DESIGN, SYNTHESES AND APPLICATIONS OF FLUORESCENT DYES

A Dissertation

by

LIANGXING WU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2009

Major Subject: Chemistry

DESIGN, SYNTHESES AND APPLICATIONS OF FLUORESCENT DYES

A Dissertation

by

LIANGXING WU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Kevin Burgess
Committee Members,	Daniel Romo
	Coran Watanabe
	Gregory Reinhart
Head of Department,	David Russell

August 2009

Major Subject: Chemistry

ABSTRACT

Design, Syntheses and Applications of Fluorescent Dyes.
(August 2009)
Liangxing Wu, B.S., University of Science and Technology of China
Chair of Advisory Committee: Dr. Kevin Burgess

New methodologies for the efficient syntheses of 4,4-difluoro-4-bora-3a,4a-diaza-*s*-indacenes (BODIPYs) and rosamines were developed. A serendipitous discovery led to a new reaction which afforded BODIPYs in high yields. Systematic studies of the kinetics and mechanisms of the new reaction were performed. A series of BODIPYs were successfully prepared using the new approach. A simple and efficient synthesis of rosamines with cyclic-amine substituents was devised. These new rosamines showed interesting anti-tumor activities.

Several types of novel fluorescent compounds were prepared. Highly fluorescent GFP-chromophore analogs were designed and synthesized. The correlation between the optical properties and the structures was investigated. New pyronin dyes with *meso*-heteroatom substituents were efficiently prepared. The fluorescence properties of these compounds were highly dependent on the nature of the *meso*-substituents. A set of BODIPY dyes that fluoresce brightly above 600 nm were made. They were then used as acceptors to prepare water-soluble through-bond energy transfer cassettes. All the cassettes had complete energy transfer and high quantum yields in MeOH. A few also had good fluorescence properties in aqueous media and even on proteins.

The through-bond energy transfer cassettes were used to monitor protein-protein interactions. In order to test our hypothesis, an artificial protein interaction system was built by utilizing the biotin/(strept)avidin interactions. Thus Atto425-BSA-biotin, streptavidin-cassette1 and avidin-cassette2 were prepared. The interactions between Atto425-BSA-biotin and cassette labeled (strept)avidin were successfully detected *in vitro* and in living cells by fluorescence techniques.

DEDICATION

To

my wife, Yingrui Dai

and

my parents

ACKNOWLEDGEMENTS

First I would like to gratefully acknowledge my advisor Dr. Kevin Burgess for the enthusiastic supervision during my research in the group. I appreciate all his contributions of ideas, motivation, patience, and funding to make my Ph.D. experience enjoyable and stimulating. I also want to thank my committee members, Dr. Daniel Romo, Dr. Coran Watanabe and Dr. Gregory Reinhart for their guidance and support throughout the course of this research.

Thanks to Jill Powers, Angie Medina and Jade Kennedy for their assistance through out all these years.

Thanks to my friends and former colleagues Drs. Sam Reyes and Yu Li Angell for their friendship and help. Thanks also go to all the current Burgess group members, especially the dyes group.

Finally, thanks to my wife, Yingrui Dai, and my parents for their love, patience and support.

TABLE OF CONTENTS

Page

ABSTRACT iii

DEDICATION iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES ix

LIST OF TABLES xiii

LIST OF SCHEMES xv

CHAPTER I INTRODUCTION 1

1.1 Fluorescence 1
1.2 Multiplexing with FRET and TBET 4
1.3 Labeling of proteins 8
 1.3.1 Covalent labeling techniques 8
 1.3.2 Non-covalent labeling methods 17
 1.3.3 Summary 20
1.4 Delivery of protein into cells 20
1.5 Detection of protein-protein interactions 21

CHAPTER II DESIGN AND SYNTHESES OF HIGHLY FLUORESCENT GFP-CHROMOPHORE ANALOGS 25

2.1 Introduction 25
2.2 Results and discussions 33
 2.2.1 Syntheses of the highly fluorescent GFP-chromophore analogs ... 33
 2.2.2 UV and fluorescence properties of the GFP-chromophore analogs 48
2.3 Conclusions 54

CHAPTER III A NEW SYNTHESIS OF BORAINDACENE (BODIPY) DYES 56

3.1 Introduction 56

Page

3.2 A new approach to symmetric BODIPYs 61
3.3 Kinetic and mechanistic studies of the new reaction 66
3.4 Conclusion 69

CHAPTER IV ROSAMINES WITH CYCLIC AMINE SUBSTITUENTS: SYNTHESES AND SPECTROSCOPIC PROPERTIES AND APPLICATIONS 70

4.1 Introduction 70
4.2 Syntheses of aminoxanthones and rosamines 78
4.3 Spectroscopic properties of rosamine derivatives 87
4.4 Biological activity test 99
4.5 Conclusions 107

CHAPTER V NEW PYRONIN DYES WITH MESO-HETEROATOM SUBSTITUENTS: SYNTHESIS AND SPECTROSCOPIC PROPERTIES 109

5.1 Introduction 109
5.2 Syntheses of new pyronin derivatives 113
5.3 Photophysical properties of new pyronin dyes 116
5.4 Conclusion 122

CHAPTER VI WATER-SOLUBLE THROUGH-BOND ENERGY TRANSFER CASSETTES: SYNTHESES, SPECTROSCOPIC PROPERTIES AND APPLICATIONS 123

6.1 Introduction 123
6.2 Syntheses of water-soluble through-bond energy transfer cassettes 129
  6.2.1 Syntheses of donor fragments 129
  6.2.2 Syntheses of acceptor parts 131
  6.2.3 Syntheses of cassettes 138
6.3 Spectroscopic properties of BODIPY acceptors and cassettes 140
6.4 Protein-protein interaction studies 145
6.5 Conclusions 154

CHAPTER VII CONCLUSIONS AND OUTLOOK 155

7.1 Conclusions 155
  7.1.1 New methodologies for the syntheses of BODIPY and rosamine 155
  7.1.2 Syntheses of novel fluorescent compounds 156

Page

7.1.3 Monitoring of protein-protein interactions *in vitro* and in living cells 158
7.2 Outlook 160
7.2.1 Rosamine derivatives 160
7.2.2 Water-soluble through-bond energy transfer cassettes 162

REFERENCES 167

APPENDIX A 182

APPENDIX B 184

APPENDIX C 243

APPENDIX D 269

APPENDIX E 314

APPENDIX F 332

VITA 482

LIST OF FIGURES

FIGURE Page

1.1. A typical Jablonski diagram 1

1.2. Jablonski diagram for two photon excitation 2

1.3. (a) Scheme for multiplexing; (b) ideal multiplexing output; (c) common multiplexing output; (d) reasons for the diminished fluorescence output as the emission maxima occur further from excitation source 4

1.4. Dependence of the orientation factor k^2 on the direction of the dipoles . 6

1.5. A typical diagram for through-bond energy transfer cassette 7

1.6. Labeling of proteins by modification of amino acid sidechains 9

1.7. Labeling of proteins by modification of terminal amino acids 11

1.8. Other direct labeling methods 13

1.9. Illustrative examples of enzyme catalyzed labeling of proteins 16

1.10. Illustrative examples of non-covalent labeling of proteins 18

1.11. Proposed system for detecting protein-protein interactions 24

2.1 (a) X-ray structure; (b) chromophore of GFP; (c) schematic diagram of the interactions between the chromophore and its surroundings in the S65T mutant 27

2.2. Biogenesis of the chromophore in wild-type GFP 28

2.3. Structure of the red fluorescent protein (DsRed) chromophore 30

2.4. Some synthetic analogs of chromophore in fluorescent proteins (FP).... 31

2.5. Thought process that led to the design of analogs **1** and the derivatives discussed in this chapter 32

2.6. Isomerization of *E*,*Z*-mixtures of compound **4c**........ 37

FIGURE Page

2.7. Trends in chemical shift differences for compounds **5** relative to compounds **1** illustrate that these can be used to differentiate between them 39

2.8. Top and side views of structures based on single molecule X-ray diffraction analyses of (a) compound **1c**; and (b) compound **5c** 41

2.9. Normalized UV absorbance and fluorescence of ***Z*-4c** and **1c** in MeOH, (*ca* 10^{-6} M for UV and 10^{-6} - 10^{-7} M for fluorescence) 49

2.10. UV absorbance and fluorescence for: (a) GFP-chromophore analogs **1** in MeOH; and, (b) GFP-chromophore analogs **5** in MeOH 51

2.11. Normalized UV absorbance and fluorescence of **1h** and the labeled streptavidin **1s** in 0.1 M lithium phosphate buffer, pH 7.4 54

3.1. Spectra of BODIPYs **9** in CH_2Cl_2 65

3.2. UV study of the self-condensation reaction in CH_2Cl_2 at 25 °C, and 6.7 x 10^{-5} M of pyrrole-2-CHO **8a** 67

3.3. (a) and (b) Two possible mechanisms for the formation of dipyrromethene 69

4.1. Structures of some common rhodamines 71

4.2. Syntheses of rosamine dyes via: (a) the conventional condensation approach; and, (b) nucleophilic additions to 3,6-diamino-xanthones 75

4.3. Reported syntheses of 3,6-diaminesubstituted-xanthones 76

4.4. X-ray structures for rosamines (a) **26** and (b) **27** 85

4.5. Normalized (a) UV-Vis absorption (10^{-6} M); and, (b) fluorescence spectra (10^{-7} M, excited at $\lambda_{\text{max abs}}$), for representative rosamines in CH_2Cl_2 91

4.6. Fluorescence intensity (excited at 562 nm) of **11i** at 1.4×10^{-7} M in a CH_2Cl_2/MeOH mixed solvent system 93

FIGURE Page

4.7. (a) Fluorescence spectra of **11i** in EtOH at different concentrations (excited at 564 nm); (b) effect of triton X-100 on the fluorescence of **11i** in EtOH (10^{-6} M, ex. 564nm) 94

4.8. Dependence of electronic spectra of **15h** on pH 96

4.9. UV-Vis absorption and fluorescence spectra of rosamine **25** and **25** labeled avidin in 0.1 M phosphate buffer, pH 7.4 98

4.10. Structure-activity relationship of rosamine derivatives in HL60 cells.... 102

4.11. Intracellular localization of compound **11d** in HSC2 cells 105

4.12. Compound **11d** induced apoptosis 106

4.13. Effects of **11d** on cell cycle 107

5.1. X-ray structure of novel pyronin **30c** 116

5.2. Normalized (a) UV absorbance and (b) fluorescence (excited at $\lambda_{\text{max abs}}$) of novel pyronins **30** in CH_2Cl_2 (*ca* 10^{-6} M for absorbance; 10^{-7} M for fluorescence) 118

5.3. Dependence of electronic spectra of **30c** on pH 119

5.4. Likely acid/base equilibrium for **30c** 122

6.1. (a) Through-space and (b) through-bond energy transfer cassettes 123

6.2. Through-bond energy transfer cassettes based on fluorescein and rosamines 124

6.3. Indication of the direction of the $S_1 \leftrightarrow S_0$ transition moments in different through-bond energy transfer cassettes 125

6.4. A water-soluble through-bond energy transfer cassette 127

6.5. Structures of through-bond energy transfer cassettes **31-33** 128

6.6. Donor fragments for water-soluble cassettes 129

6.7. BODIPY acceptor fragments for cassettes 132

FIGURE Page

6.8. Normalized (a) absorbance; and (b) fluorescence spectra of BODIPY donor **34** and illustrative acceptors **37** in MeOH (10^{-6} M for absorbance; 10^{-7} M for fluorescence) 141

