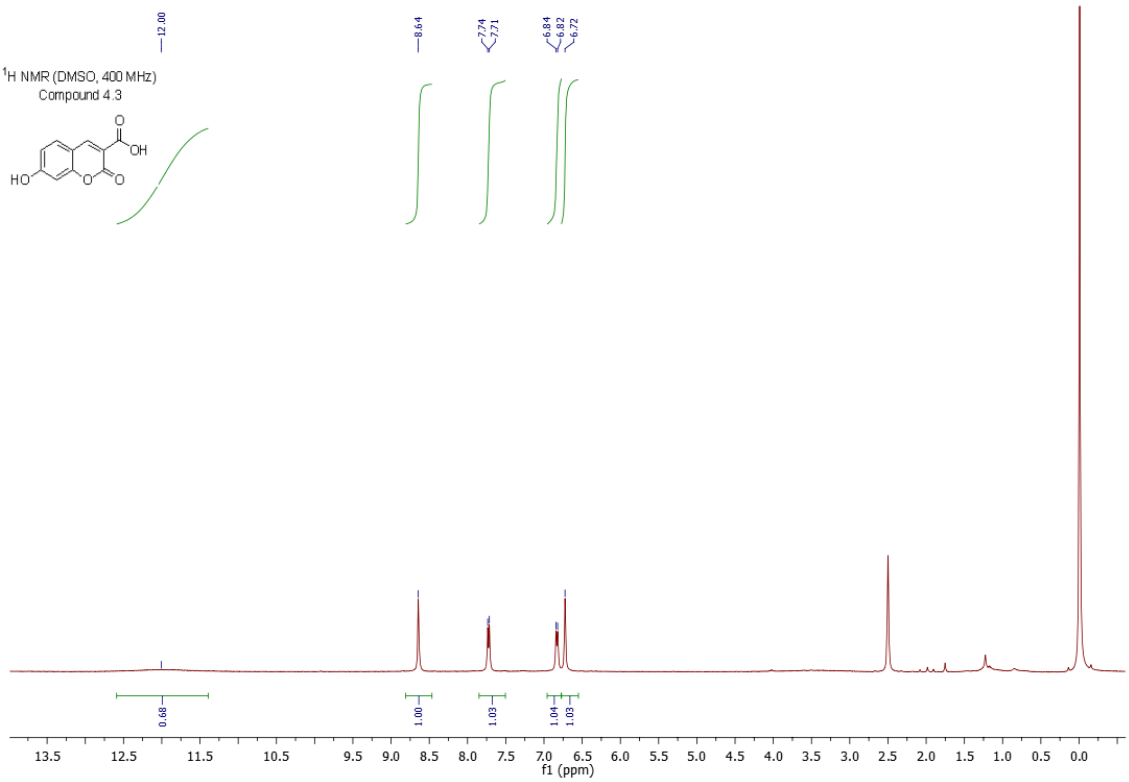
(39) Nguyen, T.; Francis, M. B. (2003) Practical Synthetic Route to Functionalized

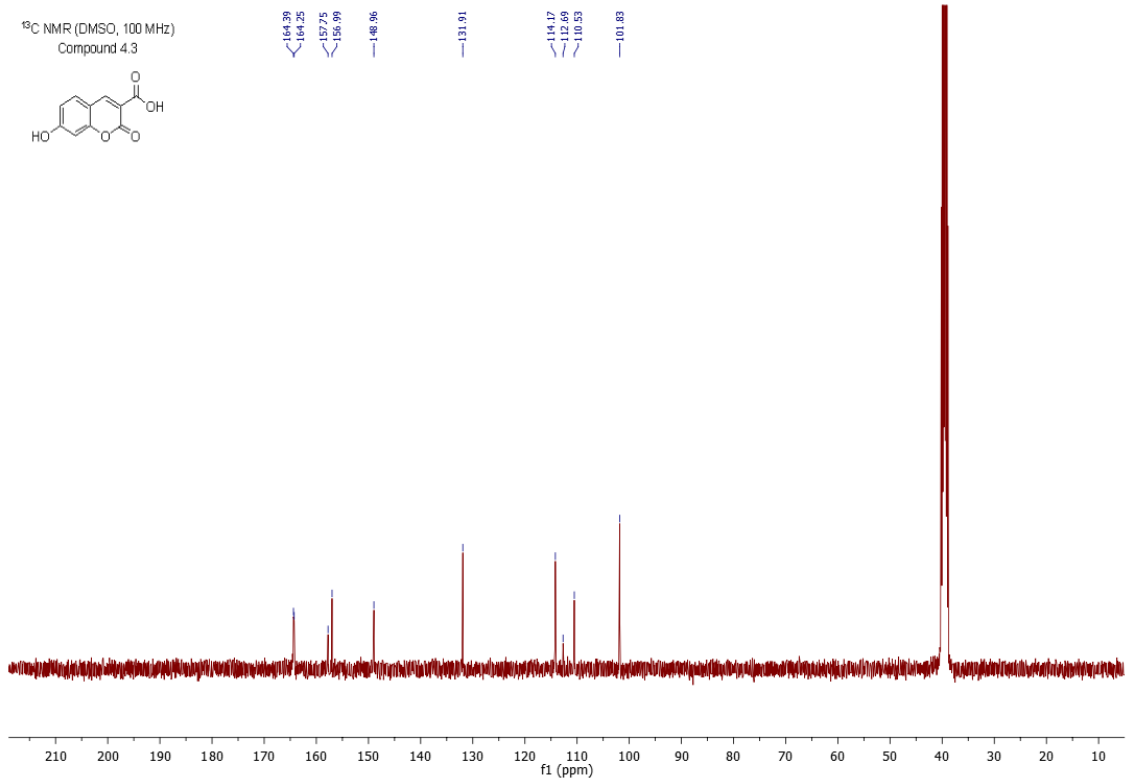
Rhodamine Dyes. *Org. Lett.* 5, 3245–3248.

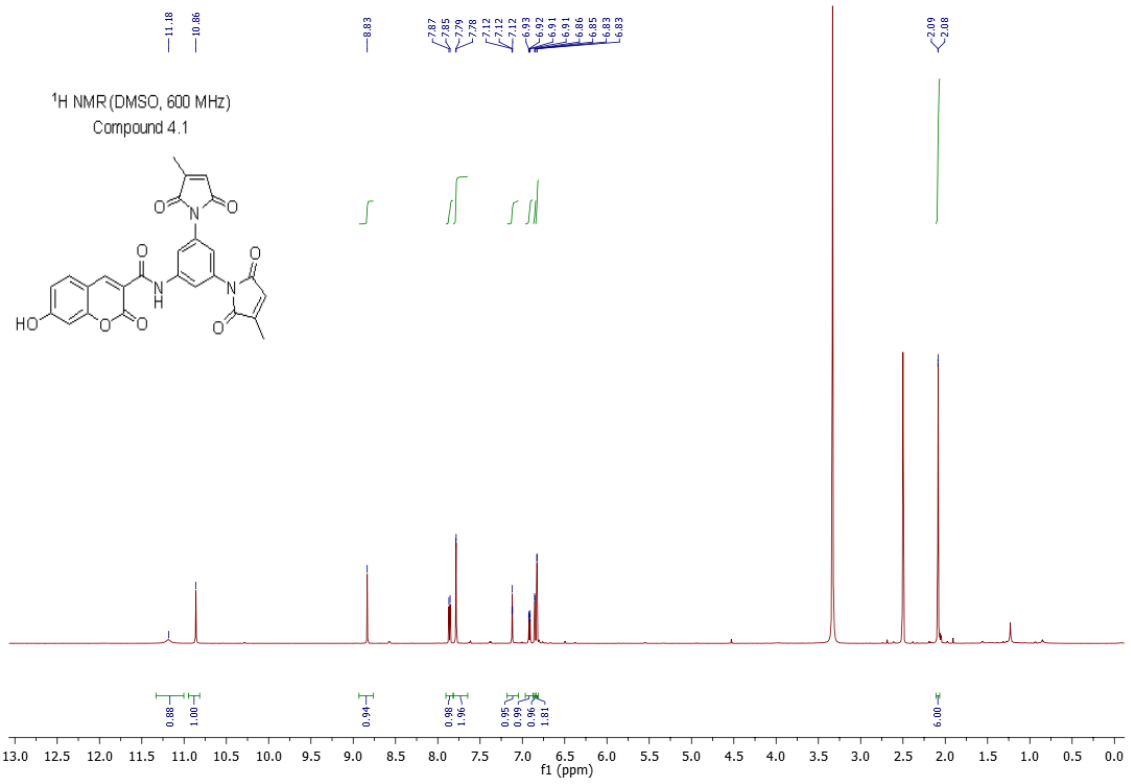
4.7. Chapter 4 Appendices

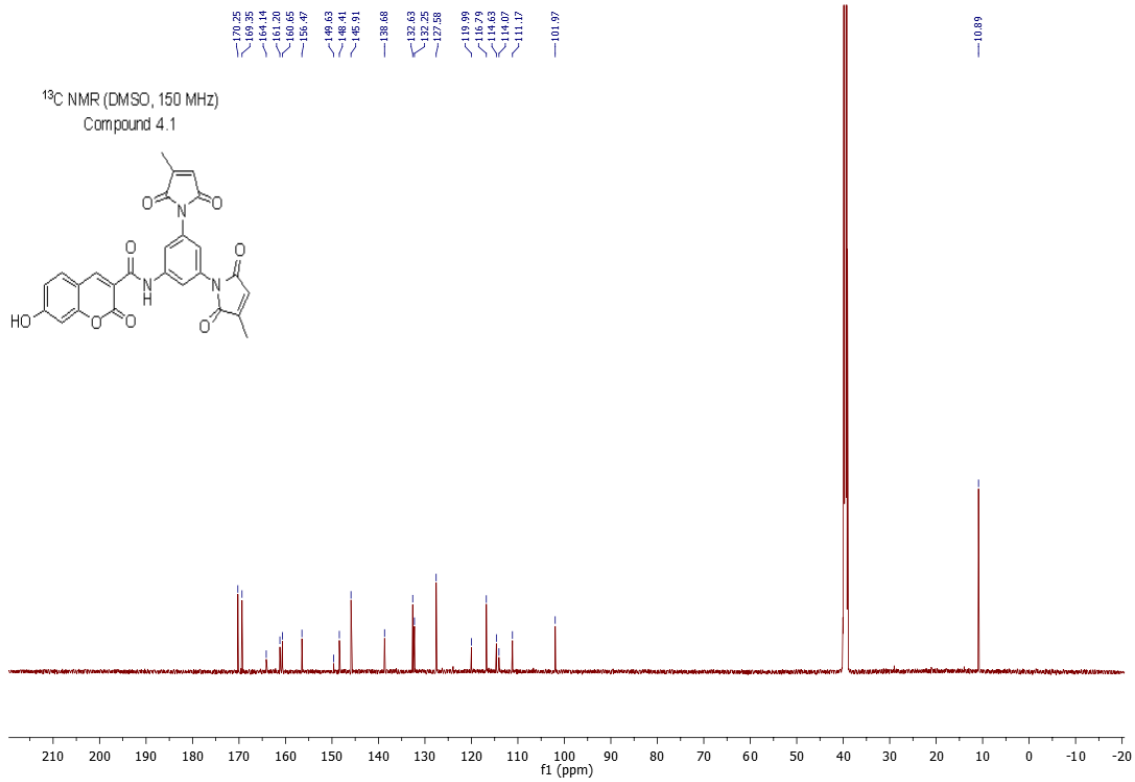


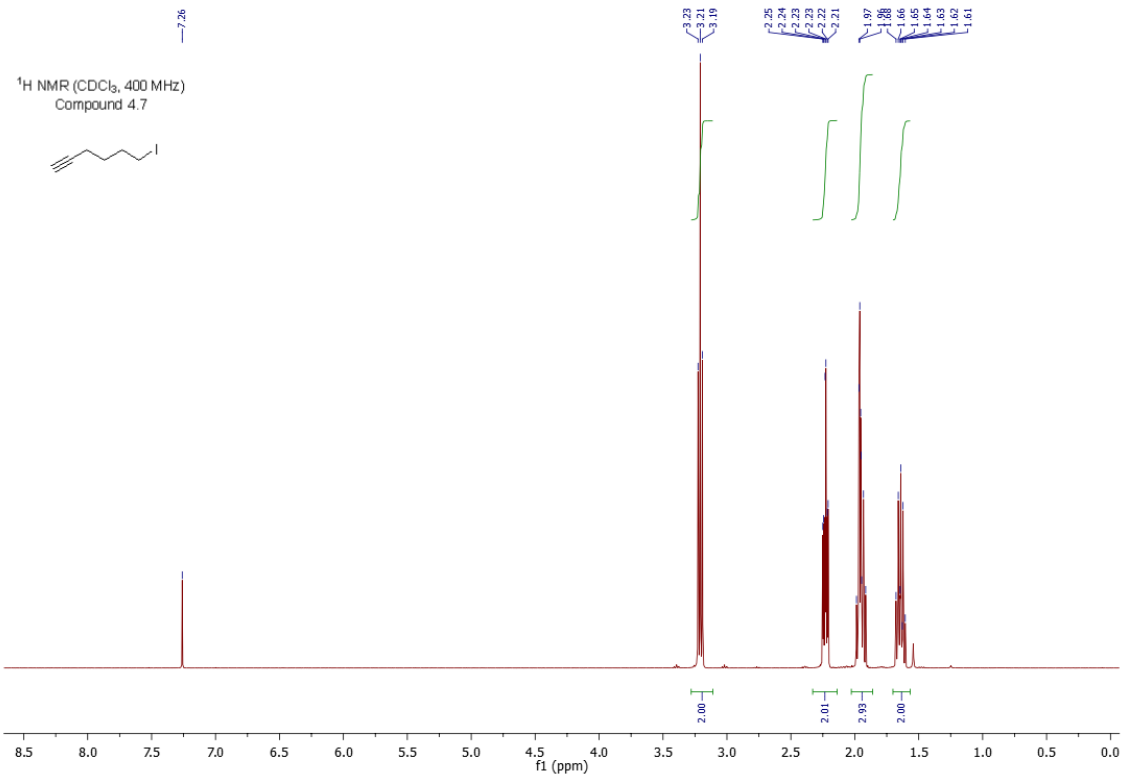


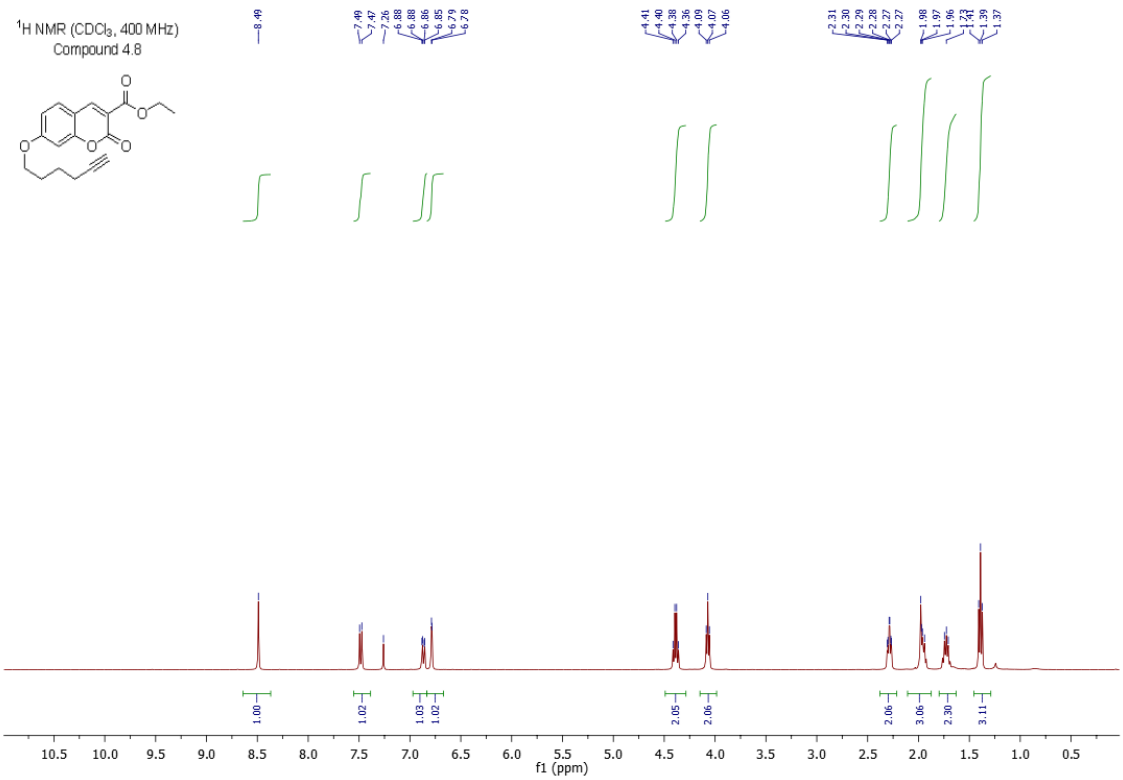

^{1}H NMR (DMSO, 400 MHz)
Compound 4.3
12.00
8.64
7.74
7.71
6.84
6.82
6.72
0.68
1.00
1.03
1.04
1.03
f1 (ppm)