6.9. Normalized (a) absorbance, (b) fluorescence in MeOH; (c) absorbance, (d) fluorescence in 0.1 M phosphoate buffer (pH 7.4) of BODIPY donor **34** and cassettes **31** 142

6.10. FRET donor Atto425, cassette **1** and cassette **2**: (a) structures; (b) absorbance and fluorescence spectra in 0.1 M phosphate buffer (pH 7.4) 145

6.11. FRET from one protein can be used to excite the donors of TBET-cassettes on another protein to reveal interactions 147

6.12. *In vitro* two protein interaction studies 149

6.13. *In vitro* three protein interactions 151

6.14. Studies of protein interactions in COS-7 cells 152

7.1. Highly fluorescent GFP-chromophore analog **1c** and its precursor **4c** ... 157

7.2. Novel pyronin derivatives 157

7.3. Illustrative examples of water-soluble BODIPY acceptor and through-bond energy transfer cassette 158

7.4. Scheme for detecting protein-protein interactions 159

7.5. Rosamine derivatives for anti-tumor activity study 161

7.6. Water-soluble through-bond energy transfer cassettes 162

LIST OF TABLES

TABLE Page

1.1. Important spectroscopic characteristics of fluorophores 3

1.2. Summary of protein labeling methods 19

2.1. Spectroscopic properties of *Z*- and *E*-**4c** isomers 35

2.2. Spectroscopic properties of **1c** and **5c** 38

2.3. Spectroscopic properties of **4c**, **1c** and **5c** in MeOH 49

2.4. Spectroscopic properties of GFP-chromophore analogs 53

3.1. Synthesis of symmetric BODIPYs **9** 63

3.2. Photophysical properties of BODIPYs **9** in CH_2Cl_2 64

4.1. Amination of the ditriflyl xanthone **16** 78

4.2. Synthesis of rosamine derivatives 80

4.3. Spectral properties of rosamines 89

4.4. Spectral properties of rosamine **11i** in different solvents 92

4.5. UV absorption and fluorescence maxima of **15h** in phosphate solutions 95

4.6. The structure-activity relationship (SAR) and *in vitro* cytotoxicity of rosamine analogues in HL60 cells 99

4.7. Cytotoxic effects of rhodamine analogues on carcinoma and immortalized normal human epithelial cell types 103

5.1. Synthesis of new pyronin dyes 115

5.2. Spectral properties of **30** measured in CH_2Cl_2 117

6.1. Summary of the yields for preparation of **47**, **48**, **37** and cassettes **31** 140

TABLE Page

6.2. Photophysical properties of BODIPY acceptors 141

6.3. Photophysical properties of BODIPY cassettes 144

7.1. Optical properties of cassette and its protein conjugates 159

LIST OF SCHEMES

SCHEME Page

2.1. Chemical synthesis of model chromophore of GFP 29

2.2. Routes to target compounds **1** 33

2.3. Revised procedure for synthesis of compounds **1f** and **5f** 42

2.4. Method to recycle the undesired compounds **5** into the target material **1** 44

2.5. Syntheses of the water-soluble probes **1g** and **1h** 46

2.6. Synthesis of water-soluble thiol-reactive analog **1i** 47

3.1. Typical syntheses of symmetric BODIPYs with meso-substituents from (a) acyl chloride; (b) anhydride; (c) aldehyde 57

3.2. (a) Two different methods for production of ketopyrroles from magnesium derivatives of pyrrole, and (b) application of these starting materials in the production of unsymmetrical BODIPYs 59

3.3. (a) Conventional synthesis of tetramethyl-BODIPY **9a**; and, (b) the new approach described here 61

3.4. Synthesis of pyrrole-2-aldehyde **8** 62

4.1. Synthesis of water-soluble rosamine **25** 83

5.1. Synthesis of 3,6-diamino-xanthones **17** and **29** 113

5.2. Failed attempt to synthesize new pyronin dyes 114

6.1. Synthesis of donor fragment **34** 130

6.2. Synthesis of donor **35** 131

6.3. Syntheses of pyrrole fragments **38-40** 133

6.4. Syntheses of iodo-benzaldehyde fragments **44-46** 135

SCHEME Page

6.5. Illustrative synthesis of BODIPY acceptor fragment **37a**...................... 136

6.6. Syntheses of acceptors **37h-i**.. 137

6.7. Illustrative synthesis of BODIPY cassettes.. 138

6.8. Synthesis of cassette **32**... 139

7.1. Illustrative syntheses of symmetric BODIPYs without meso-substituents 155

7.2. Syntheses of rosamine derivatives 156

7.3. New synthesis of rosamine derivatives 160

7.4. Synthesis of water-soluble rosamines 160

7.5. Synthesis of new water-soluble cassette **49** 163

7.6. Synthesis of cassette **50**... 164

7.7. Syntheses of NIR water-soluble BODIPY-BODIPY cassettes **51-52** 165

CHAPTER I

INTRODUCTION

1.1 Fluorescence

Fluorescence detection is highly sensitive and easy to operate. There has been dramatic growth in the use of fluorescence in the past few decades. Fluorescence technology has now been widely used by scientists from many different disciplines.

Fluorescence is the emission of light from any substance and occurs from electronically excited singlet states. Emission of light from triplet excited states is phosphorescence.[1] The processes that occur between the absorption and emission of light are usually illustrated by the Jablonski diagram (Figure 1.1.). Absorption of photons will excite a fluorophore into singlet excited states S_1 or S_2. This process is very quick and happens within 10^{-15}s. Fast internal conversion (10^{-12}s) brings the fluorophore to the lowest vibrational level of S_1. Emission of light around 10^{-8}s returns the molecules to the ground state. Molecules in the S_1 state can also undergo spin conversion to the first triplet state T_1. Transition from T_1 to S_1 is forbidden, thus emissions from T_1 are much slower (ms to s).

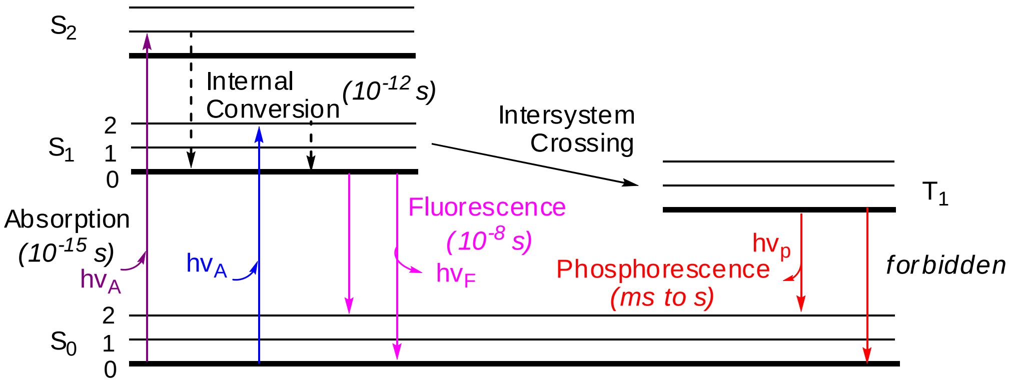

S2
Internal Conversion
(10-12 s)
Intersystem Crossing
S1
2
1
0
T1
Absorption
(10-15 s)
hvA
hvA
Fluorescence
(10-8 s)
hvF
hvp
Phosphorescence
(ms to s)
forbidden
S0
2
1
0

Figure 1.1. A typical Jablonski diagram.

This dissertation follows the style of *Journal of Organic Chemistry*.

Since the energy spacing between the vibrational levels in S_0 and S_1 is similar, there often exists mirror image symmetry between the emission spectrum and the $S_0 \rightarrow S_1$ absorption spectrum. Another general property of fluorescence is that the same fluorescence emission spectrum is generally observed regardless of the excitation wavelength.

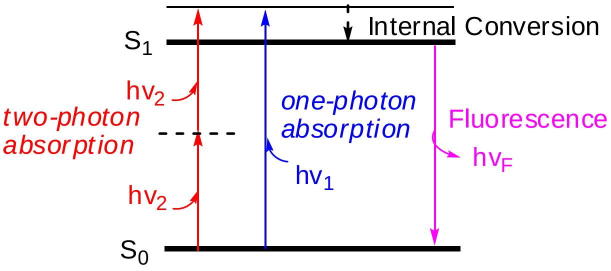


Figure 1.2. Jablonski diagram for two-photon excitation.

A fluorophore can also be excited by absorbing two photons simultaneously (Figure 1.2.). This process is two-photon excitation.[2] Irrespective of the excitation wavelength, the fluorophore still emits from the lowest energy level of the relaxed S_1 state. Two-photon absorption cross-section describes the probability of the simultaneous absorption of two photons and transition of the fluorophore to an excited state that differs energetically from the ground state by the energy of the two photons.[3] The physical origin of the two-photon cross-section can be understood by some simple comparison.[1] For one-photon absorption, the absorption coefficients are usually expressed as the molar extinction coefficients in units of $M^{-1}cm^{-1}$. The single molecule one photon absorption σ_1 can be described in units of cm^2, which is the effective area over which a single molecule absorbs the incident light. The number of photons absorbed per second (NA_1) is given by:

$$NA_1\ (\text{photon/s}) = \sigma_1\ (\text{cm}^2)\ I(\text{photon/cm}^2\text{s})$$

where I is the light intensity and σ_1 is the cross-section for one-photon absorption. The units are given within the parentheses. Similarly, for two-photon absorption:

$$NA_2\ (\text{photon/s}) = \sigma_2\ I^2(\text{photon/cm}^2\text{s})^2$$

In order for the units to match, the units of two-photon cross-section σ_2 must be cm^4s/photon. The values of σ_2 are usually reported in terms of GM (Gopper-Mayer) units, where 1 GM = 10^{-50} cm^4s/photon.

The important spectroscopic characteristics for fluorescent compounds are listed in Table 1.1. These parameters can be used as guidelines for selecting appropriate fluorophores for different applications.

Table 1.1. Important spectroscopic characteristics of fluorophores

terms	description
Molar extinction coefficient (ε: $M^{-1}cm^{-1}$)	A measurement of how strongly a fluorophore absorbs light at a given wavelength; absorbance divided by the absorption pathlength and the concentration of the fluorophore
Stokes shift (cm^{-1})	Difference (usually in frequency unit) between the spectral positions of the maxima of the lowest energy absorption and the fluorescence arising from the same electronic transition
Fluorescence quantum yield (Φ_f)	The number of emitted photons relative to the number of absorbed photons; typically determined relative to a fluorophore of known fluorescence quantum yield
Fwhm: nm	Fluorescence full width at half maximum height (fwhm): measurement of fluorescence sharpness
Brightness ($\varepsilon\Phi_f$)	Product of molar extinction coefficient multiplied by fluorescence quantum yield; a measurement for the intensity of the fluorescent signal obtained upon excitation at a specific wavelength
Fluorescence lifetime (τ: ns)	The average time the molecule spends in the excited state prior to return to the ground state; after a pulsed excitation, the time-dependent fluorescence intensity I(t) is measured. The lifetime is calculated from the slope of a plot of logI(t) versus t
TPA cross-section (σ_2: GM)	Two-photon absorption (TPA) cross-section: the probability of the simultaneous absorption of two photons and transition of the fluorophore to an electronically excited state
Photobleaching	Photochemical decompositioin of fluorophores; complex mechanisms
Blinking	The fluorophore randomly enter transient, non-fluorescent states; reversible
Fluorescence anisotropy	Measure for the polarization of the emitted light upon excitation with linearly polarized light; reflects the rotational freedom of fluorophores

1.2 Multiplexing with FRET and TBET

In biological studies especially high throughput screening, it is often desirable to simultaneously detect several fluorescently tagged components, *ie* multiplexing, using a single excitation source. Ideally, the fluorescent tags for multiplexing should emit very strongly at significantly different wavelengths when irradiated at the source wavelength so that the labeled tags can be easily detected and distinguished. However, multiplexing with one excitation source inevitably results in less intense fluorescence from the long wavelength emitting tags. This is because dyes that emit further into the red usually have red-shifted absorption and absorb less photons at the excitation wavelength (Figure 1.3.).

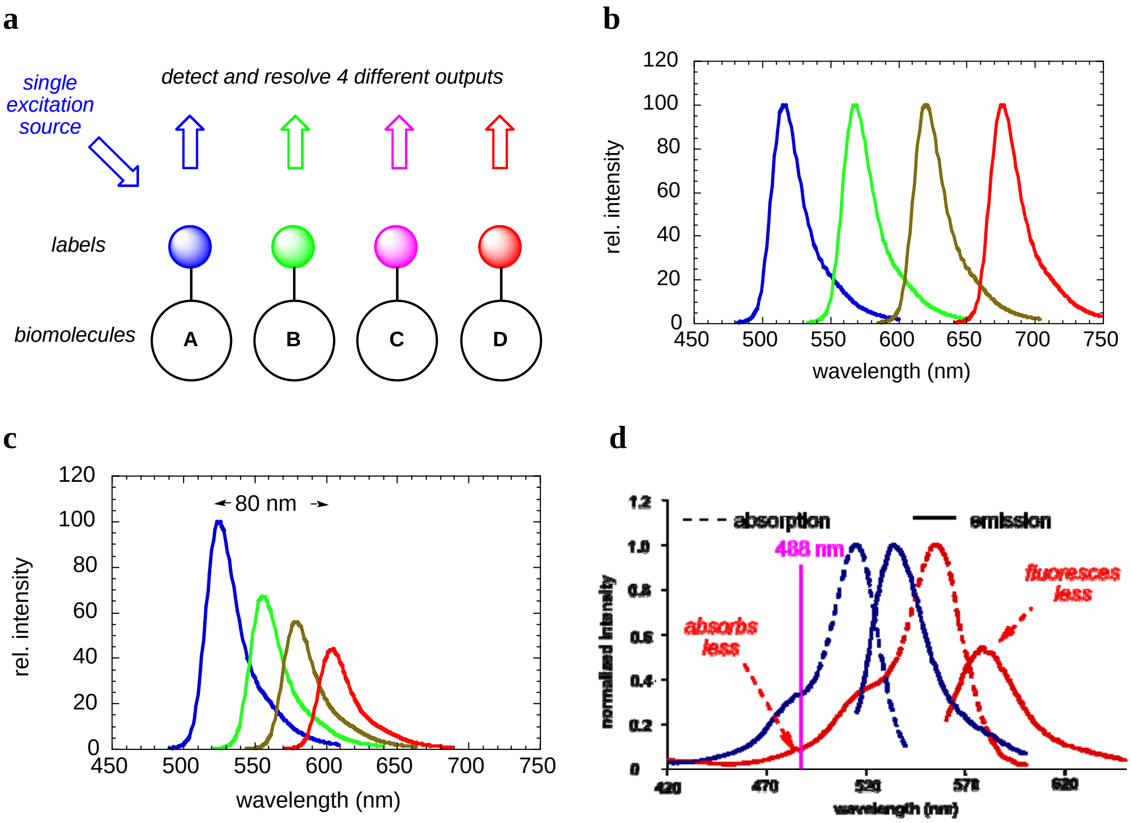


Figure 1.3. (**a**) Scheme for multiplexing; (**b**) ideal multiplexing output; (**c**) common multiplexing output; (**d**) reasons for the diminished fluorescence output as the emission maxima occur further from excitation source.

Combinations of dyes arranged to maximize fluorescence resonance energy transfer (FRET) have been used to alleviate the above mentioned problems.[4,5] FRET occurs between an excited donor (D) molecule and a ground state acceptor (A) as a result of long-range dipole-dipole interactions. The donor molecules typically emit at shorter wavelengths that overlap with the absorption spectrum of the acceptor. The rate of the energy transfer for a donor and acceptor separated by a distance *r* is given by[6,7]

$$k_T(r) = \frac{Q_D \kappa^2}{\tau_D r^6}\left[\frac{9000(\ln 10)}{128\pi^5 N n^4}\right] J(\lambda)$$

where Q_D is the quantum yield of donor in the absence of the acceptor. τ_D is the lifetime of donor in the absence of the acceptor. *r* is the distance between donor and acceptor. *N* is Avogadro's number (6.02×10^{23}). *n* is the refractive index of the medium. The overlap integral $J(\lambda)$ expresses the degree of spectral overlap between the donor emission and the acceptor absorption and can be calculated from

$$J(\lambda) = \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda$$

where $F_D(\lambda)$ is the normalized donor fluorescence intensity at wavelength λ in the corrected fluorescence spectrum. $\varepsilon_A(\lambda)$ is the extinction coefficient of the acceptor at wavelength λ.

κ^2 is a factor describing the relative orientation in space of the donor and acceptor transition dipoles. The orientation factor κ^2 can be given by

$$\kappa^2 = (\cos\theta_T - 3\cos\theta_D\cos\theta_A)^2 \qquad \textit{or}$$

$$\kappa^2 = (\sin\theta_T\sin\theta_A\cos\phi - 2\cos\theta_D\cos\theta_A)^2$$

where θ_T is the angle between the emission transition dipole of the donor and the acceptor absorption dipole, θ_D and θ_A are the angles between these dipoles and the vector joining the donor and the acceptor, and ϕ is the angle between the planes (Figure 1.4.). κ^2 can range from 0 to 4. For head to tail parallel dipoles $\kappa^2 = 4$, and for parallel dipoles $\kappa^2 = 1$. If the dipoles are perpendicular to each other, $\kappa^2 = 0$. κ^2 is generally assumed to be 2/3, which is appropriate for dynamic random averaging of the donor and acceptor.

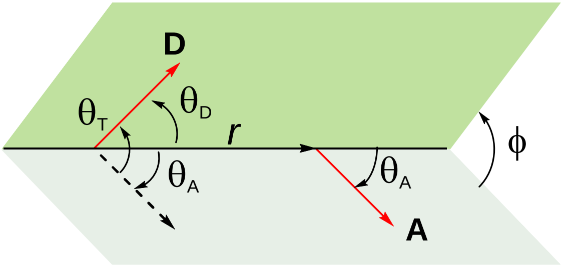


Figure 1.4. Dependence of the orientation factor κ^2 on the direction of the dipoles.

The Förster distance (R_0) is the distance at which half the donor molecules decay by energy transfer and half decay by usual radiative and non-radiative pathways. So when $r = R_0$, $K_T = (1/\tau_D)$. Thus

$$R_0^{\ 6} = \left[\frac{9000(\ln 10)Q_D k^2}{128\pi^5 N n^4}\right]\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda$$

R_0 is typically in the range of 20 to 60 A° for organic fluorophores. R_0 is usually reported for an assumed κ^2 value (2/3). Once the R_0 is known, the rate of energy transfer can be easily calculated using

$$\mathrm{k}_T = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6$$

The efficiency of energy transfer E is the fraction of photons absorbed by the donor that are transferred to the acceptor. E is given by

$$E = \frac{\mathrm{k}_T}{\tau_D^{-1} + \mathrm{k}_T}$$

which is the ratio of the energy transfer rate to the total decay rate of the donor. E can also be expressed as:

$$E = \frac{R_0^{\ 6}}{R_0^{\ 6} + r^6} = 1 - \frac{\tau_{DA}}{\tau_D} = 1 - \frac{F_{DA}}{F_D}$$

The efficiency of energy transfer is typically measured using the relative fluorescence intensity or fluorescence lifetime of the donor, in the absence and presence of acceptor.

Since energy transfer is governed by the spectra overlap and dipole orientations, FRET systems are only partial solutions to the multiplexing problems. On the other hand, through-bond energy transfer (TBET) cassettes (Figure 1.5.), in which donor and acceptor systems are connected via a twisted but conjugated linker, might solve all these problems. In such cassettes, rapid energy transfer from the donor to the acceptor may occur though bonds. Extensive studies by our group and others showed that the through-bond energy transfer is mechanistically different to the Förster basis for FRET, and there is no requirement for overlap of the emission of the donor part with the absorption of the acceptor fragment.[8-11] It therefore is possible to design dyes that absorb strongly at a short wavelength and emit brightly with similar intensities at several long wavelengths which are critical for multiplexing.

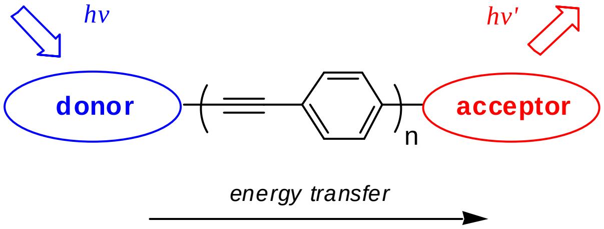


Figure 1.5. A typical diagram for through-bond energy transfer cassette.

The exact mechanism for through-bond energy transfer is still not clear. But it is thought to occur via Dexter energy transfer or conjugated linker mediated superexchange. Dexter energy transfer is an electron-exchange interaction between donor and acceptor.[12] In dexter energy transfer, the excited electron from the donor is transferred to the acceptor with concomitant transfer of a ground state electron from the acceptor to the donor. Dexter energy transfer depends on the orbital energy levels and symmetries of the donor relative to those of the acceptor. It is a short-range process compared with FRET. Superexchange is a stepwise energy transfer process in which the electron is transferred through the conjugated bridge, thus superexchange takes place over longer distances than Dexter energy transfer.[13]

Even though they are referred to as "through-bond energy transfer cassettes", the energy transfer process in such systems is usually the combination of FRET, through-bond and other transfer mechanisms. It is not possible to determine how much energy is transferred via FRET or through-bond. The observed energy transfer efficiency is the overall transfer via all possible pathways. Our group is interested in such energy transfer cassettes for biological applications.

1.3 Labeling of proteins

Labeling of proteins with fluorescent probes has facilitated the studies of protein structures, functions, intracellular locations, dynamics and protein-protein interactions. Traditionally this has been done by *in vitro* labeling of proteins with fluorescent probes, followed by transferring them back into cells and monitoring them in real time using advanced imaging techniques such as confocal microscopy and others. Recent advances in genetic engineering have made it possible to directly generate fluorescently labeled proteins in living cells. A brief summary of the available protein labeling methods is described in this section.

1.3.1 Covalent labeling techniques

(1) Direct fluorescent labeling of target proteins

The widely used labeling strategies rely on selective reaction of fluorescent probes with amino acid side chains in the proteins of interest (Figure 1.6.). For instance, the primary amine in lysine sidechains are often targeted by fluorescent N-hydroxysuccinimidyl ester, isocyanates or aldehydes to give amides, ureas or reductive amination products. Carboxylic acid groups from Asp or Glu side chain can also be used to react with amino-fluorophores. Similarly, fluorescent maleimides, iodoacetamides or disulfides are used to react with cysteine sidechains to afford the corresponding bioconjugates.[14] Recent development expands the labeling scope to different amino acid side chains such as tyrosine, tryptophan and so on by choosing appropriate reactions and carefully controlling the reaction conditions (Figure 1.6.).[15] However, these methods are not site-specific and only good for *in vitro* labeling, because the same amino acid can appear at different locations in the target protein and it will also be present in many other proteins.