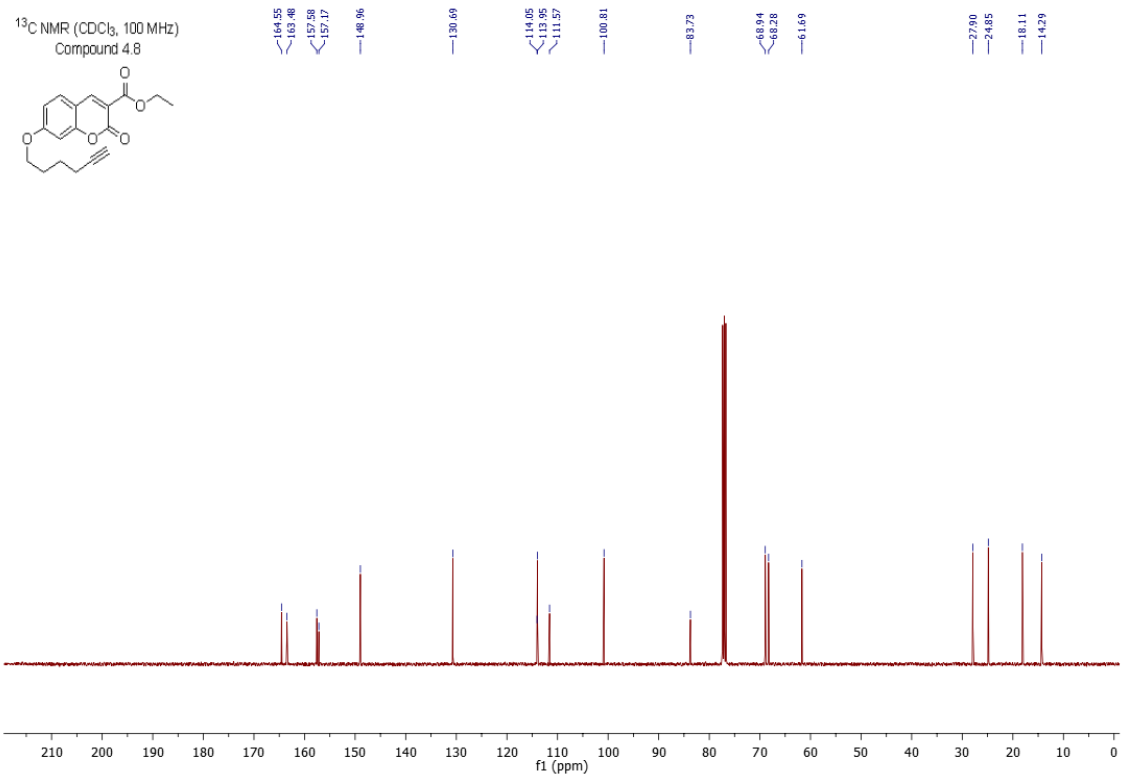

^{13}C NMR (DMSO, 100 MHz)
Compound 4.3
164.39
164.25
157.75
156.99
148.96
131.91
114.17
112.69
110.53
101.83
f1 (ppm)

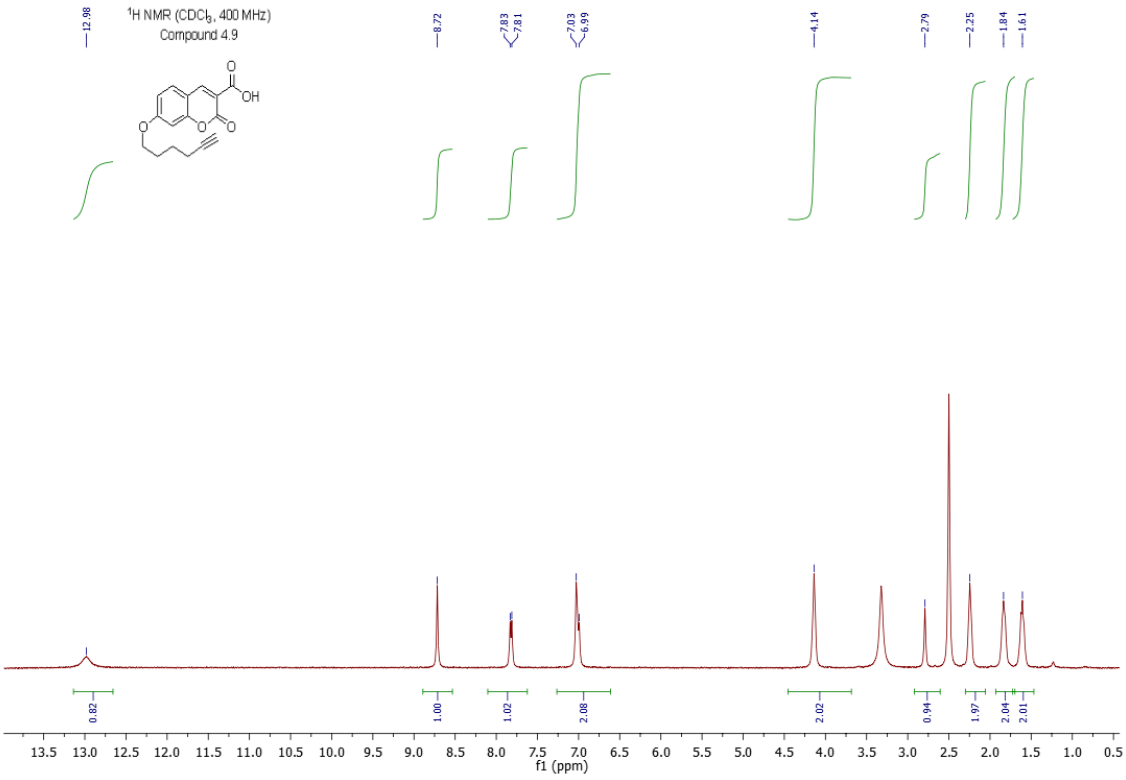

11.18
10.86
8.83
7.87
7.85
7.79
7.78
7.12
7.12
7.12
6.93
6.92
6.91
6.91
6.86
6.85
6.83
6.83
2.09
2.08
1H NMR (DMSO, 600 MHz)
Compound 4.1
0.88
1.00
0.94
0.98
1.96
0.95
0.99
0.96
1.81
6.00
f1 (ppm)

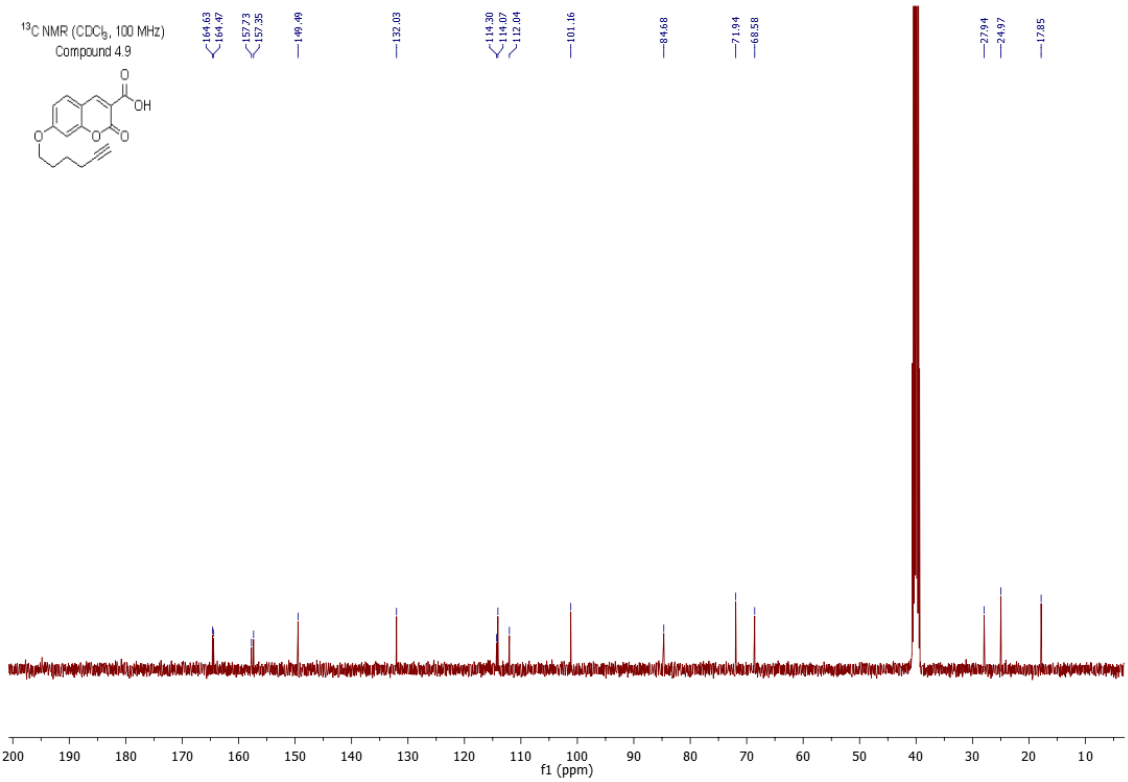

170.25
169.35
164.14
161.20
160.65
156.47
149.63
148.41
145.91
138.68
132.63
132.25
127.58
119.99
116.79
114.63
114.07
111.17
101.97
10.89
13C NMR (DMSO, 150 MHz)
Compound 4.1
f1 (ppm)

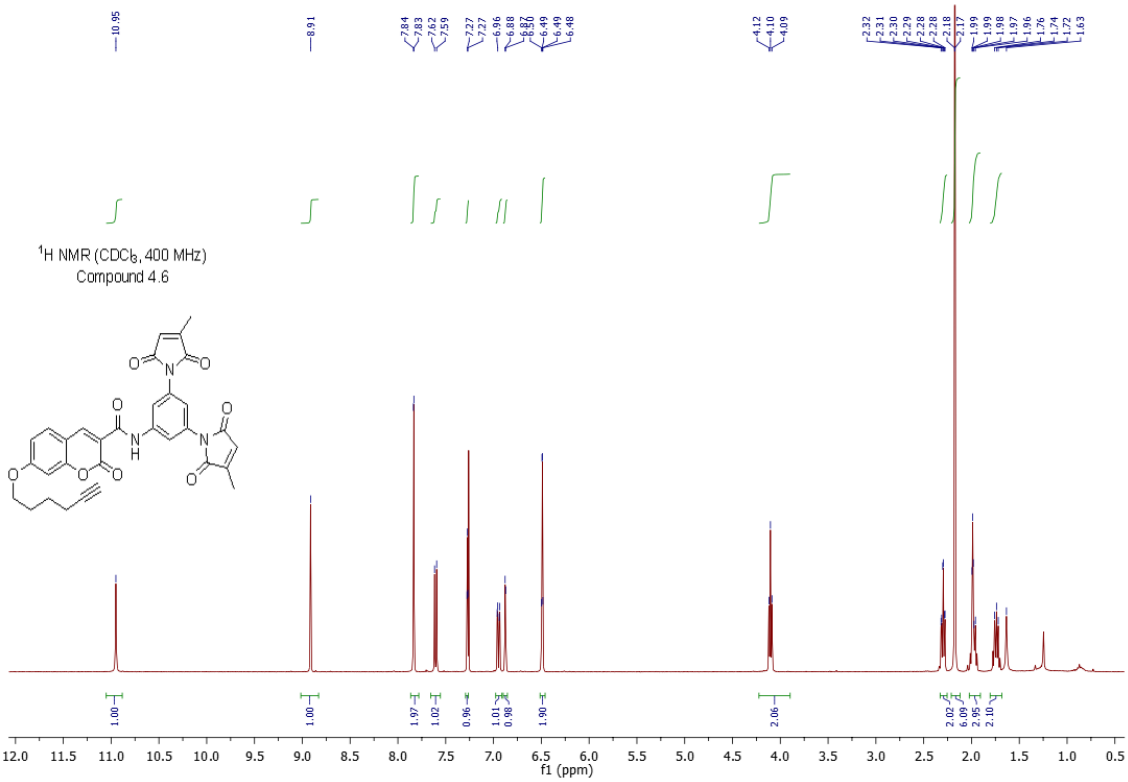

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 4.7
I
f1 (ppm)

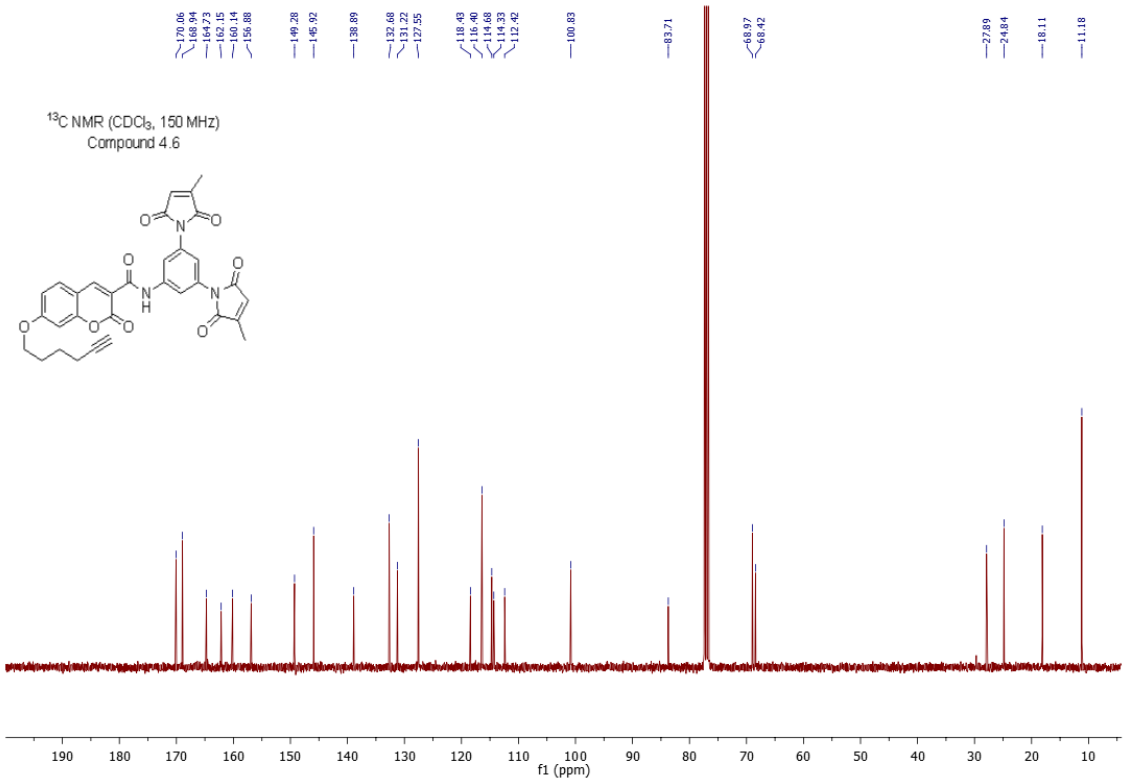

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 4.8
O
O
O
O
O
f1 (ppm)