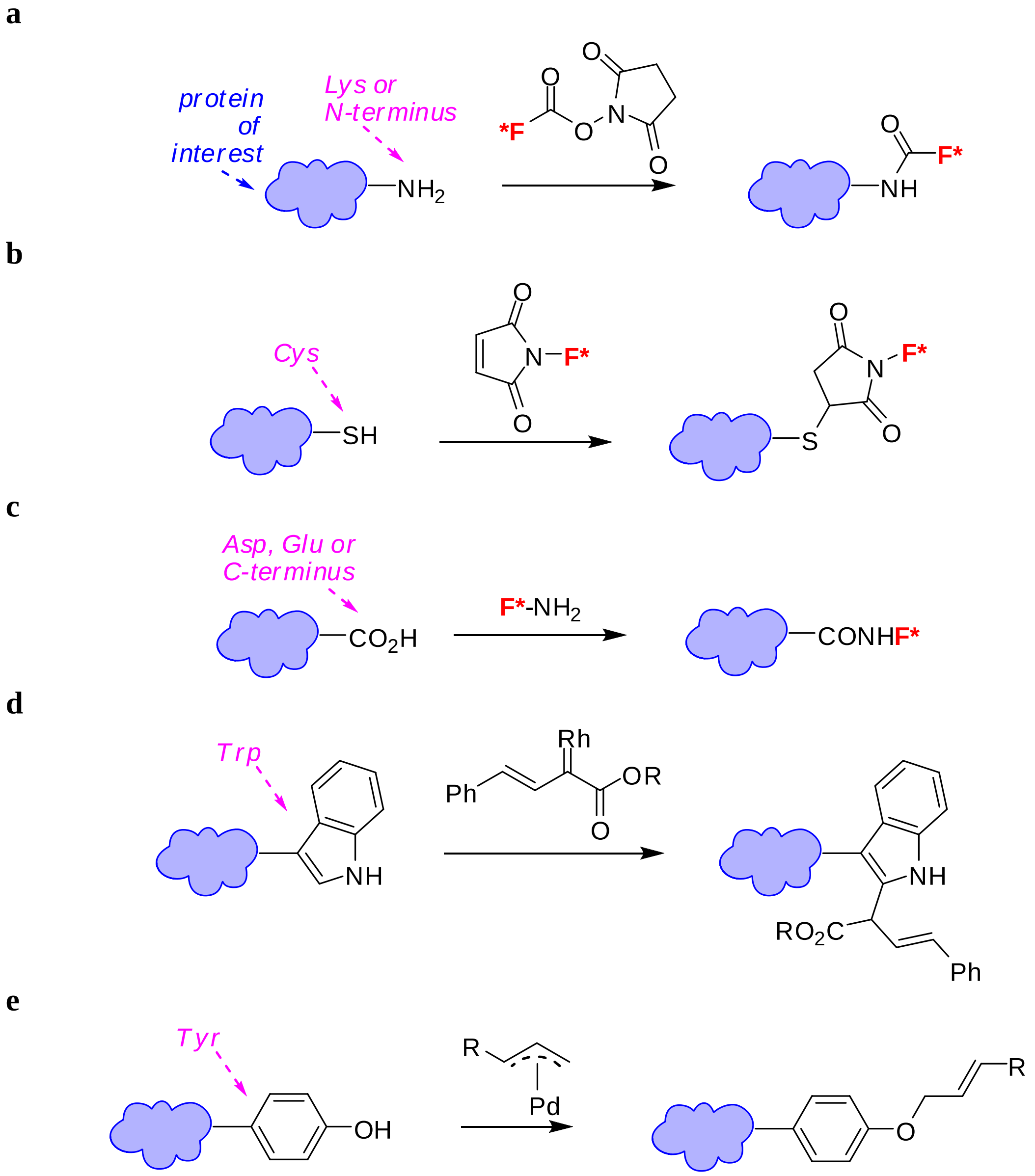

a
protein of interest
Lys or N-terminus
NH2
*F
O
N
O
O
O
NH
F*
b
Cys
SH
O
N
F*
O
O
N
F*
S
O
c
Asp, Glu or C-terminus
CO2H
F*-NH2
CONHF*
d
Trp
NH
Rh
Ph
OR
O
NH
RO2C
Ph
e
Tyr
OH
R
Pd
O
R

Figure 1.6. Labeling of proteins by modification of amino acid sidechains. Modification of **a** lysines; **b** cysteines; **c** aspartic acids or glutamic acids; **d** tryptophans; **e** tyrosines.

Selective modification of the terminal amino acid can be used to alleviate the problems (Figure 1.7.). For instance, N-terminal tryptophan residues can be modified by aldehydes through Pictet-Spengler reaction.[16] Oxidation of N-terminal serine and threonine residues can form reactive aldehydes/ketones allowing further selective derivatization.[17] The N-terminal amino group can also be oxidized by pyridoxal-5-phosphate (PLP) to form an imine, tautomerization followed by hydrolysis of the resulting glyoxyl imine to yield an aldehyde or ketone specifically at the N terminus for further functionalization.[18] Since endogenous N-terminal cysteine-containing proteins are rare, protein ligation has been used for site-specific labeling of proteins.[19] The protein of interest with an N-terminal cysteine is expressed inside a live cell by using intein-mediated protein splicing, methionyl aminopeptidases, the ubiquitin-fusion strategy or selective proteolysis with highly specific proteases like the TEV (tobacco etch virus) protease. Incubation of the cell with a thioester-containing, cell-permeable, fluorescent probe leads to chemoselective native chemical ligation reaction between the thioester of the probe and the N-terminal cysteine of the protein, giving rise to the resulting specifically labeled protein. Muir et al successfully adopted intein-mediated splicing to selectively label proteins by expressing a target protein with the first half of the naturally occurring Ssp DnaE split intein fused to its C-terminus. The second half of the split intein was covalently tagged by a fluorescent probe. Upon combining the split intein halves, self-splicing occurred, leaving the protein of interest labeled at its C-terminus with the probe (Figure 1.8.a).[20] Although modification of terminal amino acids are more selective than labeling of amino acid sidechains, these labeling protocols still lack generality as it is difficult to achieve selective labeling of the protein of interest in protein mixtures, because several proteins might have the same terminal amino acid.

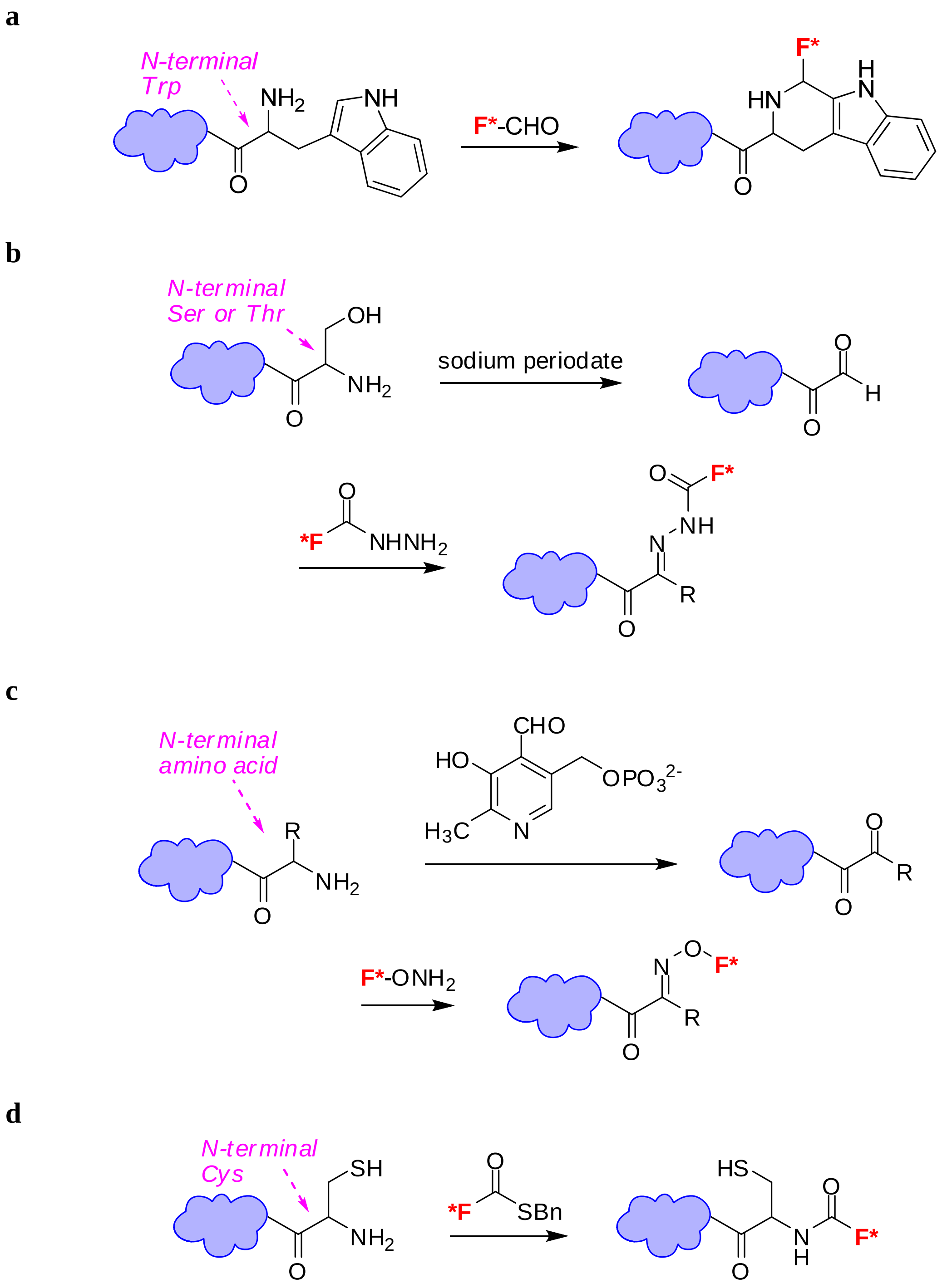

a
N-terminal Trp
NH2
NH
O
F*-CHO
F*
H
N
HN
O
b
N-terminal Ser or Thr
OH
NH2
O
sodium periodate
O
H
O
O
*F
NHNH2
O
F*
NH
N
R
O
c
N-terminal amino acid
R
NH2
O
CHO
HO
OPO3 2-
H3C
N
O
R
O
F*-ONH2
N
O
F*
R
O
d
N-terminal Cys
SH
NH2
O
O
*F
SBn
HS
O
N
H
F*
O

Figure 1.7. Labeling of proteins by modification of terminal amino acids. Modifcations of: **a** tryptophan; **b** serine or threonine; **c** N-terminal amino acids; **d** cysteine.

More specific labeling methods are usually achieved by genetic incorporation of unnatural functional groups into the proteins of interest (Figure 1.8.b). The first demonstration of this concept was by Dibowski and Schimdtchen in 1998 through introducing a *p*-iodophenylalanine residue into an undecapeptide using reverse proteolysis.[21] This residue was modified through metal catalyzed coupling reactions with different probes. Important extensions of this approach were made by Schultz *et al.*[22] They have successfully incorporated various unnatural amino acids with different functional groups into proteins with high efficiency and site-specificity. For instance, unnatural amino acids bearing ketones, azide or alkynes have been incorporated, allowing further derivatization using hydrazone formation, Staudinger ligation or azide/alkyne cycloaddition reactions. The major advantages of unnatural amino acid labeling are the excellent specificity, versatility and minimal structural perturbation to the target protein. However, unnatural amino acid mutagenesis is not yet broadly applicable in all cell types.