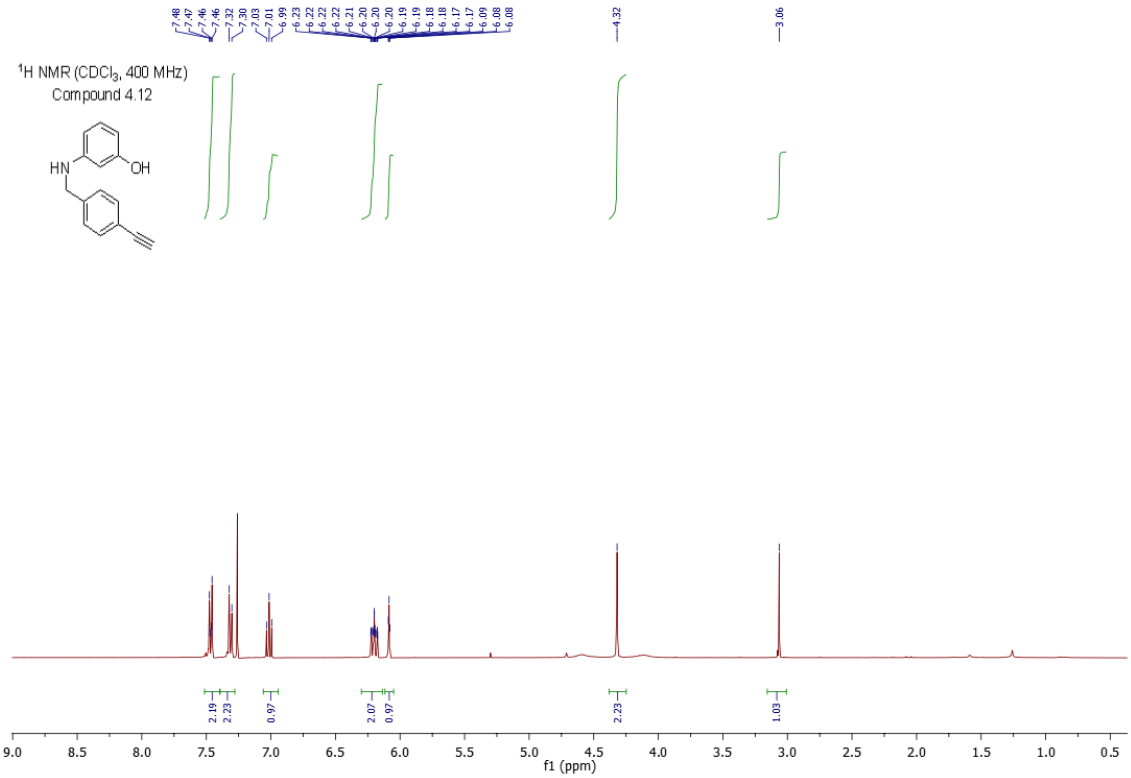

^{13}C NMR ($CDCl_3$, 100 MHz)
Compound 4.8
164.55
163.40
157.58
157.17
148.96
130.69
114.05
113.95
111.57
100.81
83.73
68.94
68.28
61.69
27.90
24.85
18.11
14.29
f1 (ppm)

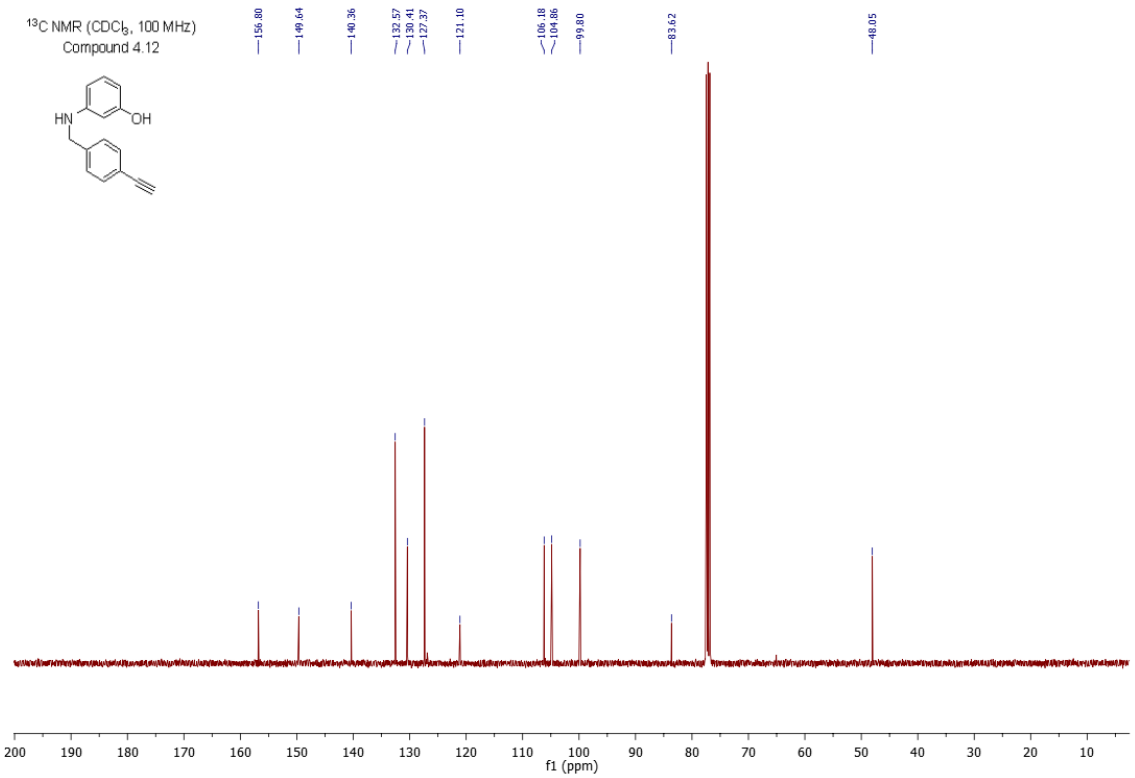

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 4.9
12.98
8.72
7.83
7.81
7.03
6.99
4.14
2.79
2.25
1.84
1.61
0.82
1.00
1.02
2.08
2.02
0.94
1.97
2.04
2.01
f1 (ppm)

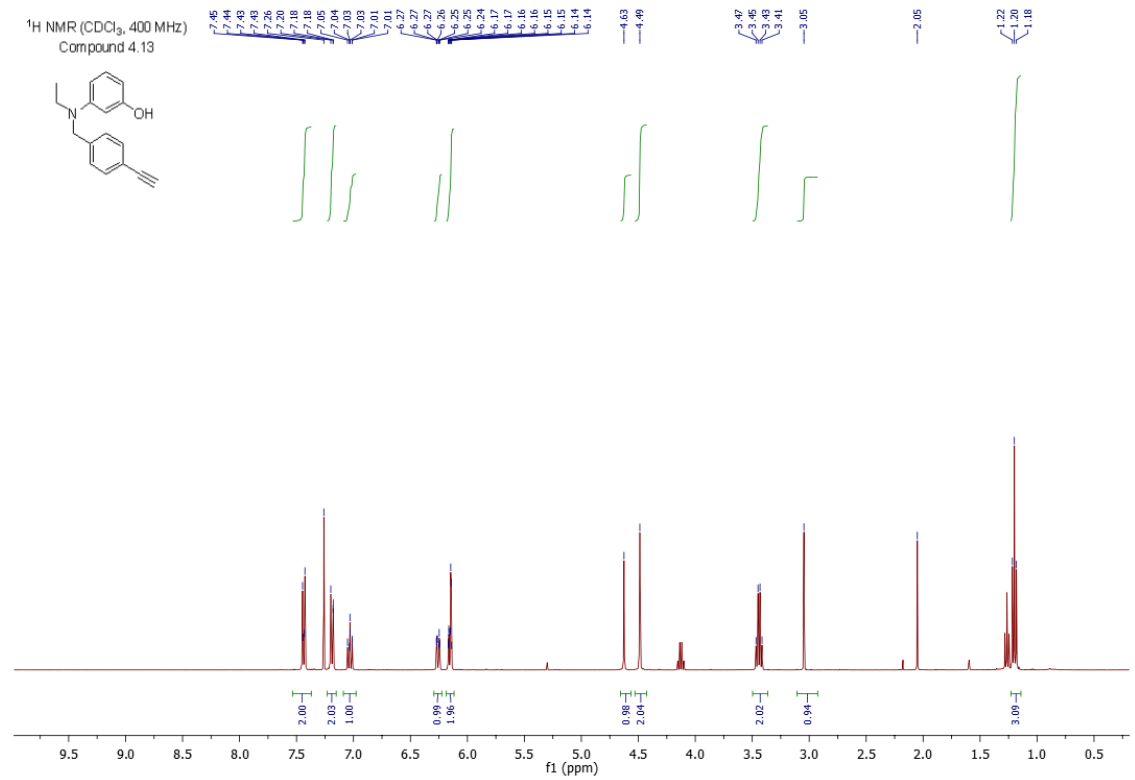

13C NMR (CDCl3, 100 MHz)
Compound 4.9
164.63
164.47
157.73
157.35
149.49
132.03
114.30
114.07
112.04
101.16
84.68
71.94
68.58
27.94
24.97
17.85
f1 (ppm)

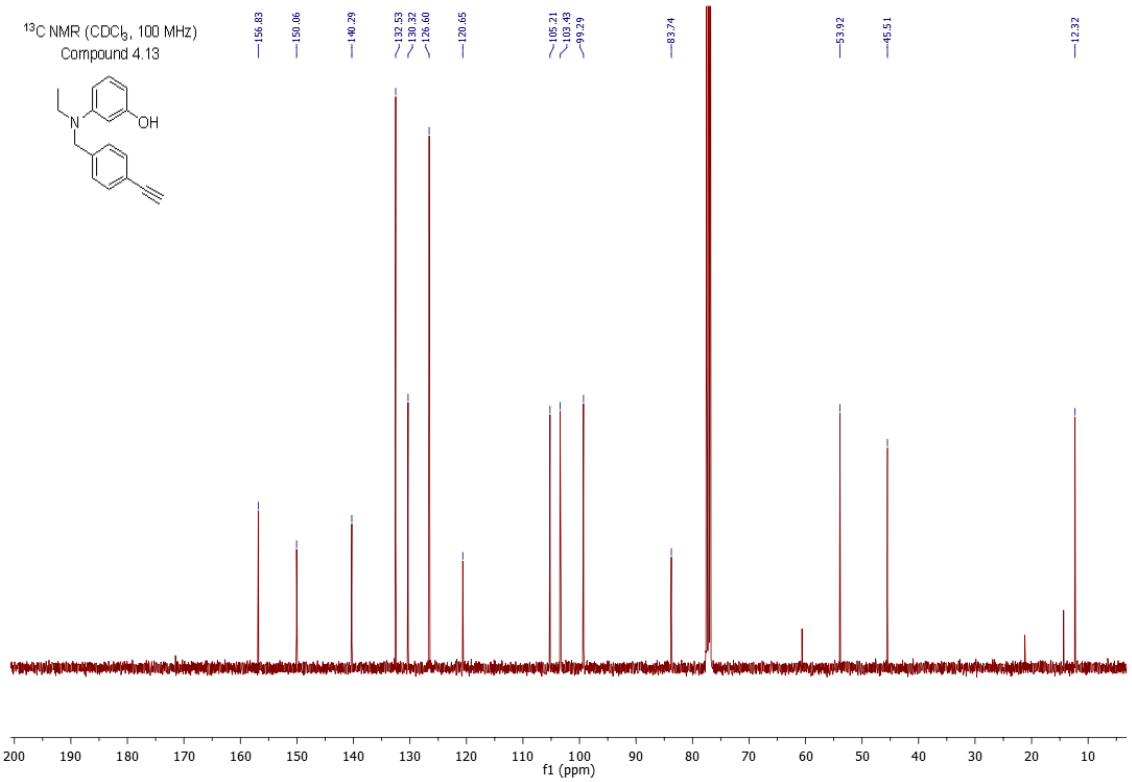

1H NMR (CDCl3, 400 MHz)
Compound 4.6
10.95
8.91
7.84
7.83
7.62
7.59
7.27
7.27
6.96
6.88
6.87
6.49
6.49
6.48
4.12
4.10
4.09
2.32
2.31
2.30
2.29
2.28
2.28
2.18
2.17
1.99
1.99
1.98
1.97
1.96
1.76
1.74
1.72
1.63
1.00
1.00
1.97
1.02
0.96
1.01
0.98
1.90
2.06
2.02
6.09
2.95
2.10
f1 (ppm)

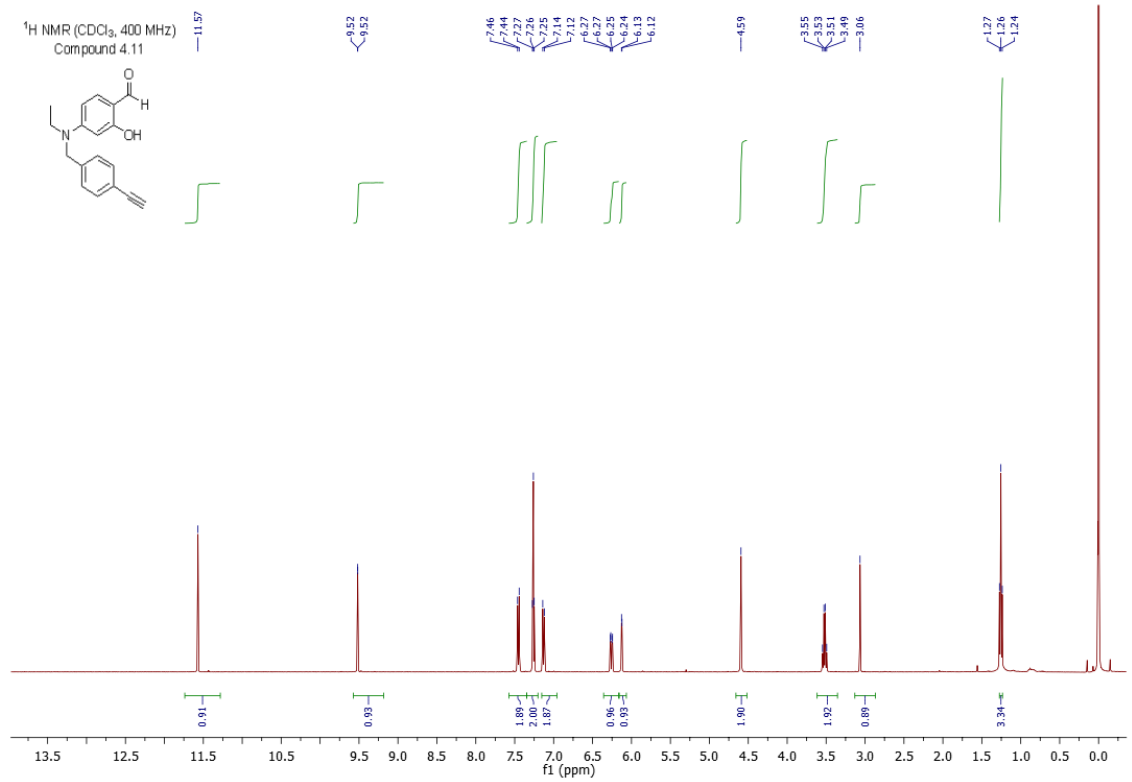

170.06
168.94
164.73
162.15
160.14
156.88
149.28
145.92
138.89
132.68
131.22
127.55
118.43
116.40
114.68
114.33
112.42
100.83
83.71
68.97
68.42
27.89
24.84
18.11
11.18
13C NMR (CDCl3, 150 MHz)
Compound 4.6
f1 (ppm)