Recently, a ligand-directed protein labeling method has been reported by Hamachi and co-workers.[23] In this approach, the ligand binds to the target protein with high affinity, driving a chemical reaction of the reactive group on the ligand with an amino acid located on the protein surface through the proximity effect (Figure 1.8.c). The reactive group used by Hamachi is an electrophilic phenylsulfonate ester group. The surface of the target protein is specifically labeled using tosyl chemistry: an S_N2-type reaction with the concomitant release of the ligand molecule. This approach was successfully used to label proteins in living cells, tissues and mice. However, a high affinity ligand for the protein of interest needs to be identified in order to apply this labeling method, which is usually not convenient especially for the unknown proteins.

The most prominent method of protein labeling is to genetically encode green fluorescent protein (GFP) or one of its variants as a fusion to the protein of interest[24] (Figure 1.8.d). The resulting fusion gene is expressed and the fluorescence of GFP is detected microscopically.[25] Although GFP and its variants have proven to be extremely useful for *in vivo* studies, these probes are limited by their large size and thus may interfere with the structure and function of the target proteins to which they are fused.

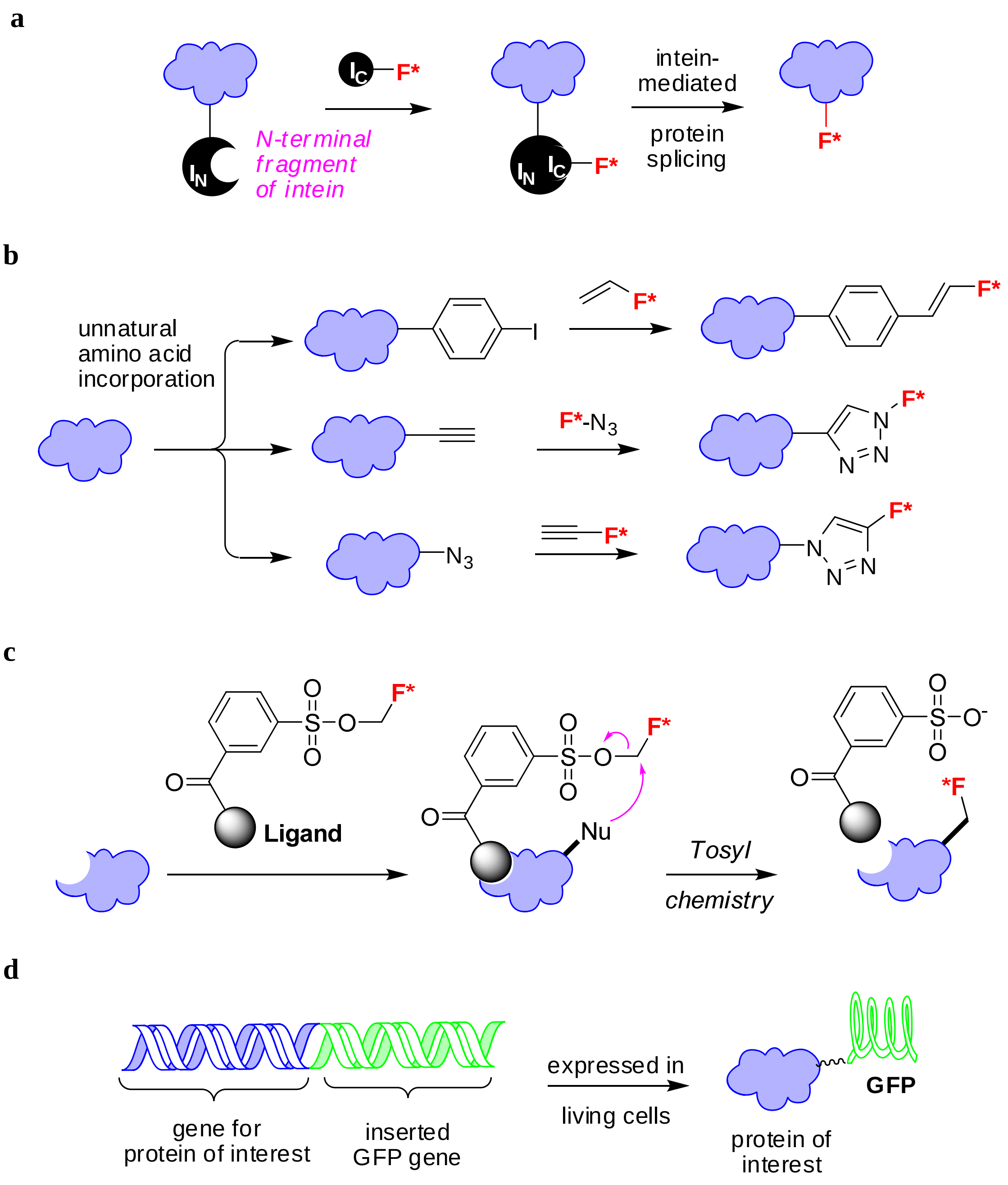

a
I_C
F*
N-terminal
fragment
of intein
I_N
I_N I_C
F*
intein-
mediated
protein
splicing
F*
b
unnatural
amino acid
incorporation
I
F*
F*
F*-N_3
N
N
N
F*
N_3
F*
N
N
N
F*
c
O
S
O
O
F*
O
Ligand
O
S
O
O
F*
O
Nu
Tosyl
chemistry
O
S
O
O$^-$
O
*F
d
gene for
protein of interest
inserted
GFP gene
expressed in
living cells
GFP
protein of
interest

Figure 1.8. Other direct labeling methods. **a** split intein labeling (I_N, N-terminal portion of a split intein; I_c, C-terminal fragment); **b** functionalization of incorporated unnatural amino acids; **c** ligand-directed labeling; **d** labeling with GFP.

(2) Enzyme-catalyzed labeling by post-translational modification

More recent approaches have employed transferase and ligases for covalent labeling of target proteins (Figure 1.9.). Recongnition between the enzyme and its protein substrate ensures that the labeling reaction is fast and specific.

Acyl carrier protein (ACP) and peptide carrier protein (PCP) have been used to specifically label proteins of interest by fusion to the target proteins followed by modification by the enzyme phosphopantetheine transferase (PPtase).[26,27] PPtase transfers 4′-phosphopantetheine-linked probes from coenzyme A (CoA) to a serine sidechain of ACP or PCP to form a specific covalent attachment.

Biotin ligase has also been exploited as an alternative strategy for covalent protein labeling. A 15 amino acid peptide sequence (‘acceptor peptide’, or AP) was identified by Ting for specific biotinylation of proteins.[28,29] The biotin ligase catalyzed the ligation of biotin to the AP which can be fused to proteins of interest. A ketone isostere of biotin was also transferred to the AP, providing a unique functionality for subsequent derivatization with different probes.

Transglutaminases (TGases) are enzymes that catalyze amide bond formation between the γ-carboxyamide group of an acceptor glutamine (Gln or Q) residue and a primary amine donor, with the loss of ammonia.[30] TGase exhibits high specificity for its glutamine-containing protein substrate, but wide tolerance for the structure of the amine donor, which makes it ideal for protein labeling applications. Recently, Ting and co-workers have utilized a seven-amino-acid tag (PKPQQFM, ‘Q tag’) that is substrate for a TGase.[31] Target proteins were expressed as fusions with Q tag then site-specifically labeled by TGase with amine-containing probes.

Several other enzymes have also been utilized to achieve specific labeling of proteins by post-translational modification, such as farnesyltransferase (FTase)[32] and formylglycine-generating enzyme (FGE).[33] FTase catalyzes the covalent attachment of modified farnesyl moieties through a thioether bond to a cystein residue in the C-terminal four-aminoacid farnesylation motif “CAAX”, where “C” is a cysteine residue, “A” is an aliphatic residue, and “X” is usually serine, methionine, glutamine, alanine or threonine. Farnesyl analogs with azide/alkyne functional groups were successfully incorporated into proteins with high specificity. The products were subsequently labeled with different probes. Formylglycine generating enzyme can generate the sulfatase active site by converting a conserved cysteine to formyl glycine. A six-residue sulfatase motif (LCTPSR) containing this cystein has been used as a genetic tag to introduce the aldehyde functionality into proteins for subsequent labeling with aminooxy- or hydrazide-functionalized probes.[34]

Most of the above mentioned methods are restricted to the cell surface, because of the use of membrane impermeable dyes or competition from endogenous substrates. A covalent system that can be used intracellulary relies on the self-alkylation reaction of human O^6-alkylguanine transferase (hAGT or SNAP-tag) to label hAGT fusion proteins with a number of cell permeable, fluorescent O^6-benzylguanine (BG) substrates.[35-37] The irreversible self-labeling reaction involves transfer of the functionalized benzyl groups from the BG substrates to an active site cysteine in hAGT to form a stable covalently modified protein. BG can be derivatized with a wide range of probes without influencing the efficiency of the reaction. Labeling of SNAP-tag fusion proteins in different cellular compartments such as nucleus, cytoplasm, endoplasmic reticulum (ER), cytoplasmic and extracellular face of the plasma membrane have been successfully illustrated.

Another self-labeling protein tag (HaloTag) has also been developed for specific protein labeling.[38] The engineered halogen dehydrogenase cleaves the carbon halogen bond of fluorescent ligands to become covalently labeled with the fluorophore. Covalent bond formation between HaloTag and the fluorescent alkylhalides is highly specific, occurs rapidly under physiological conditions and essentially irreversible. This system was successfully used to specifically label HaloTag fusion proteins *in vitro* and in living cells.