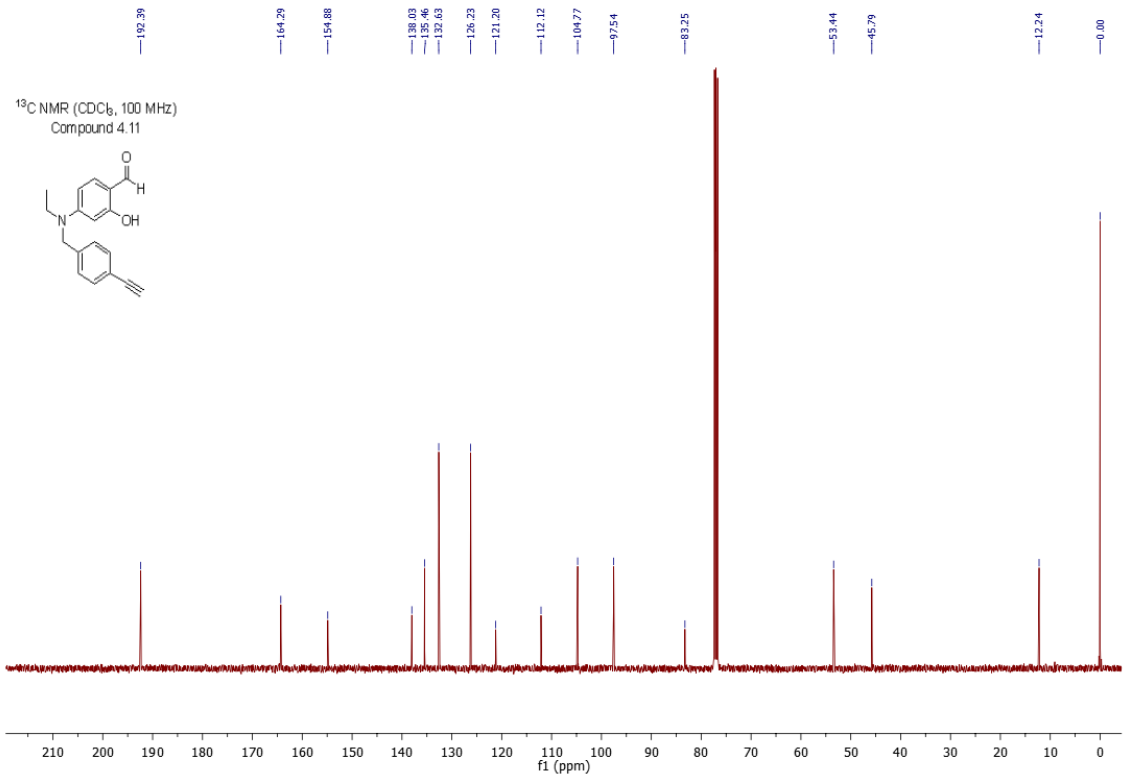

7.48
7.47
7.46
7.46
7.32
7.30
7.03
7.01
6.99
6.23
6.22
6.22
6.21
6.20
6.20
6.20
6.19
6.19
6.18
6.18
6.17
6.17
6.09
6.08
6.08
4.32
3.06
1H NMR (CDCl3, 400 MHz)
Compound 4.12
HN
OH
2.19
2.23
0.97
2.07
0.97
2.23
1.03
f1 (ppm)

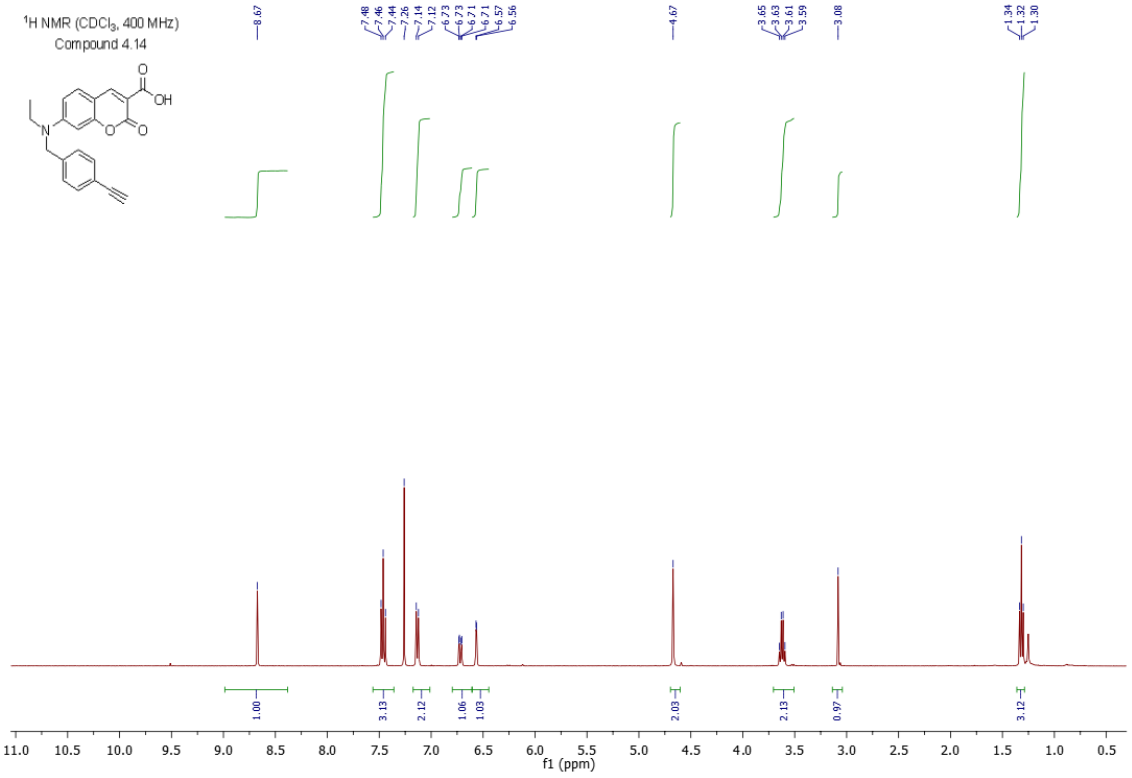

13C NMR (CDCl3, 100 MHz)
Compound 4.12
HN
OH
156.80
149.64
140.36
132.57
130.41
127.37
121.10
106.18
104.86
99.80
83.62
48.05
f1 (ppm)

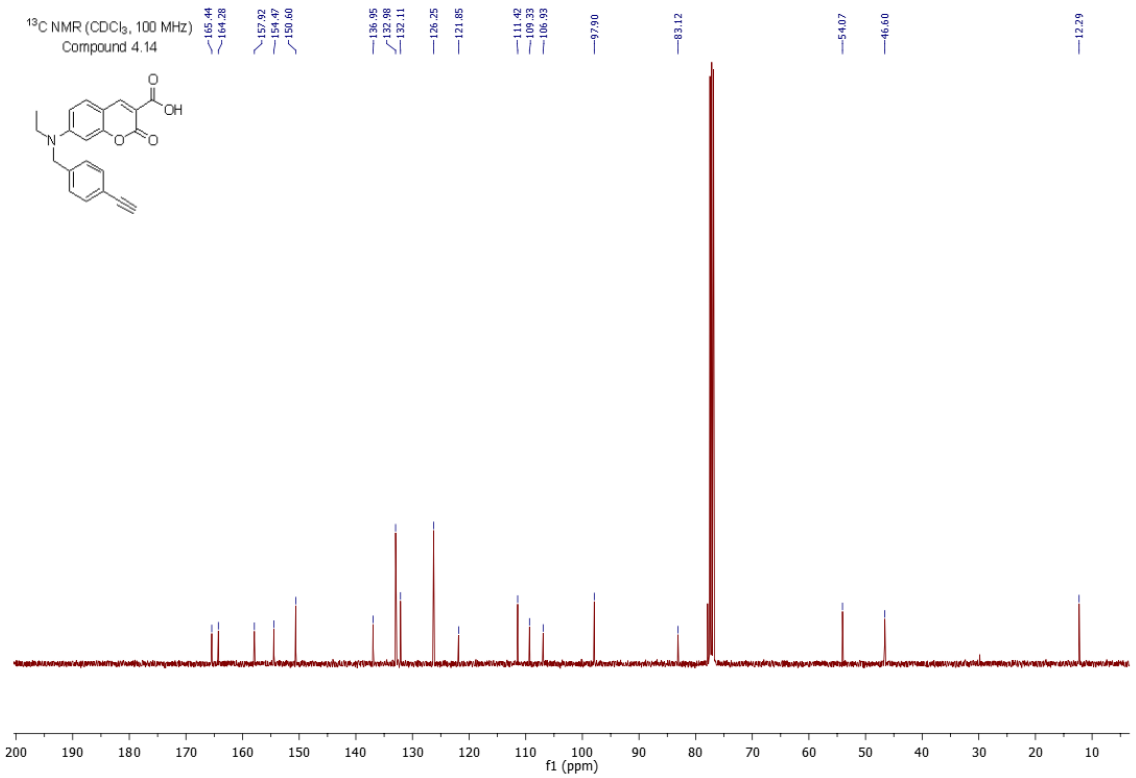

1H NMR (CDCl3, 400 MHz)
Compound 4.13
N
OH
4.63
4.49
3.47
3.46
3.43
3.41
3.05
2.05
1.22
1.20
1.18
2.00
2.03
1.00
0.99
1.96
0.98
2.04
2.02
0.94
3.09
f1 (ppm)

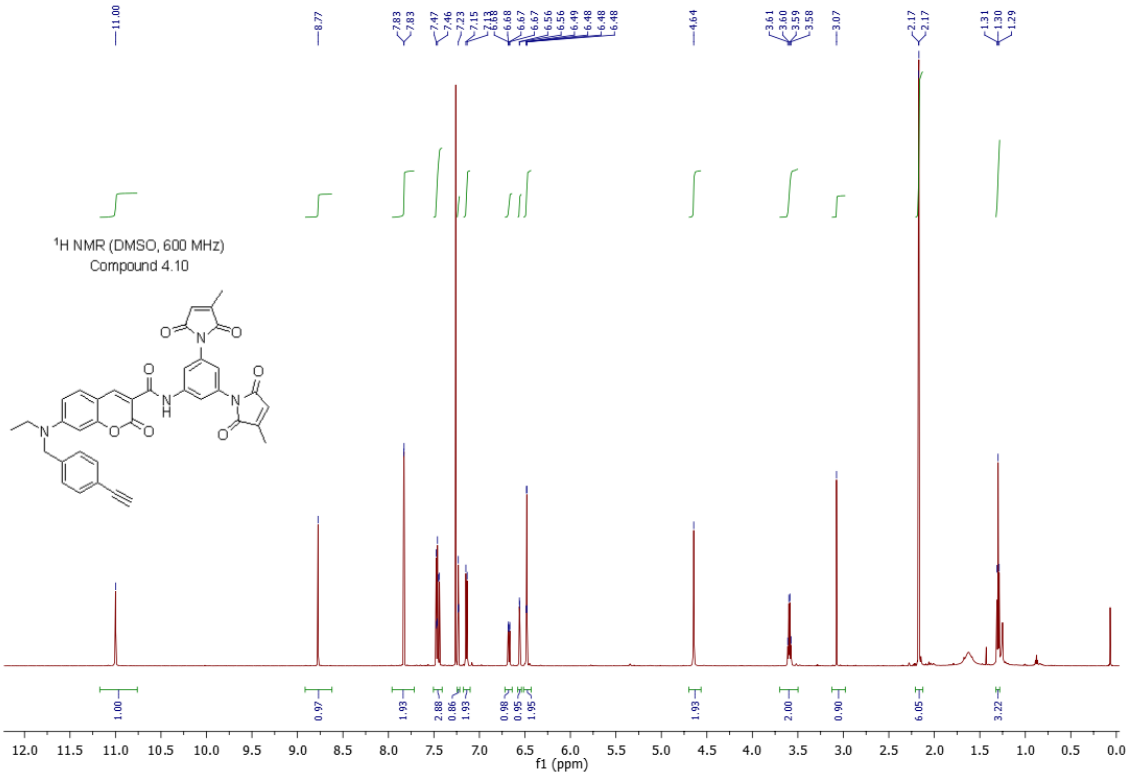

13C NMR (CDCl3, 100 MHz)
Compound 4.13
OH
156.83
150.06
140.29
132.53
130.32
126.60
120.65
105.21
103.43
99.29
83.74
53.92
45.51
12.32
f1 (ppm)

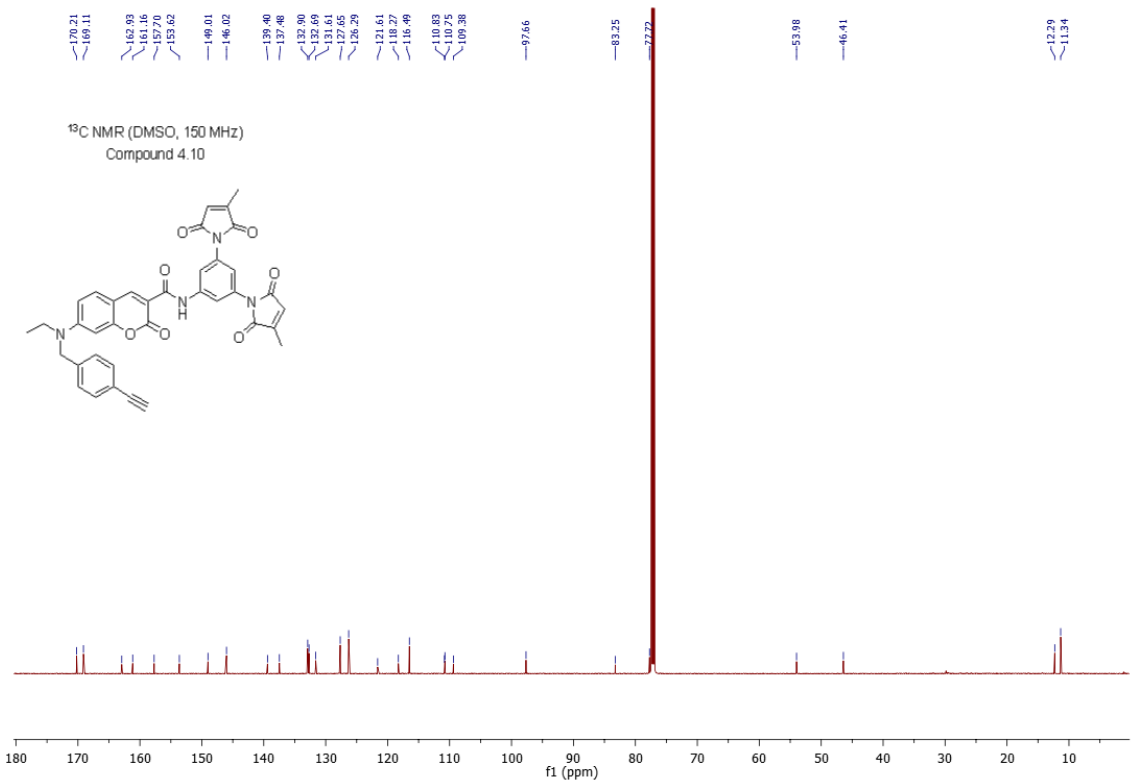

1H NMR (CDCl3, 400 MHz)
Compound 4.11
O
H
OH
11.57
9.52
9.52
7.46
7.44
7.27
7.26
7.25
7.14
7.12
6.27
6.27
6.25
6.24
6.13
6.12
4.59
3.55
3.53
3.51
3.49
3.06
1.27
1.26
1.24
0.91
0.93
1.89
2.00
1.87
0.96
0.93
1.90
1.92
0.89
3.34
f1 (ppm)