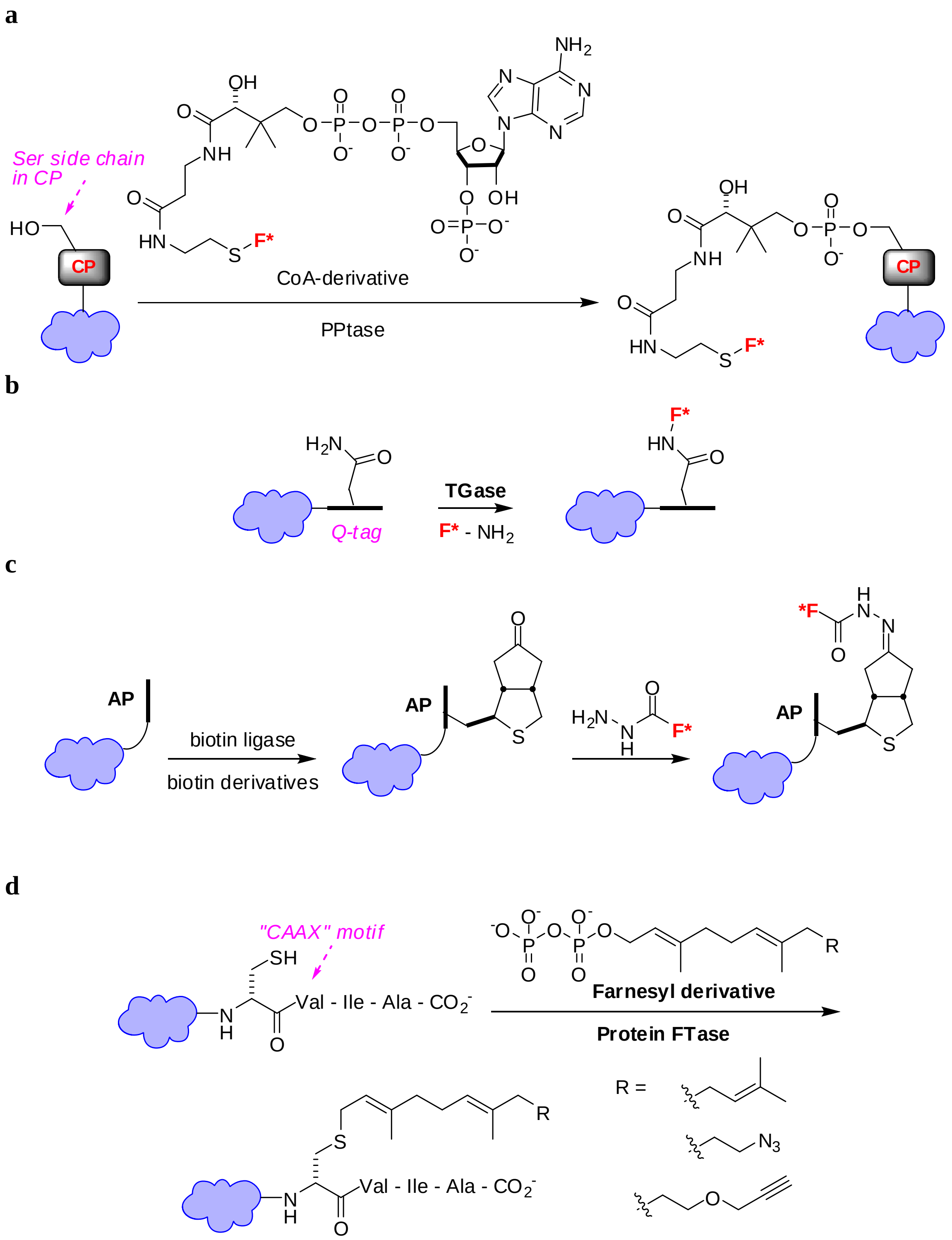

a
Ser side chain in CP
HO
CP
CoA-derivative
PPtase
F*
b
H2N
O
Q-tag
TGase
F* - NH2
HN
c
AP
biotin ligase
biotin derivatives
S
H2N
N
H
*F
d
"CAAX" motif
SH
Val - Ile - Ala - CO2-
Farnesyl derivative
Protein FTase
R =
N3

Figure 1.9. Illustrative examples of enzyme catalyzed labeling of proteins. **a** ACP (PCP)/PPtase; **b** Q-tag/TGase; **c** AP/biotin ligase; **d** farnesyl/FTase; **e** BG/AGT (SNAP-tag); **f** Halo-tag.

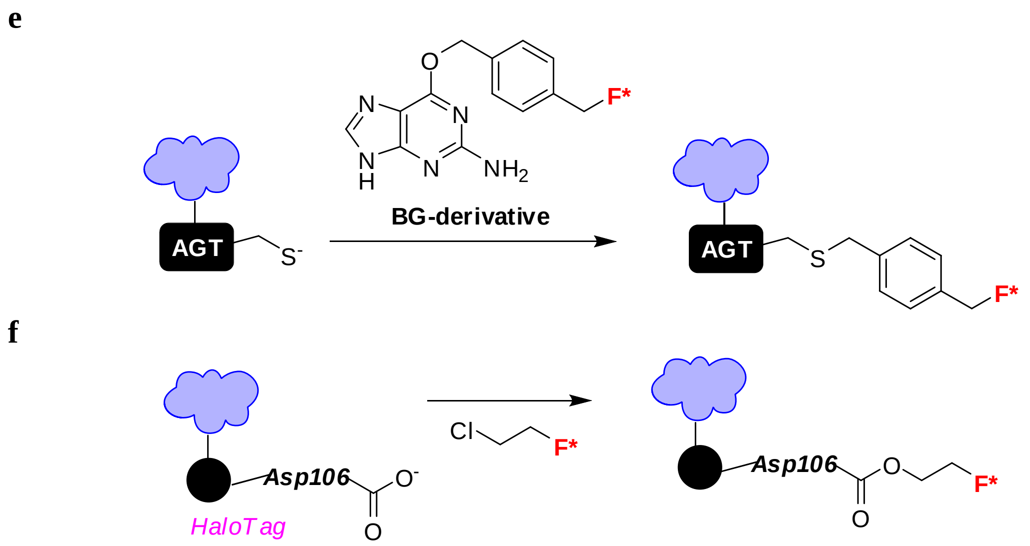


Figure 1.9. Continued.

1.3.2 Non-covalent labeling methods

The most straightforward method for targeting a small molecule tag to a protein sequence is to use a specific, high-affinity non-covalent binding interaction between a ligand and its corresponding receptor (Figure 1.10.). In this approach, the protein of interest is expressed as fusion to a protein tag that is capable of binding to a small-molecule ligand.

One of the first ligand-receptor pairs for *in vivo* protein labeling was reported by Farinas and Verkman, who used antibody tags fused to localization signal sequences to target various hapten-fluorophore conjugates to specific subcellular compartments in living cells.[39] Similar strategy was exploited to label avidin fusion proteins with fluorescent biotin derivatives using the biotin-avidin interactions.[40] More recent approaches have employed the interaction between an FKBP12 mutant (F36V) and a synthetic ligand (SLF').[41] The specific interaction between dihydrofolate reductase (DHFR) and its small molecule-ligand methotrexate was also used to specifically label DHFR fused proteins with fluorescently labeled methotrexates.[40,42,43]

The tetracysteine/biarsenical method introduced by Tsien and coworkers is one of the most used strategies to target chemical labels to specific proteins in living cells.[44]

The membrane-permeant fluorogenic biarsenical (FlAsH) compounds become strongly fluorescent upon high-affinity binding (K_d : 2-4 pM) to the six amino acid tetracysteine motif (CCXXCC, where X is any amino acid except cysteine) in the target protein.[45] The FlAsH method has already made significant contributions to cell biology.

In addition to the FlAsH tag, several groups have developed conceptually similar approaches. For example, Nolan has evolved a 38 amino acid peptide ('fluorette') that binds to Texas Red fluorophore with a K_d of 25 pM.[46,47] Labeling of this sequence fused protein in cells with Texas Red derivatives was demonstrated. In a separate approach, Vogel used nickel-nitrilotriacetic acid (Ni-NTA) functionalized fluorophores to label hexahistidine tagged cell-surface proteins.[48] Another system consisting of a polyaspartate tag for labeling of cell-surface proteins with fluorescent zinc conjugates has also been reported.[49]

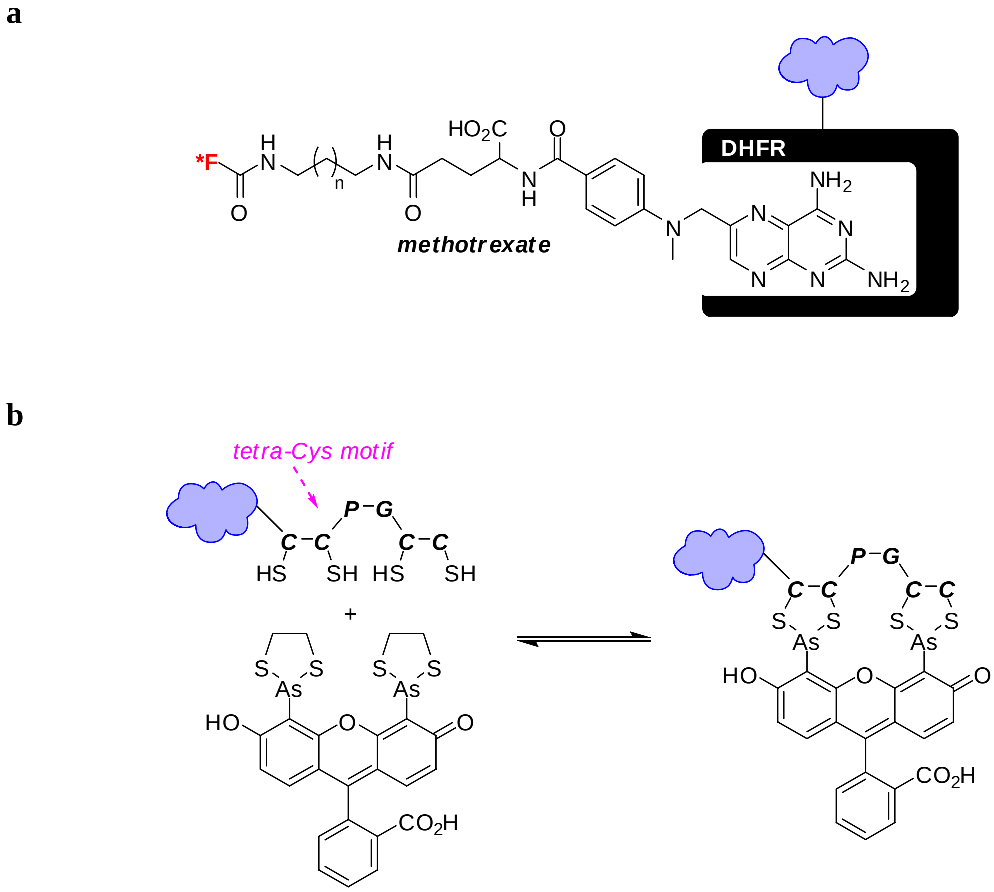

a
*F
H
N
n
H
N
O
O
HO2C
O
N
H
methotrexate
N
DHFR
NH2
N
N
N
N
N
NH2
b
tetra-Cys motif
P-G
C-C
C-C
HS
SH
HS
SH
+
S
S
As
S
S
As
HO
O
O
CO2H
P-G
C-C
C-C
S
S
As
S
S
As
HO
O
O
CO2H

Figure 1.10. Illustrative examples of non-covalent labeling of proteins. **a** DHFR/methotrexate; **b** tetracysteine/biarsenical (FlAsH).

Table 1.2. Summary of protein labeling methods

labeling method	tag size (aa)	comments
Covalently labeling		
(1) direct attachment to proteins		
modification of amino acid side chains (Lys, Cys, Glu, Asp, Trp, Tyr)	0	selective but not specific, only for *in vitro* labeling
terminal amino acids modification (Ser, Thr, Cys, Trp)	0	specific but only for *in vitro* labeling
unnatural amino acids	0	excellent specificity, not yet broadly applicable in all cell types
intein mediated protein splicing	~ 150 (temporary)	specific, tag size problem even though temporary
ligand-directed labeling	0	high affinity ligand is needed
GFP	238	excellent specificity, tags are big
(2) enzyme catalyzed labeling		
ACP (PCP)/PPTase	77 (80)	specific but tags are big and only for cell surface protein labeling
AP/biotin ligase	12	cell surface protein labeling
Q-tag/TGase	6	moderate specificity, background labeling
CAAX/FTase	4	specificity needs improvement
Sulfatase motif/FGE	6	specificity needs improvement
BG/hAGT	182	specific and intracellular labeling, but tag is big
Alkylhalide/HaloTag	296	tag size problem
Non-covalently labeling		
Hapten (phOx)/antibody (sFv)	25 KDa	tag size problem
biotin/avidin	516	tag is big
SLF'/FKBP12 (FK36V)	108	protein-ligand dissociation, tag size problem
Methotrexate/DHFR	157	protein-ligand dissociation, tag size problem
Biarsenicals/tetracystein	4	background labeling, only for reducing environment
Texas Red/fluorette	38	Texas Red only
Ni-NTA/His_6	6	low affinity
Zn labels/polyaspartate	12	low affinity

1.3.3 Summary

In summary, a diverse set of protein labeling systems are available (Table 1.2.). As each labeling method has its advantages and restrictions, successful experiments require selection of a labeling approach that is matched with the biology of the research question. The approaches reported to date point the way toward effective design criteria for site-specific labeling of proteins in living systems.[40] Receptors must be amenable to genetic encoding as fusions to the protein of interest. The receptors must be small so as not to perturb protein functions. Probes should bind (or be covalently attached to) receptors with high specificity and stability. The kinetics of the binding or attachment should be fast enough to facilitate the applications to study fast biological processes.

1.4 Delivery of proteins into cells

The *in vitro* labeled proteins generally can not enter cells by passive diffusion, but require active transport. Cell-penetrating peptides (CPPs) are very attractive for this purpose because of their ability to mediate cellular uptake of proteins. The most widely used CPPs are derived from HIV-1 tat (TAT) and *Drosophila* Antennapedia homoprotein.[50,51] However, the delivery with these CPPs requires covalent attachment of proteins of interest to the CPPs thus the processes are experimentally inconvenient, time consuming and restrictive with respect to the scope of the experiments that can be performed. Conversely, carrier vehicles that can import proteins without being covalently attached have the potential to circumvent all these problems. The most common “Non-covalent carriers” include “Pep-1” (also known as Chariot)[52,53], oligo-arginine[54,55] and polyethyleneimine (PEI)[56]. Pep-1 and oligo-arginine have been used in our group to import proteins into cells.

Pep-1 is a synthetic peptide carrier capable of introducing proteins into cells. The peptide has 21 amino acid residues (KETWWETWWTEWSQPKKKPKV), which can be divided into three different domains with specific functions: (i) a hydrophobic tryptophan-rich motif, forming the main interactions with macromolecules and required for efficient targeting to the cell membrane; (ii) a hydrophilic lysine-rich domain derived from the nuclear localization sequence (NLS) of simian virus 40 (SV-40) large T antigen, required to improve intracellular delivery and solubility of the peptide vector; and, (iii) a spacer domain, which improves the flexibility and the integrity of both the hydrophobic and hydrophilic domains.[57]

Pep-1 associates with the protein cargo via its tryptophan-rich motif through noncovalent interactions to form the Pep-1/cargo complex. The cellular translocation mechanism of the Pep-1/cargo complexes has become very controversial. Deshayes *et al.* postulated that the membrane crossing process involved formation of a transient transmembrane pore-like structure.[57] In contrast, Henriques *et al.* found no evidence for pores or other lytic structures using photophysical techniques. Their work suggested that the main driving force for the peptide translocation was charge asymmetry between the outer and inner leaflet of biological membranes.[58] Zerbe *et al.* proposed a model which was compatible both with the active endocytic mechanism and with the passive inverted micelle pathway.[59]

The Pep-1 peptide is non-toxic and has high transport efficiency (65-95%). It has been shown to rapidly co-transport large fluorescent proteins (>119kDa, and mAb's *ca*. 250 kDa) into cells. Once internalized, the fluorescent protein/Pep-1 complex rapidly dissociates, thereby allowing the fluorescent-tagged protein to proceed to its intracellular target while concomitantly the Chariot peptide is rapidly degraded.[52,53]

1.5 Detection of protein-protein interactions

Protein-protein interactions (PPIs) are known to play an important role in all cellular functions, and thus technologies for identifying and characterizing PPIs are extremely important. PPIs are commonly detected by two-hybrid systems, reconstitution methods, immunoprecipitations, protein mass spectrometry and more recently FRET-based systems.[60,61]

The two-hybrid system is a method that uses transcriptional activity as a measure of protein-protein interactions.[62,63] The system relies on the modular organization of many transcription factors which have separable domains for DNA-binding and transcriptional activation. One of the target proteins (often termed the "bait") is fused to DNA-binding domain (DBD). The other protein of interest (often termed the "prey") is fused to the activation domain (AD) of the transcription factor. Interactions between the "prey" and "bait" bring the AD close to the DBD resulting in activation of transcription then expression of a reporter gene.

Reconstitution of a split protein's function can also be used to detect PPIs. For example, the N terminal fragment and C-terminal fragment of β-galactosidase were fused to interacting protein partners. Upon interaction, β-galactosidase was reconstituted and detected histochemically.[64] Interaction-mediated reconstitution has also been described for ubiquitin, mouse dihydrofolate reductase (DHFR) and so on.[60]

Immunoprecipitation is another commonly used method to detect PPIs.[61] In this approach, an antibody that is specific for one of the target proteins is used to precipitate that particular protein out of a solution. The interacting partners will also be pulled out of the solution as intact protein complexes which can be further identified by mass spectrometry or other methods.

Fluorescence technologies have been used extensively for assaying PPIs in living cells and in real time since they are highly sensitive even at very low protein concentrations and easy to operate. Fluorescence assays for protein interaction have usually been designed based on relative changes in the photophysical parameters such as anisotropy, lifetime, emission wavelength and quantum yield of the fluorescence emission upon interaction.

Intermolecular FRET between two spectrally overlapping fluorophores attached to two different host proteins offers a unique opportunity to monitor real-time protein-protein interactions. This is usually performed with a pair of fluorescent proteins used for FRET such as BFP-GFP and CFP-YFP.[65] Several approaches, such as mathematical correction, internal calibration, external calibration, donor lifetime measurement and acceptor bleaching, have been employed to quantify of the extent of FRET in the interacting protein mixtures.

The strong spectral overlap of the donor and acceptor that is required for FRET leads to a substantial fluorescence background (*eg* spectral bleedthrough or cross-talk) that interferes with the detection of the FRET signals and makes the quantification even more complicated. Moreover, the spectral overlap requirement makes it extremely difficult to detect multiple PPIs simultaneously because it is hard to find several acceptors that have strong spectral overlap with the same donor but emit at significantly different wavelengths.

The through-bond energy transfer cassettes can strongly fluoresce at a much longer wavelength than the excitation source, effectively giving a much larger Stoke's shifts, and more intense fluorescence from a single excitation source than is possible via FRET. Therefore, the through-bond energy transfer cassettes have the potential to increase both the resolution and fluorescence intensities obtained from several tags excited by the same laser source operating at a single wavelength. Sets of energy-transfer cassettes can then be used to monitor communication between proteins in cells. Combination of FRET with through-bond energy transfer cassettes enables several contacts to be detected using one laser excitation source; several through-bond energy transfer cassettes can be observed simultaneously due to their identical absorption maxima, high quantum yields and huge (different) Stoke's shifts. As showed in Figure 1.11., target protein **1** is labeled with a FRET donor and the interacting partner proteins are attached with different through-bond energy transfer cassettes. After excitation of the FRET donor on protein **1**, it will be close enough for the FRET donor to transfer energy through space to the FRET acceptors, which are also donors in the through-bond energy transfer cassettes, on the interacting partner(s), then the energy will be transferred again through bond to the cassette acceptor and the fluorescence of that acceptor is going to be detected. So the interacting partner(s) can be determined simply by the fluorescence detected. This method also enables multiple protein-protein interactions to be observed simultaneously.

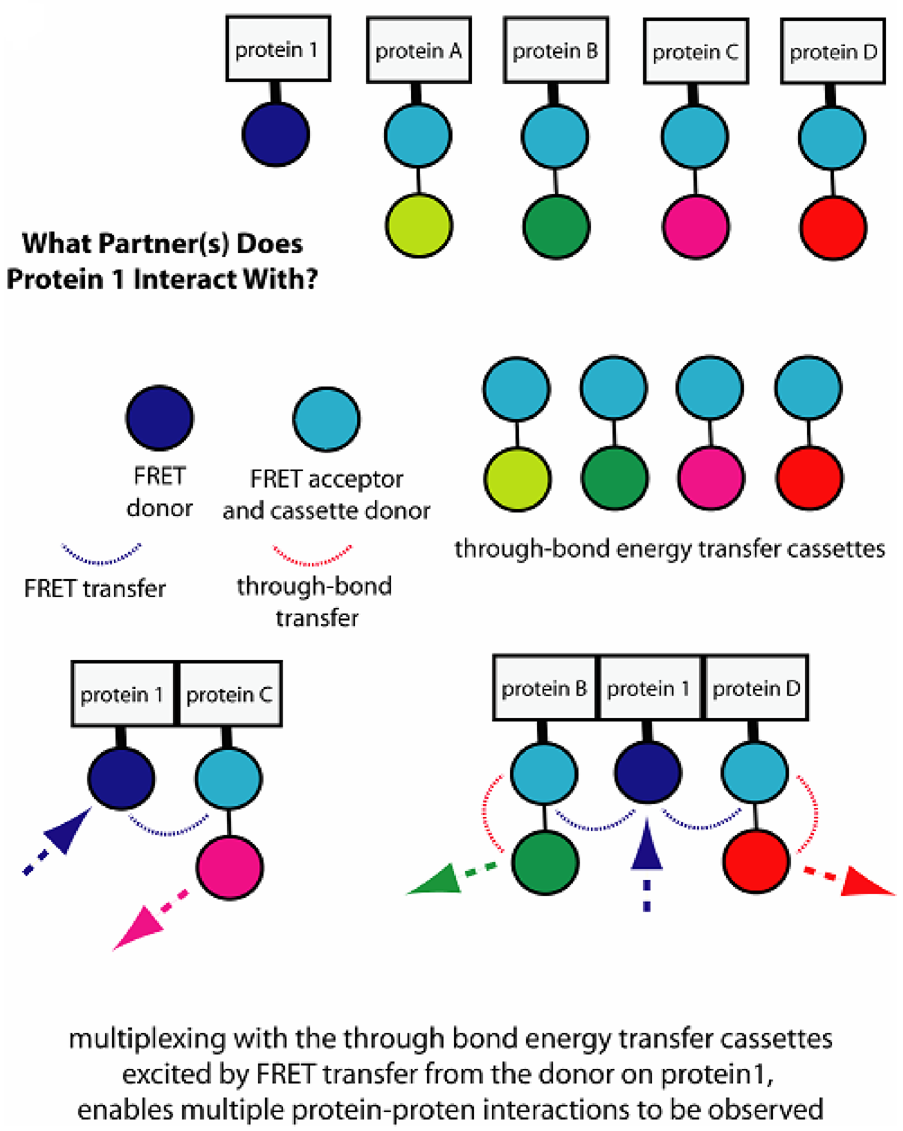

protein 1
protein A
protein B
protein C
protein D
What Partner(s) Does Protein 1 Interact With?
FRET donor
FRET acceptor and cassette donor
through-bond energy transfer cassettes
FRET transfer
through-bond transfer
protein 1
protein C
protein B
protein 1
protein D
multiplexing with the through bond energy transfer cassettes excited by FRET transfer from the donor on protein1, enables multiple protein-proten interactions to be observed

Figure 1.11. Proposed system for detecting protein-protein interactions.

CHAPTER II

DESIGN AND SYNTHESES OF HIGHLY FLUORESCENT GFP-CHROMOPHORE ANALOGS*

2.1 Introduction

The green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* is one of the most widely studied and exploited proteins in biochemistry and cell biology due to its strong intrinsic visible fluorescence, which is easily detectable by fluorescence spectroscopy.[66,67] The phenomenon of green fluorescence from jellyfish has been known for years.[68] However, the material that causes the fluorescence is not known until 1962 when it was isolated from *Aequorea victoria* and identified as a protein by Osamu Shimomura.[69] Since then, this green fluorescent protein (GFP) has been extensively studied. Shimomura was only interested in GFP's bioluminescence and wanted to understand the biochemistry involved in the process. Prasher was the first person to realize the potential of GFP as a tracer molecule. He was able to isolate and characterize GFP gene.[70] Two years later in 1994, Charlfie successfully expressed GFP in *E. coli* so that they fluoresced green when irradiated with blue light.[24] For the first time, Chalfie demonstrated the value of GFP as a biological marker. While they were all interested in GFP and demonstrated that it can be used as a fluorescent tag, it is Roger Tsien who is responsible for much of our understanding of how GFP works and for developing mutants of GFP. His group has developed mutants that are brighter and have different colors.[71-73] In fact, most of the GFP samples used today are products from Tsien's lab. For their discovery and development of the green fluorescent protein, Shimomura, Chalfie and Tsien were awarded the 2008 Nobel Prize in Chemistry.