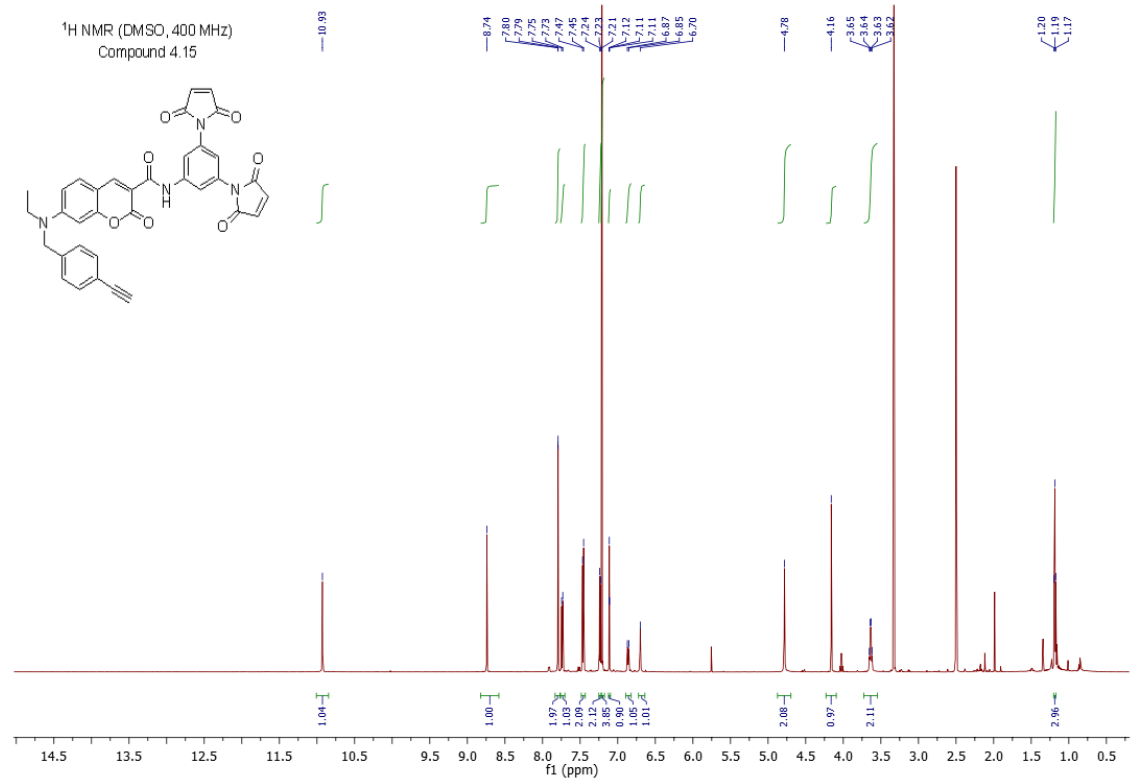

192.39
164.29
154.88
138.03
135.46
132.63
126.23
121.20
112.12
104.77
97.54
83.25
53.44
45.79
12.24
0.00
13C NMR (CDCl3, 100 MHz)
Compound 4.11
O
H
OH
N
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

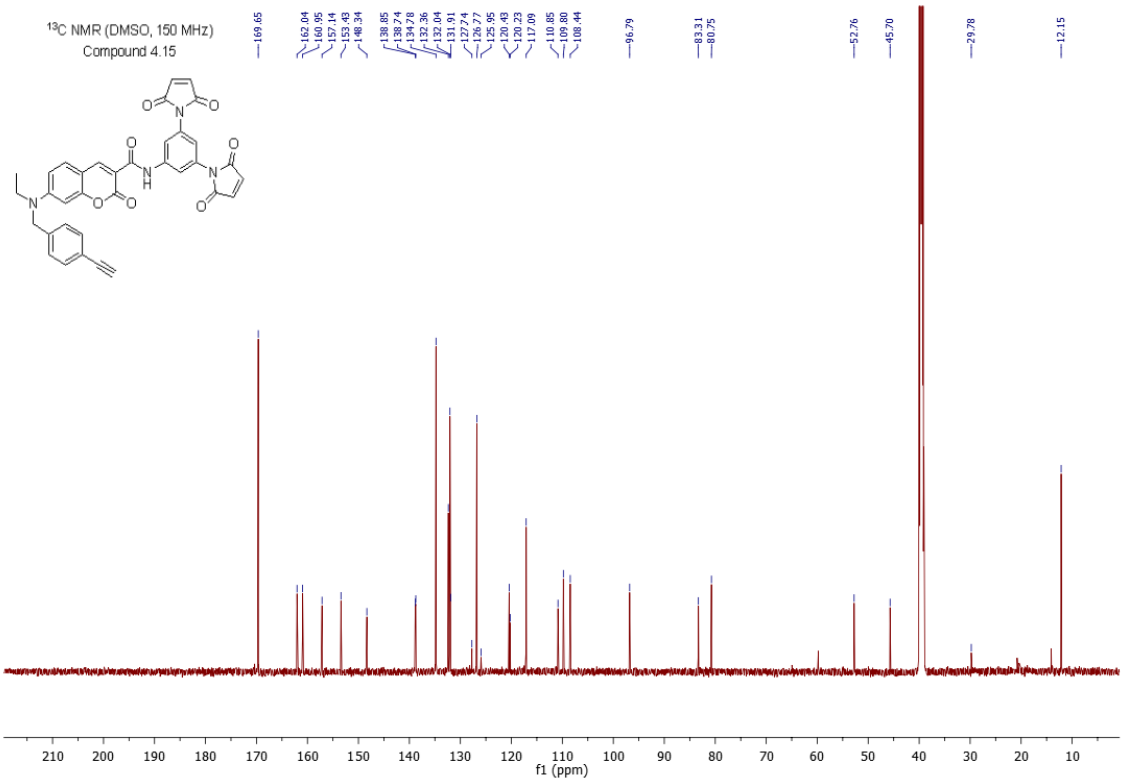

1H NMR (CDCl3, 400 MHz)
Compound 4.14
O
OH
O
O
N
8.67
7.48
7.46
7.44
7.26
7.14
7.12
6.73
6.73
6.71
6.71
6.57
6.56
4.67
3.65
3.63
3.61
3.59
3.08
1.34
1.32
1.30
1.00
3.13
2.12
1.06
1.03
2.03
2.13
0.97
3.12
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)

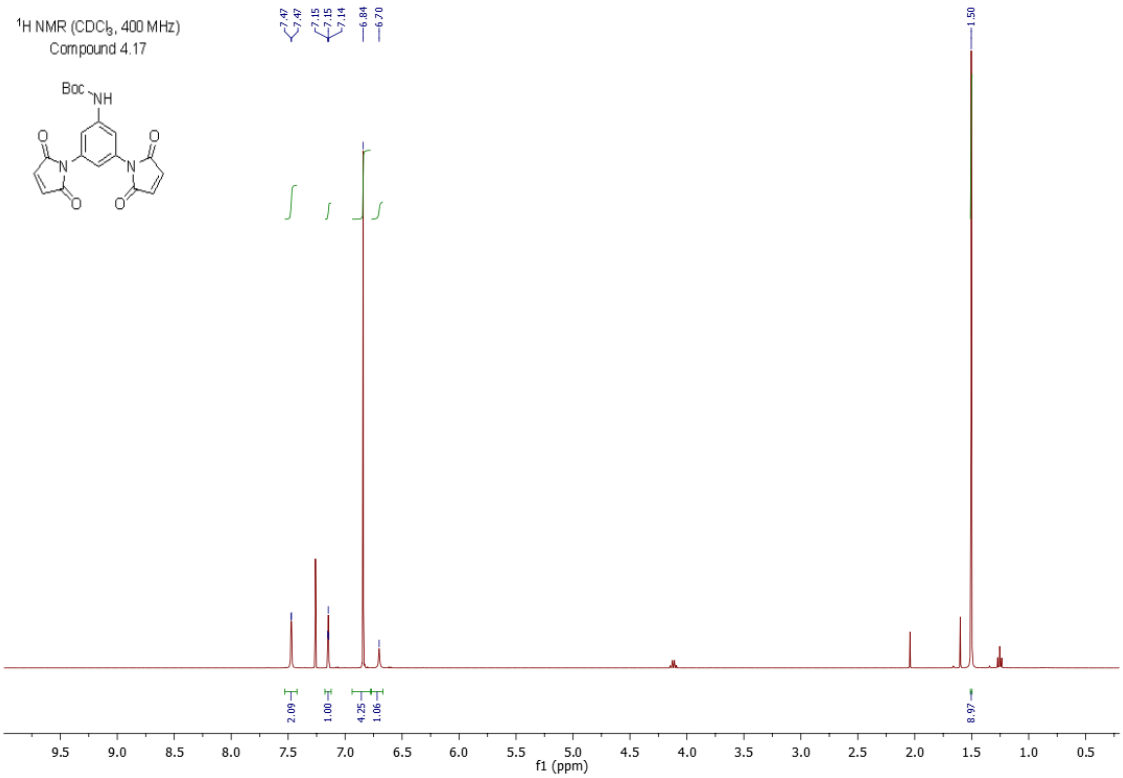

^{13}C NMR (CDCl$_3$, 100 MHz)
Compound 4.14
165.44
164.28
157.92
154.47
150.60
136.95
132.98
132.11
126.25
121.85
111.42
109.33
106.93
97.80
83.12
54.07
46.60
12.29
OH
f1 (ppm)

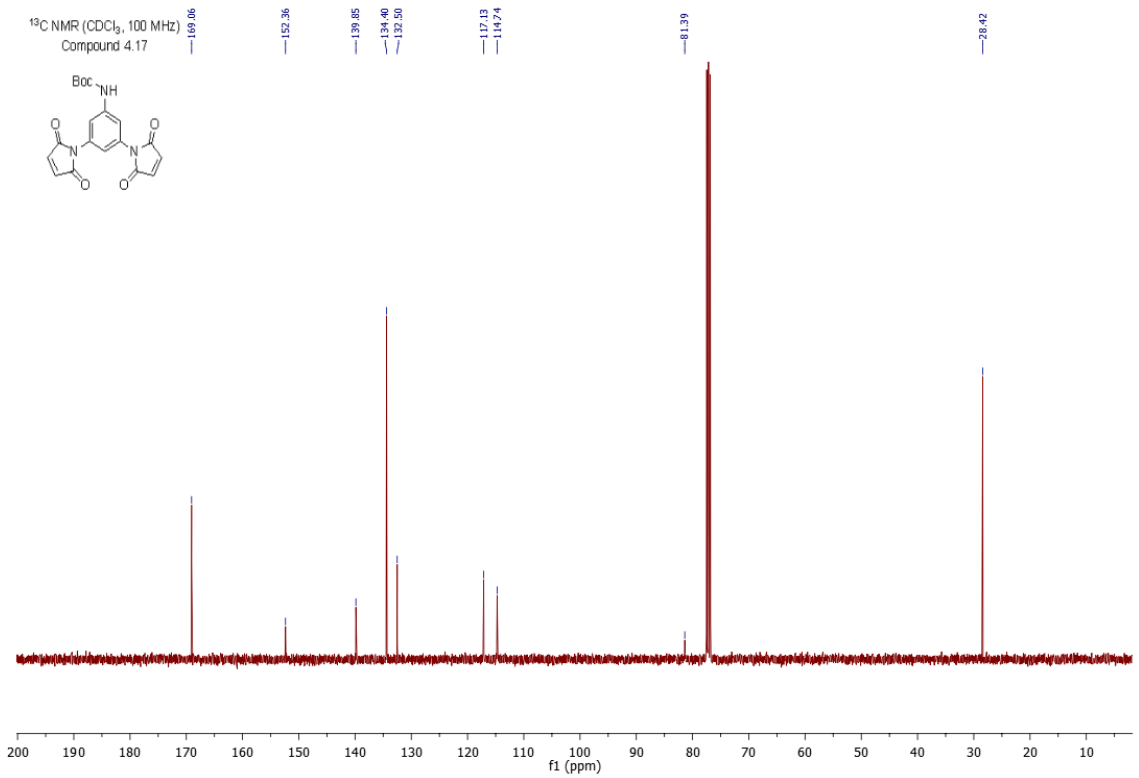

^{1}H NMR (DMSO, 600 MHz)
Compound 4.10
11.00
8.77
7.83
7.47
7.46
7.23
7.15
6.68
6.67
6.56
6.49
6.48
6.40
4.64
3.61
3.60
3.59
3.58
3.07
2.17
2.17
1.31
1.30
1.29
f1 (ppm)

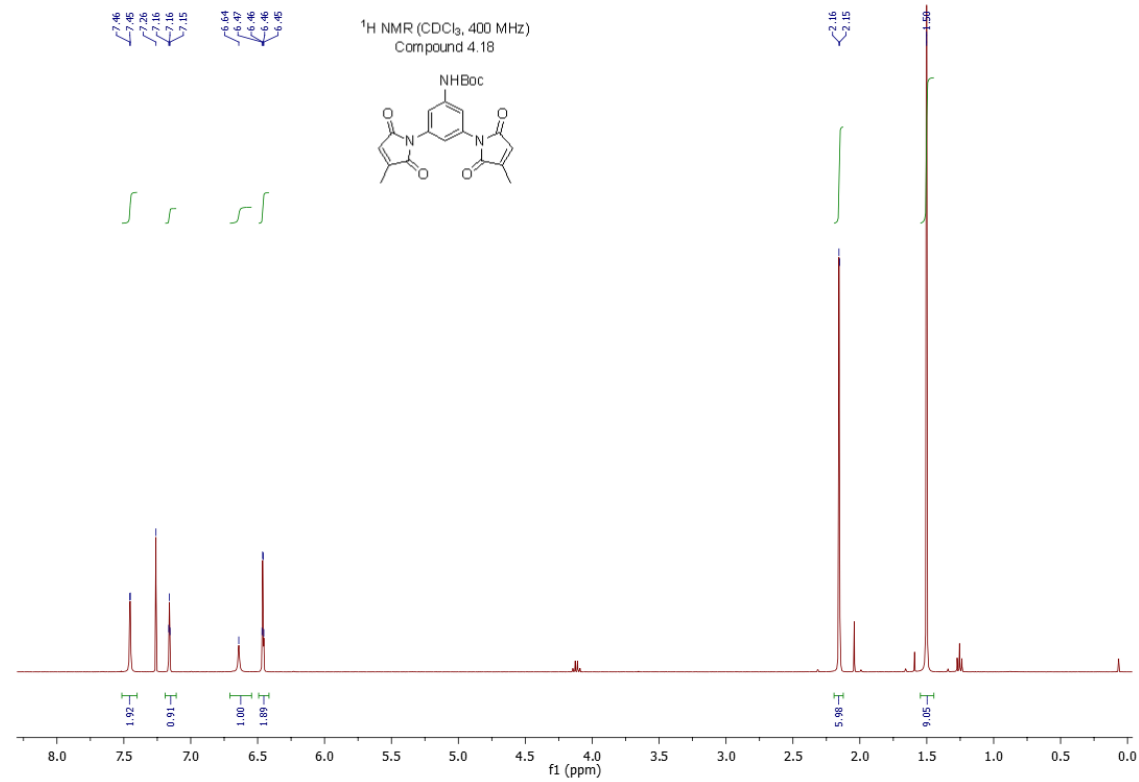

170.21
169.11
162.93
161.16
157.70
153.62
149.01
146.02
139.40
137.48
132.90
132.69
131.61
127.65
126.29
121.61
118.27
116.49
110.83
110.75
109.38
97.66
83.25
77.22
53.98
46.41
12.29
11.34
13C NMR (DMSO, 150 MHz)
Compound 4.10
f1 (ppm)

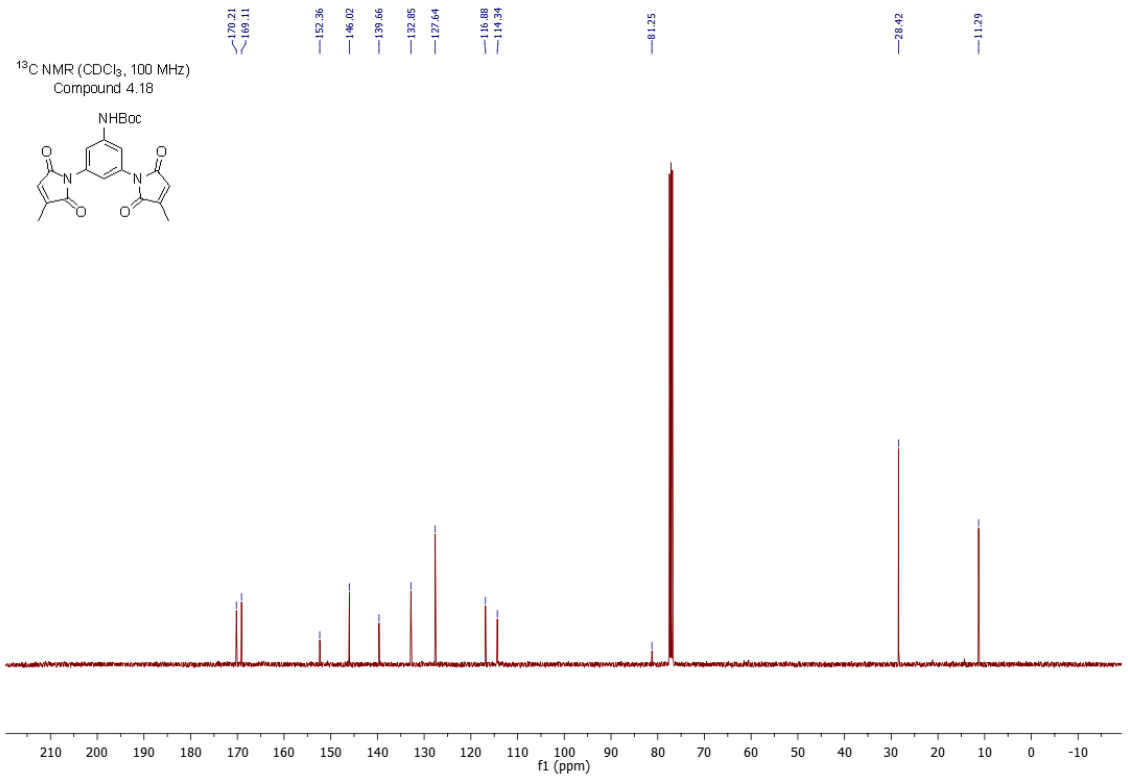

10.93
8.74
7.80
7.79
7.75
7.73
7.47
7.45
7.24
7.23
7.21
7.12
7.11
7.11
6.87
6.85
6.70
4.78
4.16
3.65
3.64
3.63
3.62
1.20
1.19
1.17
1H NMR (DMSO, 400 MHz)
Compound 4.15
1.04
1.00
1.97
1.03
2.09
2.12
3.85
0.90
1.05
1.01
2.08
0.97
2.11
2.96
f1 (ppm)

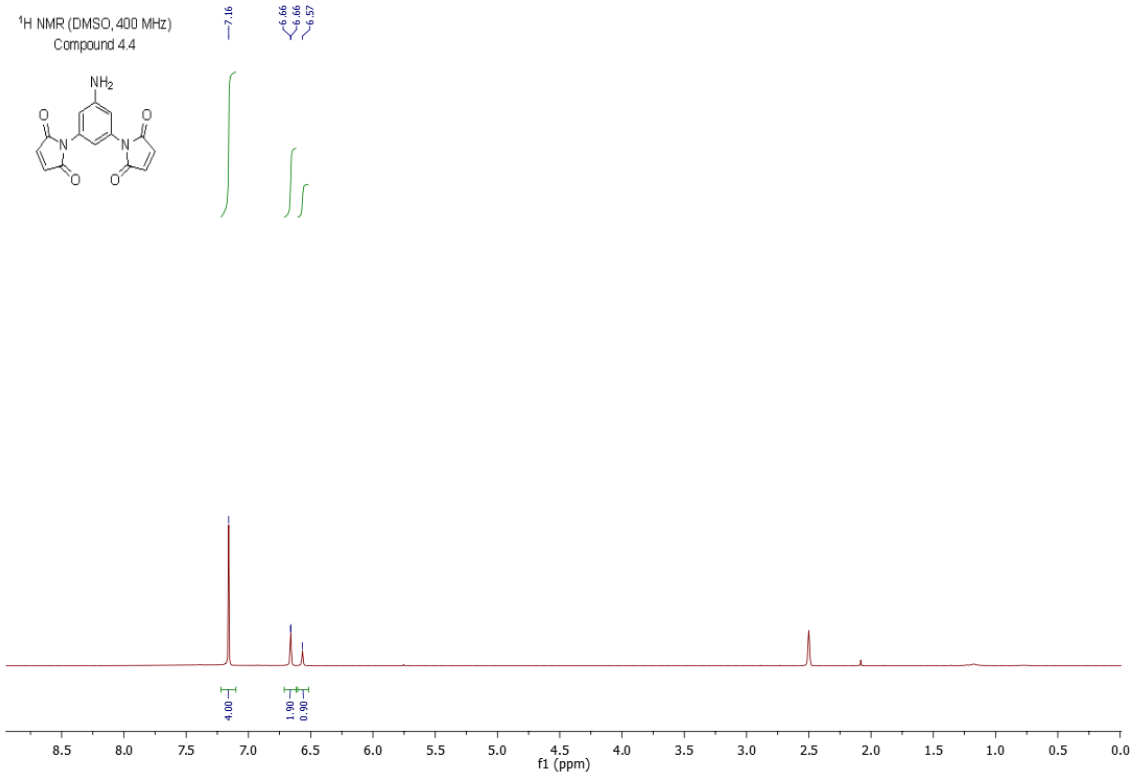

^{13}C NMR (DMSO, 150 MHz)
Compound 4.15
169.65
162.04
160.95
157.14
153.43
148.34
138.85
138.74
134.78
132.36
132.04
131.91
127.74
126.77
125.95
120.49
120.23
117.09
110.85
109.80
108.44
96.79
83.31
80.75
52.76
45.70
29.78
12.15
f1 (ppm)

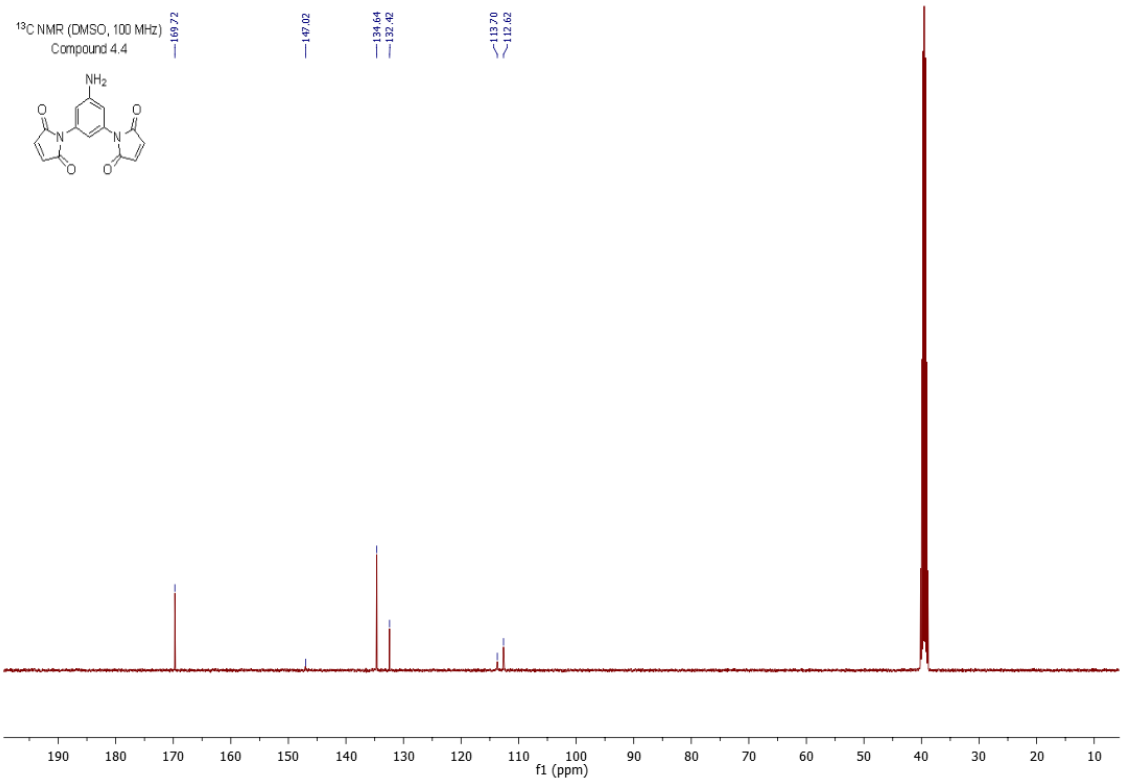

^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 4.17
7.47
7.47
7.15
7.15
7.14
6.84
6.70
1.50
2.09
1.00
4.25
1.06
8.97
f1 (ppm)

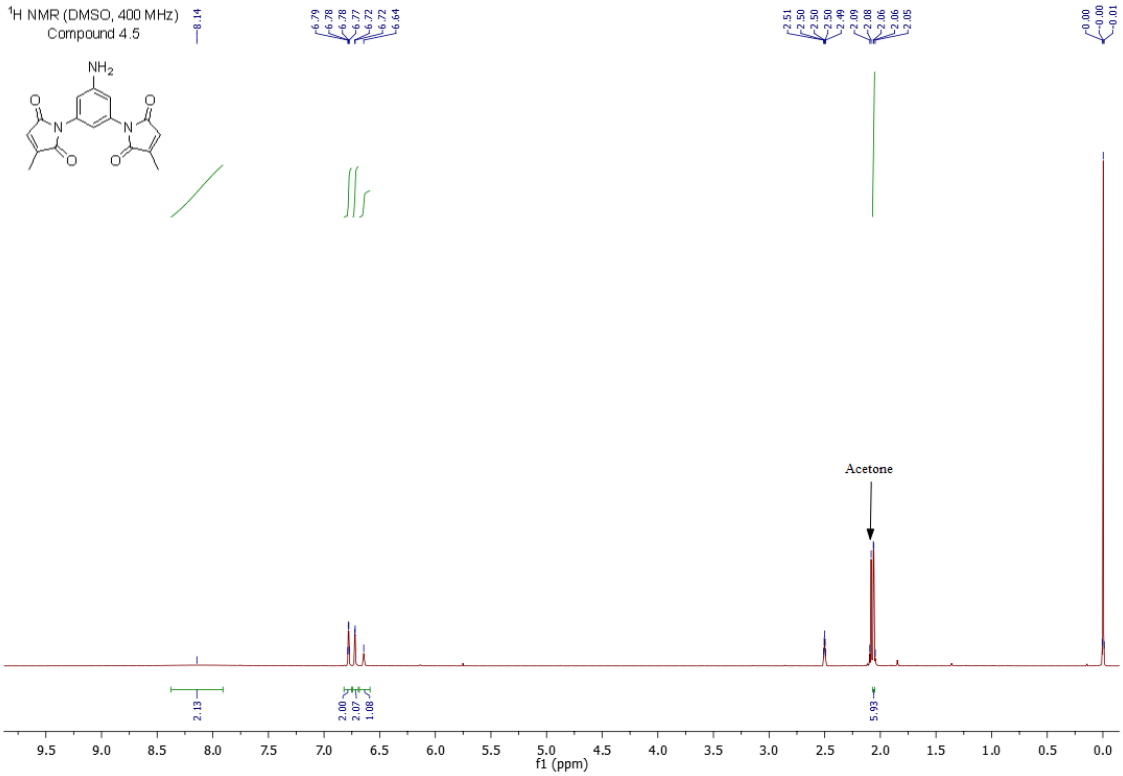

13C NMR (CDCl3, 100 MHz)
Compound 4.17
Boc
NH
169.06
152.36
139.85
134.40
132.50
117.13
114.74
81.39
28.42
f1 (ppm)

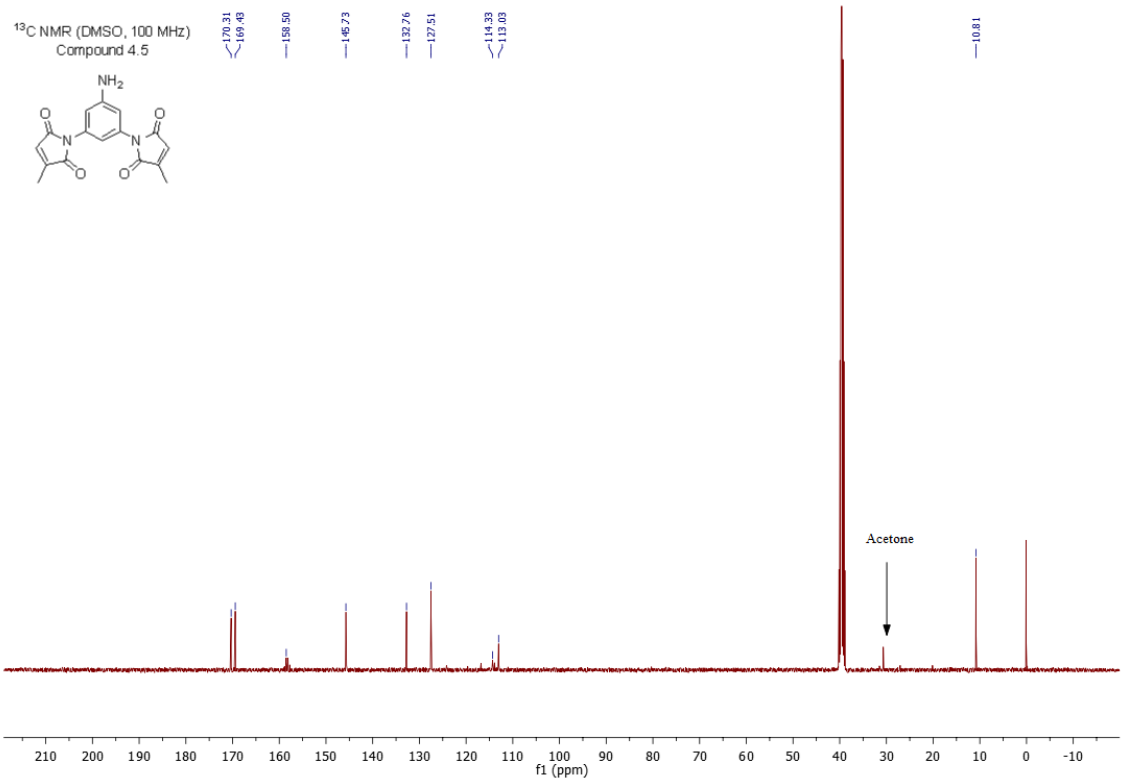

1H NMR (CDCl3, 400 MHz)
Compound 4.18
NHBoc
7.46
7.45
7.26
7.16
7.16
7.15
6.64
6.47
6.46
6.46
6.45
2.16
2.15
1.50
1.92
0.91
1.00
1.89
5.98
9.05
f1 (ppm)

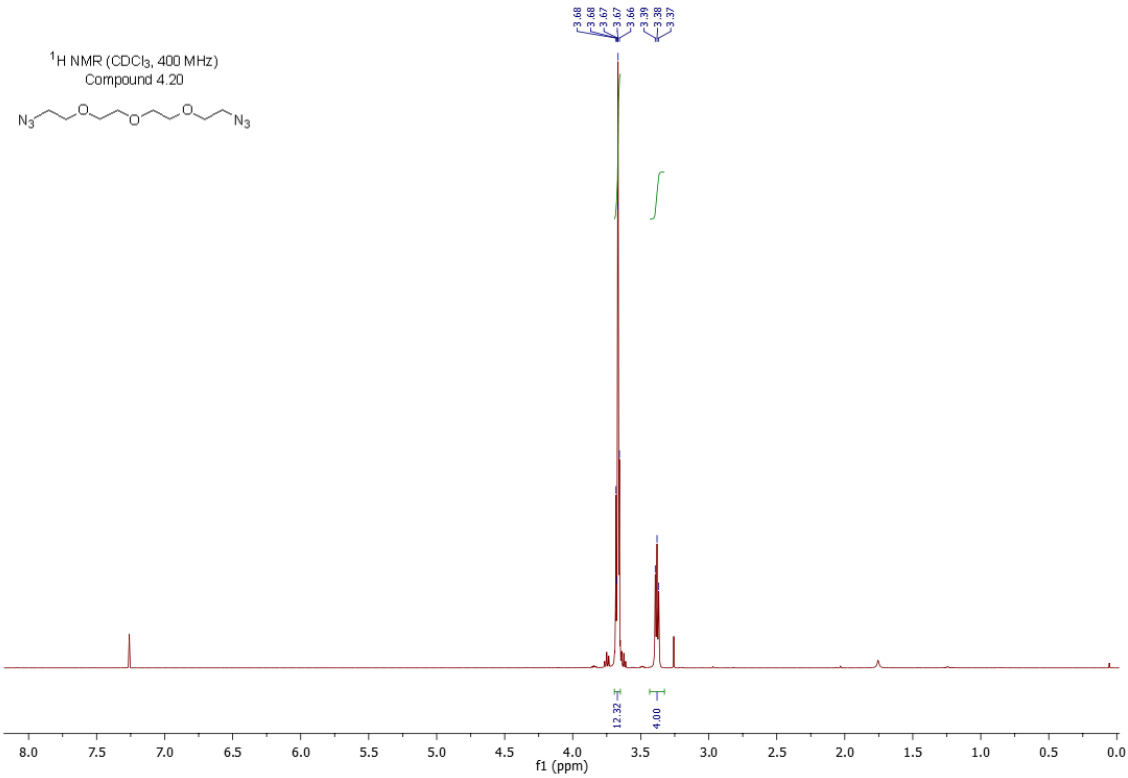

170.21
169.11
152.36
146.02
139.66
132.85
127.64
116.88
114.34
81.25
28.42
11.29
13C NMR (CDCl3, 100 MHz)
Compound 4.18
NHBoc
f1 (ppm)

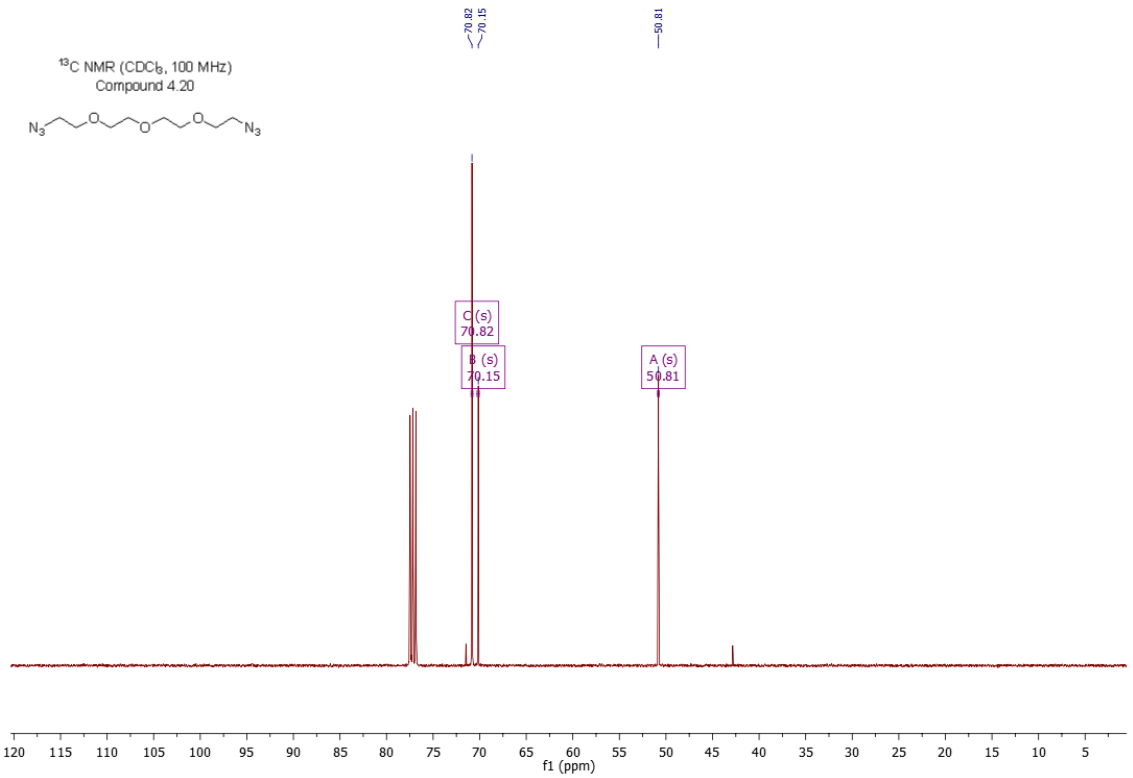

7.16
6.66
6.66
6.57
1H NMR (DMSO, 400 MHz)
Compound 4.4
NH2
4.00
1.90
0.90
f1 (ppm)

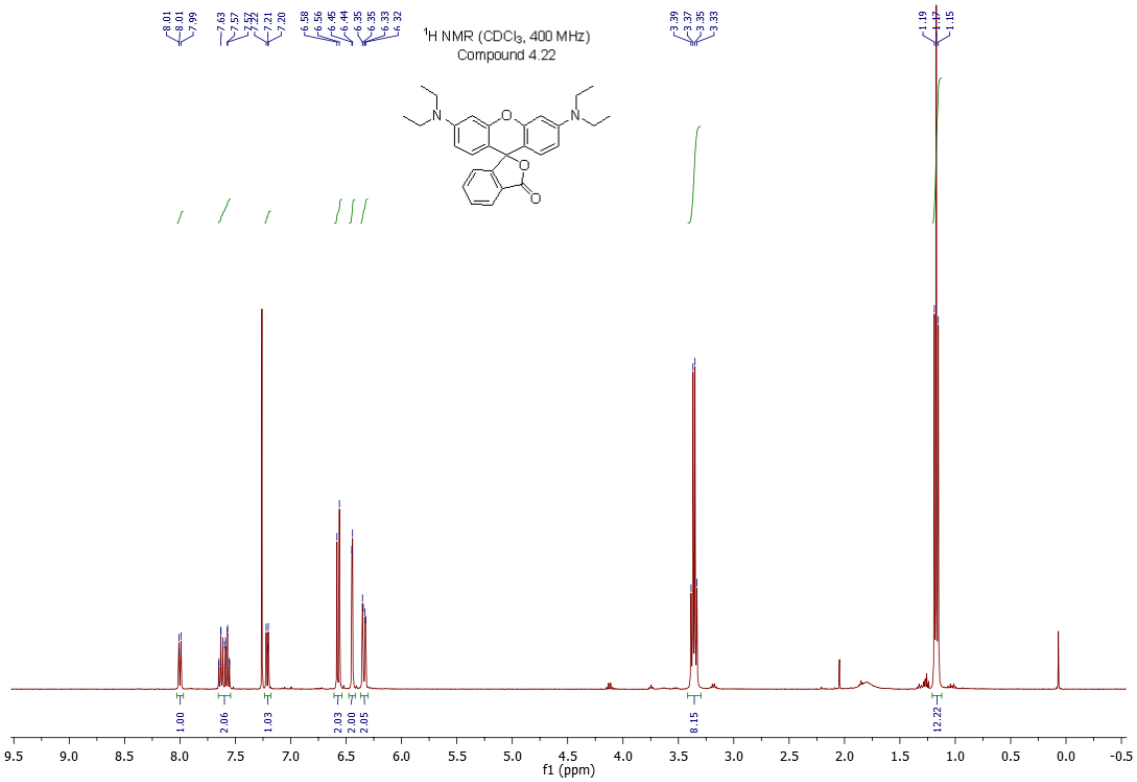

13C NMR (DMSO, 100 MHz)
Compound 4.4
169.72
147.02
134.64
132.42
113.70
112.62
NH2
f1 (ppm)

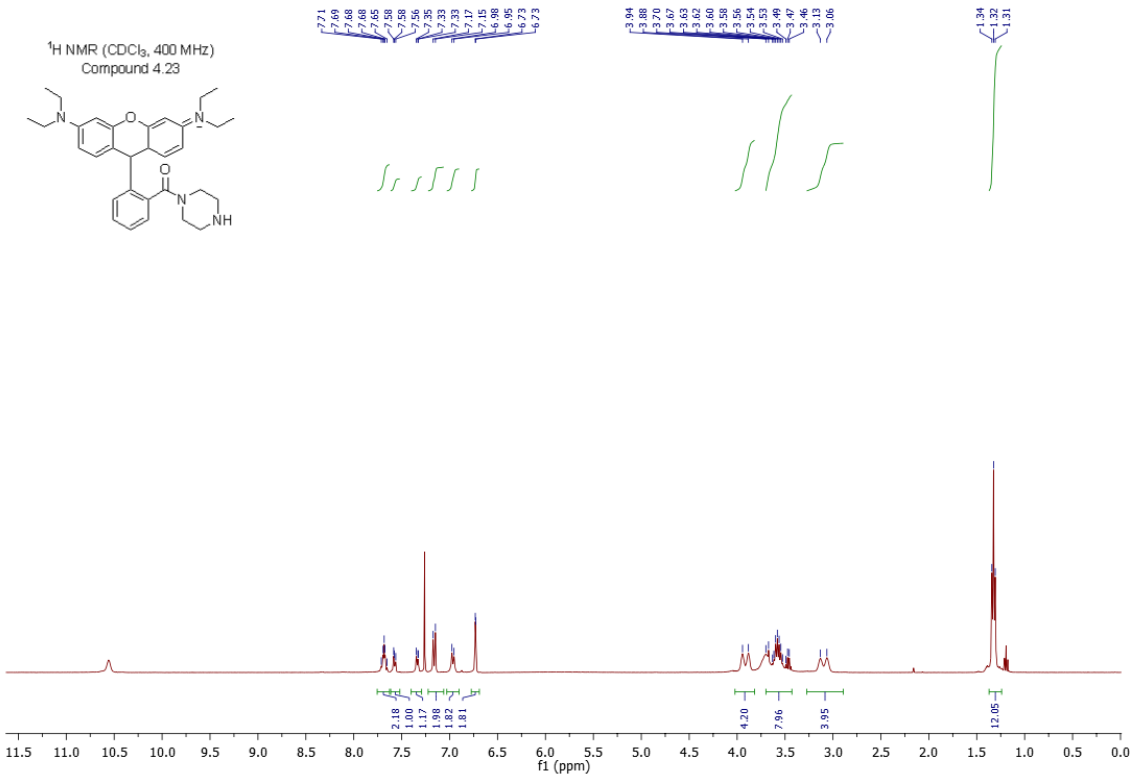

1H NMR (DMSO, 400 MHz)
Compound 4.5
8.14
6.79
6.78
6.78
6.77
6.72
6.72
6.64
2.51
2.50
2.50
2.50
2.49
2.09
2.08
2.06
2.06
2.05
0.00
-0.00
-0.01
NH2
Acetone
2.13
2.00
2.07
1.08
5.93
f1 (ppm)

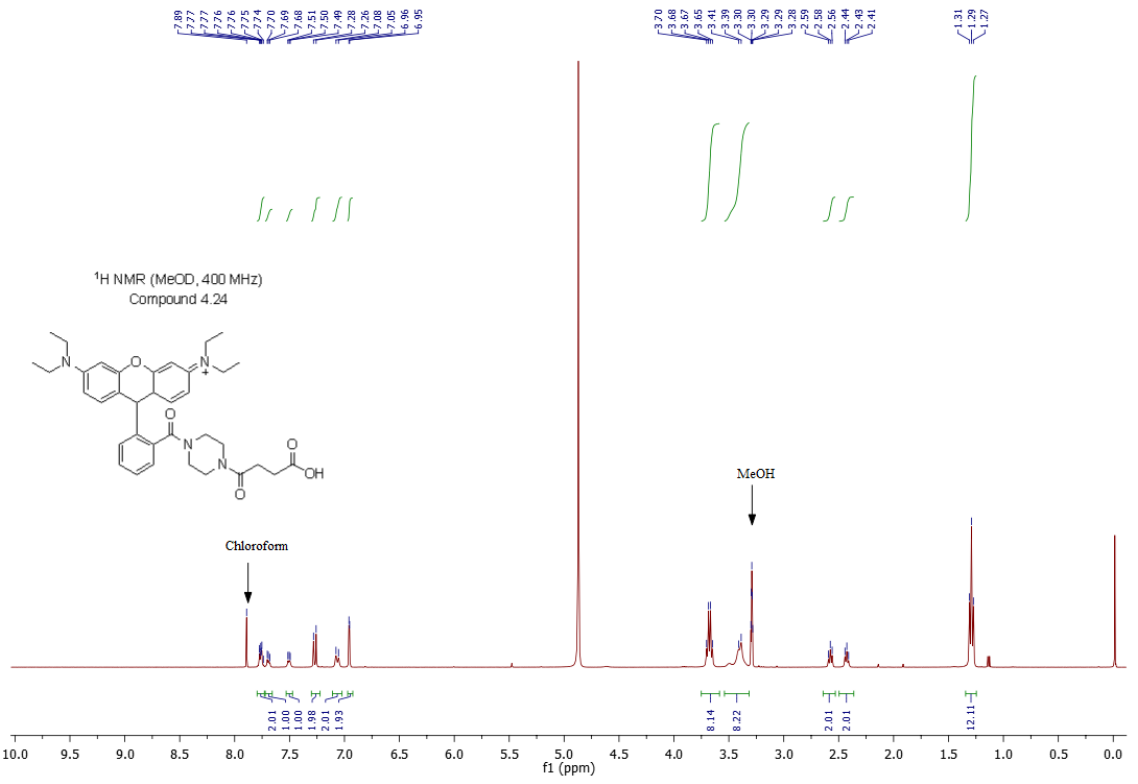

13C NMR (DMSO, 100 MHz)
Compound 4.5
NH2
170.31
169.43
158.50
145.73
132.76
127.51
114.33
113.03
10.81
Acetone
f1 (ppm)


1H NMR (CDCl3, 400 MHz)
Compound 4.20
N3
N3
3.68
3.68
3.67
3.67
3.66
3.39
3.38
3.37
12.32
4.00
f1 (ppm)


70.82
70.15
50.81
^{13}C NMR ($CDCl_3$, 100 MHz)
Compound 4.20
C (s)
70.82
B (s)
70.15
A (s)
50.81
f1 (ppm)


^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 4.22
f1 (ppm)


^{1}H NMR ($CDCl_3$, 400 MHz)
Compound 4.23
NH
f1 (ppm)


^{1}H NMR (MeOD, 400 MHz)
Compound 4.24
OH
MeOH
Chloroform
f1 (ppm)

Chapter 5: Conclusion

5.1. Studying Maleimide Reactivity to Improve the FlARe Probe

5.1.1. Project Goals

This project aimed to improve the selectivity and stability of the current FlARe probe design. Thus, a mechanistic study was performed to study how substituents on the maleimide C=C bond influenced thiol addition and hydrolysis using a library of fluorogenic coumarin substrates (Figure 5.1). Ideally, finding a substituted maleimide that demonstrates lower rates of thiol addition or hydrolysis than the methoxymaleimide would be beneficial in the design of a FlARe probe that is more selective for the dC10 tag and/or more stable in the intracellular milieu of the cell.

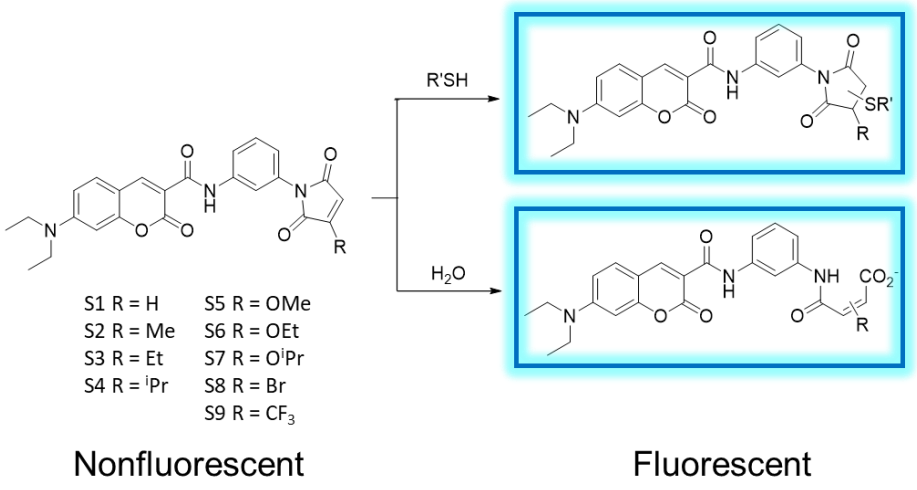


Figure 5.1. Approach to studying maleimide thiol addition and hydrolysis using fluorogenic coumarin substrates.

5.1.2. Results

Physical organic studies showed that the thiol addition of β-mercaptoethanol (BME) to a substituted maleimide was governed primarily by the electronic properties of the substituent. Additionally, thiol addition regioselectivity was also found to be substituent

dependant. Maleimide hydrolysis was found to be influenced by the steric characteristics of the substituent, and hydroxide attack always occurred on the carbonyl closest to the substituted alkene carbon.

In addition to the analysis using linear free energy relationships, the coumarin maleimide substrate carrying the isopropyl substituent demonstrated similar attenuated reactivity against BME as the methoxy substituted substrate. Since the methoxymaleimide is the substituted maleimide used in the FlARe probe design (**YC20**), we reasoned that an isopropylmaleimide variant, **KT12**, would demonstrate improved stability in aqueous environments while maintaining similar selectivity for the dC10 tag. Following **KT12's** synthesis, it was found that the novel FlARe probe showed better stability against hydrolysis than **YC20**. However, it was not deemed an improvement over **YC20** due to its lower solubility and selectivity for the dC10 tag.

5.1.3. Perspectives

The substituents explored in this study were restricted to alkyl substituents and their alkoxy variants, based on the past developments of the FlARe probe. Exploring other substituted maleimides would help complement the above study. For example, substituted maleimides carrying an electron-donating group, such as aminomaleimide, may also be attenuated against thiol addition given the reaction's electronic dependence and the similar electron-donating nature of the methoxy group. This would provide direction towards FlARe probes with greater selectivity for the target peptide tag. More sterically hindered substituted maleimides should also be explored as even though the substituent's electronic characteristics primarily influence thiol addition, its steric character also plays a role. This was made evident in the equally attenuating effects of the isopropyl substituent compared to the methoxy

substituent. As such, multiple substituents on the maleimide C=C bond, such as 2,3-dimethylmaleimide, may be worth exploring.

Further understanding the mechanisms that govern maleimide thiol addition will not only help improve the FlARe probe design but also aid in the rational design of the dC10 tag. Studying the effect of varying thiol nucleophiles and their reaction with *N*-arylmaleimides would benefit the future augmentation of the dC10 tag. This was demonstrated by the dC10* tag, which increased the likelihood of the target peptide existing as the thiolate. As a result, the reactivity of the tag with the FlARe probe was improved. Therefore, continued studies into maleimide reactivity will help not only attenuate the FlARe probe's reactivity against monothiols but also contribute to improving the dC10 tag's reactivity with the probe itself.

5.2. Imaging New Elements of the FlARe Labelling Method in Cells

5.2.1. Project Goals

A new target peptide that reacted faster with FlARe probes, dC10*, and a novel green fluorogenic FlARe probe, **YC23**, was developed and tested *in vitro*. This project's goal was to establish these novel elements of FlARe in a cellular setting. To do this, we aimed to label histone H2B tagged with dC10* with **YC20** and to label H2B-dC10 with **YC23** (Figure 5.2).

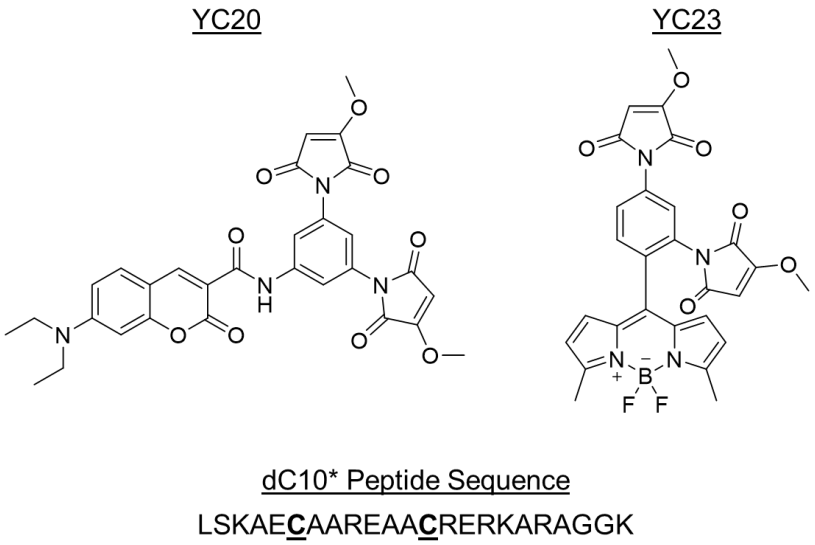


Figure 5.2. Structures of FlARe probes **YC20** and **YC23** and the peptide sequence of the dC10* tag.

5.2.2. Results

After establishing a viable procedure for cellular labelling and fluorescence imaging, histone H2B-dC10* expressed in HeLa cells was labelled with **YC20**. Specifically, the nuclear-localized blue fluorescence of activated **YC20**, observed only in cells expressing tagged protein, demonstrated the successful implementation of the improved target peptide in cells. However, a side-by-side comparison of labelling times between dC10* and dC10 tagged H2B protein showed no significant difference. This suggests that other factors, such as the ability to pass through the cell membrane, may influence labelling times.

Histone H2B-dC10 that was expressed in HeLa cells was also labelled with **YC23**. The green fluorescence from activated **YC23**, observed only in HeLa cells expressing H2B-dC10, indicated successful FlARe labelling. The fluorescence was also confirmed to be localized to the nucleus, where the histone protein is expected to be, using the DNA binding Hoechst counterstain. Although labelling mNeptune-dC10* proved more difficult, histone

H2B-dC10 labelling still established the novel green fluorogenic FlARe probe, which contributed towards the expansion of the FlARe method's colour palette.

5.2.3. Perspectives

Although counterstaining with Hoechst was used to confirm the nuclear localization of the fluorescence indicative of successful labelling, labelling the histone H2B protein with a specific antibody could also be performed to complement the colocalization studies. Furthermore, the nuclear localization of H2B-dC10* labelled with the blue fluorogenic **YC20** could not be confirmed due to the overlapping blue emission of Hoechst. Thus, a histone H2B antibody functionalized with a red-emitting fluorophore would allow for the assessment of colocalization.

While labelling mNeptune-dC10 with **YC23**, only a localized section of the cell exhibited green fluorescence as opposed to the entire cell, as one would expect for the fluorogenic labelling of a cytoplasmic protein. We hypothesized that not all the mNeptune present was labelled, whichcould be addressed by increasing labelling times or the concentration of labelling agent. Alternatively, maltose binding protein (MBP) carrying the dC10 tag could also be used, as it is both a cytoplasmic protein and the model protein used when testing FlARe probes *in vitro*. Since the FlARe probe's reactivity with MBP would already be established, labelling the same MBP-dC10 target protein that was expressed in HeLa cells would eliminate concerns regarding tag accessibility.

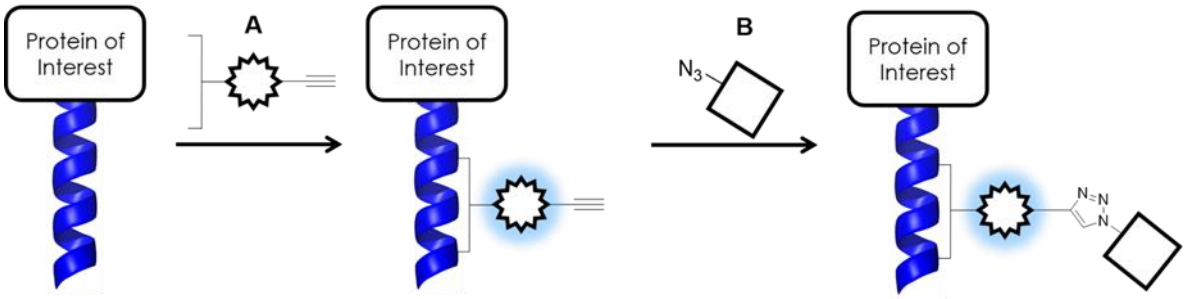
Lastly, the studies performed in this project can be expanded beyond nuclear and cytoplasmic proteins. Other localized proteins carrying the dC10 tag could also be labelled in order to fully demonstrate the flexibility and feasibility of the FlARe labelling technique.

Alternatively, localization tags can be used to demonstrate the systematic labelling of a dC10 tagged model protein in different cellular compartments.

5.3. A Fluorogenic and Site-specific FlARe Bioconjugating Linker

5.3.1. Project Goals

In this project, we aimed to expand the scope of the current FlARe probe design towards the formation of bioconjugates in a site-specific and fluorogenic manner. Therefore, a terminal alkyne was integrated into the coumarin-based FlARe probe design. As such, a fluorogenic linker was developed such that a dC10 tagged protein could be labelled with a linker that introduced an alkyne functional group to the exposed target peptide tag (Scheme 5.1). Due to the FlARe design, this initial protein labelling could be monitored *in situ* with fluorescence. Azide-functionalized cargo could then be conjugated to the protein using the copper-catalyzed azide-alkyne cycloaddition (CuAAC) click reaction for the formation of fluorescent bioconjugates *in vitro*.

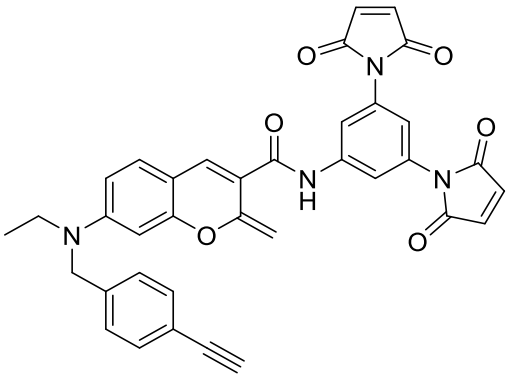


Scheme 5.1. Schematic of the site-specific fluorogenic bioconjugating linker strategy.

5.3.2. Results

The first linker, **4.6**, could conjugate rhodamine B azide (RhB-N_3) to MBP-dC10* at the target peptide tag, but failed to do so in a fluorogenic manner. However, in the absence of the hexyne tail, the corresponding FlARe probe, **4.1**, was fluorescent upon reacting with MBP-

dC10*. Therefore, we hypothesized that the flexibility of the hexyne tail on the FlARe linker disrupted the probe's fluorescence even after reacting with dC10*. A second linker, **4.10**, was developed that involved a more sterically hindered and rigid aromatic terminal alkyne. Additionally, the unsubstituted maleimides were used to allow faster reactivity with dC10* for the *in vitro* formation of bioconjugates. As such, the final FlARe bioconjugating linker, **4.15**, was synthesized (Figure 5.3). Similar to the initial design, **4.15** was able to site-specifically conjugate RhB-N_3 to MBP-dC10* but did so more rapidly and in a fluorogenic manner. This allowed the introduction of the alkyne group to MBP at the dC10* tag to be monitored *in situ* by fluorescence, after which RhB-N_3 was conjugated to MBP through **4.15** at the dC10* tag.



4.15

Figure 5.3. Structure of the fluorogenic FlARe based bioconjugating linker, **4.15**.

5.3.3. Perspectives

To demonstrate the proof-of-principle for this novel FlARe-based linker, RhB-N_3 was conjugated to MBP-dC10*. Other cargo such as azide-functionalized carbohydrates, fatty acids, or oligonucleotides could be conjugated to MBP-dC10* to complement these initial proof-of-principle demonstrations. Additionally, conjugation of azide-functionalized biotin

would allow for the downstream purification of bioconjugates. This would demonstrate the linker's versatility for the formation of therapeutically relevant and fluorescent bioconjugates.

Here, the unsubstituted maleimide was chosen for its rapid reaction with the dC10* tag. However, its reactivity with monothiols such as GSH limit **4.15** to the formation of bioconjugates *in vitro*. Using the methoxymaleimide would introduce a novel FlARe-based linker for *in vivo* bioconjugation, given that the methoxy substituent attenuates maleimide reactivity with small molecule monothiols. As such, applications involving the fluorogenic labelling of dC10 tagged protein in cells or the purification of tagged protein from cells using azide-functionalized biotin and streptavidin beads could be explored.

In demonstrating the linker's proof-of-principle using SDS PAGE, fluorescence resonance energy transfer (FRET) was observed to occur between the coumarin and rhodamine B fluorophores. Although unintentional, this FRET interaction presents the opportunity to monitor the copper catalyzed azide-alkyne cycloaddition of rhodamine B to the coumarin linker. This can be done by excitation of the coumarin and observing the increasing emission of rhodamine B over time. Furthermore, this interaction can also monitor the labelling reaction of the dC10* tag with the RhB-N_3-coumarin linker adduct since the FRET interaction will not occur until the maleimide quenching mechanism is abolished. In doing so, the steric effects of the RhB-N_3 cargo can be assessed, which was hypothesized to contribute to slowing down the dC10* labelling reaction. Alternatively, this FRET interaction could be optimized to design a novel, large Stokes shift, and red-emitting FlARe probe. Ideally, this red-emitting FlARe probe would be excited at the absorbance maximum of the coumarin fluorophore and emit from the rhodamine B fluorophore. This would be beneficial given the better penetration and visibility of probes emitting in the red region of the visible spectrum.

Evidently, the FRET observed FRET interaction introduces numerous opportunities that are worth exploring.

Although the FlARe linker's design was demonstrated to work as intended through the proof of principle experiments, further modifications could be made to the design. For example, the alkyne could be replaced with an azide functional group, allowing for the conjugation of cargo functionalized with strained cylcooctynes. This would offer a less toxic conjugation method using the strained-promoted azide-alkyne cycloaddition (SPAAC) click reaction. Positioning the azide close to the fluorogen may also render the linker doubly quenched, thereby controlling the stage of conjugation that can be monitored. Furthermore, ensuring the azide functional group remains small would limit the steric hindrance surrounding the probe, which was observed to decrease the rate of reaction with the dC10 tag.

Future work regarding the FlARe linker's development could also integrate cleavable functional groups for the triggered release of cargo in the cell. For example, ester functional groups that hydrolyze within the cell can be implemented near the terminal alkyne for this application. Alternatively, disulfide bonds would allow the triggered release of cargo in reductive cellular environments.

5.4. Conclusion

In summary, this thesis aimed to develop the current FlARe probe design, implement novel aspects of the FlARe labelling method in cells. and expand the FlARe probe's functional scope. In doing so, the work herein contributed to understanding how maleimide substituent effects govern maleimide thiol addition and hydrolysis. Furthermore, a new dC10* target peptide and green fluorogenic FlARe probe were demonstrated as viable additions to the

FlARe cellular imaging method. These additions are valuable for their efforts to improve the FlARe technique's sensitivity and speed as well as for broadening the colour palette of the protein labelling method. Lastly, a new site-specific and fluorogenic linker was designed by integrating a terminal alkyne into the current coumarin-based FlARe probe. As a result, a novel FlARe linker was designed, which could rapidly conjugate azide-functionalized cargo to dC10 tagged proteins. This process could also be monitored, allowing for the controlled assembly of therapeutically relevant bioconjugates. These results expand the current scope of the FlARe probes, demonstrating the functional versatility of the FlARe probe design and small molecule probes in general. Advancements in the FlARe protein labelling method are focused on improving the labelling times by better understanding maleimide reactivity and by improving the target peptide's reactivity. Additionally, the synthetic versatility of the FlARe probe is being used to integrate new bioorthogonal functional groups and fluorophores to expand the functional scope and colour palette of the protein labelling technique. Together, these developments contribute to cellular imaging, which is integral to both the fields of cell biology and chemical biology, by advancing the minimal maleimide-based fluorogenic protein labelling method, FlARe.