---

*Reprinted in part with permission from "Syntheses of Highly Fluorescent GFP-chromophore Analogues", Liangxing Wu, Kevin Burgess, *J. Am. Chem. Soc.,* **2008**, *103*, 4089-4096. Copyright 2009 American Chemical Society.

The primary sequence of GFP was deduced from the nucleotide sequence of its cDNA by Prasher in 1992.[70] Wild type GFP is made up of 238 amino acids residues in a single polypeptide with a calculated molecular weight *ca* 26.9 KDa. The structure of the GFP chromophore from *Aequorea victoria* was first proposed by shimomura in 1979[74] and was further confirmed and corrected by Prasher in 1993.[75] The chromophore is a *p*-hydroxybenzylideneimidazolinone formed from residue 65-67, which are Ser-Tyr-Gly in the native protein. The crystal structure of GFP was solved in 1996.[76] GFP has a unique eleven β-sheets barrel-like structure with a diameter of *ca* 24 Å and a height of 42 Å. The β-sheets form the walls of the can and an α-helix runs diagonally through the cylinder. The chromophore is attached to the α-helix and is buried in the center of the β-can (Figure 2.1.a). Several polar residues and water molecules comprise a complex hydrogen-bonding network around the chromophore[67,76] (Figure 2.1.c). This network permits the transfer of protons between the chromophore and residue sidechains, the direction of which determines the ionization state of the chromophore. In the neutral form, Glu222 and Ser205 help to buffer and neutralize the chromophore. When exposed to UV light, Glu222 donates its charge to the chromophore by proton extraction through the H-bond network involving Ser205 and water, thereby generating the ionized form of the chromophore. The H-bond pattern in the chromophore vicinity is thought to change, stabilizing the chromophore in its ionized state. His148 now stabilizes the Tyr-derived portion of the chromophore by H-bonding directly to the phenolic oxygen. The side chain of Thr203 now hydrogen bonds directly to this oxygen as well. The H-bond between Glu222 and Ser205 is lost in the ionized form of chromophore. The reestablishment of the original hydrogen bond network allows the chromophore to return the charge back to Glu222 regenerating the neutral isomer.[66]

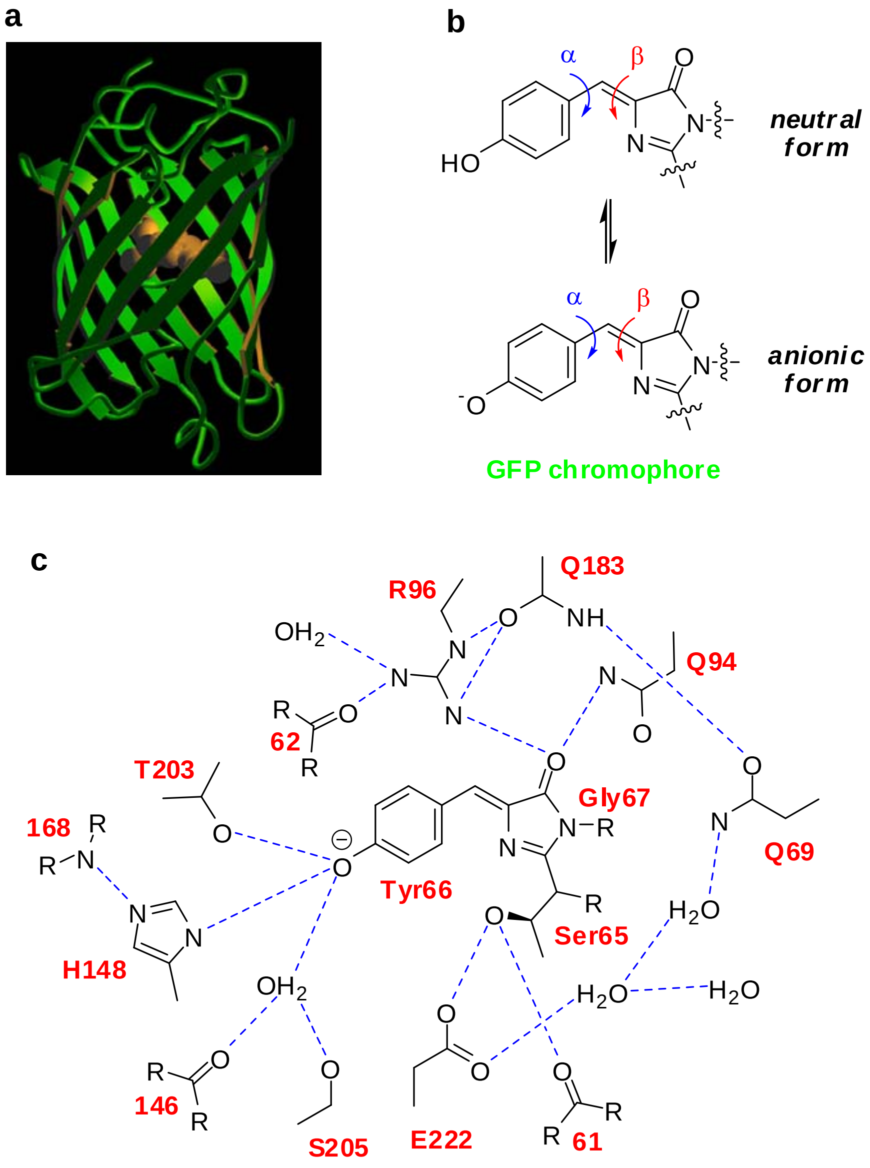

a
b
α
β
O
N
N
HO
neutral form
α
β
O
N
N
O
anionic form
GFP chromophore
c
R96
Q183
OH2
N
O
NH
N
Q94
R
O
N
N
O
62
R
O
T203
Gly67
168
R
N−R
N
Q69
R
N
O
O
N
Tyr66
R
H2O
N
N
H148
Ser65
OH2
H2O
H2O
O
R
O
O
O
146
R
O
S205
E222
R
61
R

Figure 2.1. (**a**) X-ray strucrure; (**b**) chromophore of GFP; (**c**) schematic diagram of the interactions between the chromophore and its surroundings in the S65T mutant. Possible hydrogen bonds are drawn as blue dashed lines.

The generally accepted mechanism for the biosynthesis of GFP chromophore is autocatalytic posttranslational cyclization and oxidation of the polypeptide backbone, involving Ser65, Tyr66 and Gly67 residues^{66,77,78} (Figure 2.2.). First, GFP folds into a native conformation which is non-fluorescent and called "apo-GFP". Then nucleophilic attack of the amide in Gly67 on the carbonyl of residue Ser65 followed by dehydration to form the imidazolinone. Finally, oxidation by molecular oxygen put the aromatic group of Tyr66 into conjugation with the imidazolinone. Only at this stage, the GFP is highly fluorescent.

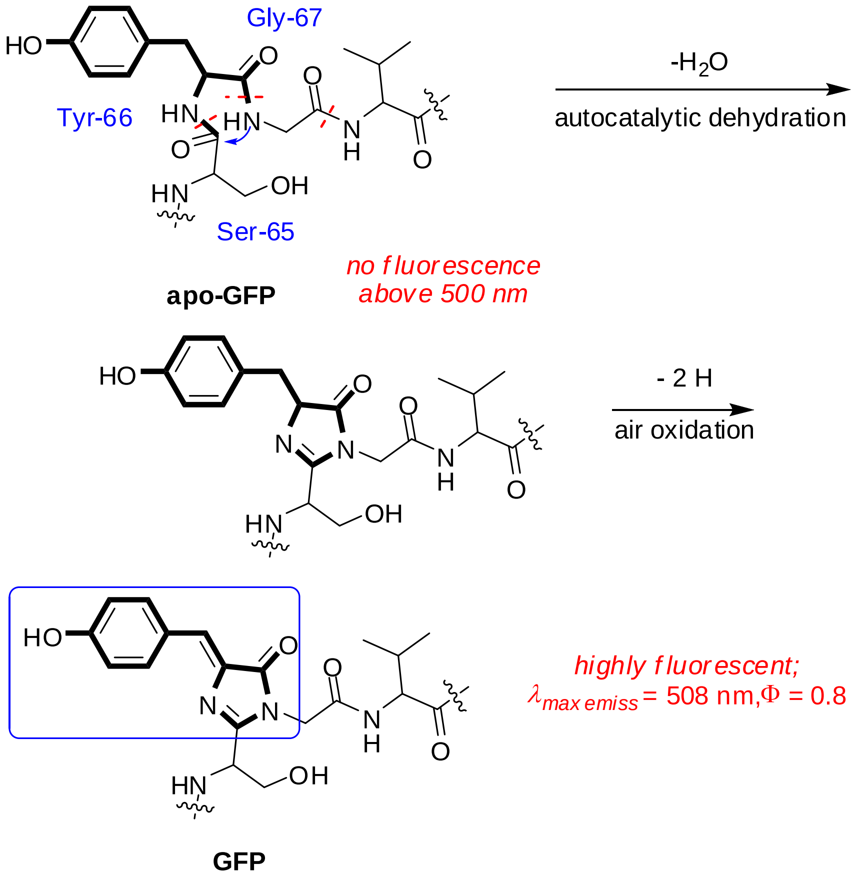
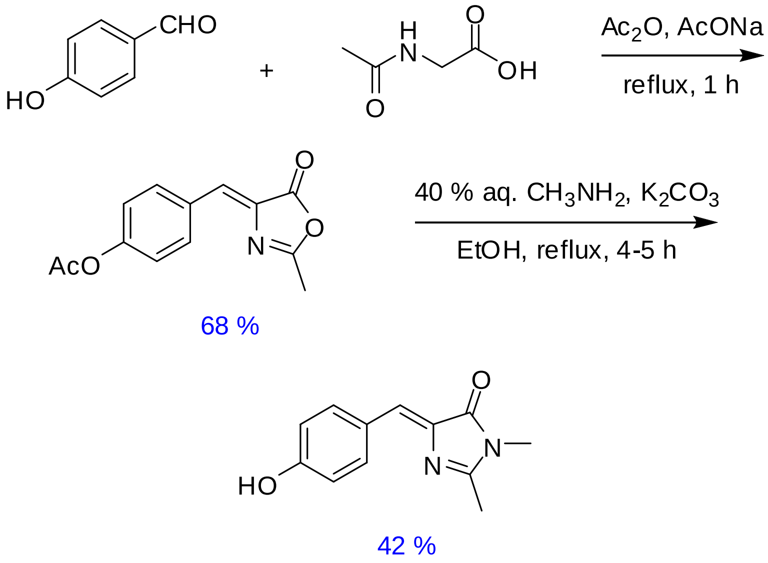

Gly-67
Tyr-66
Ser-65
apo-GFP
no fluorescence above 500 nm
-H2O
autocatalytic dehydration
- 2 H
air oxidation
highly fluorescent;
λmax emiss = 508 nm, Φ = 0.8
GFP

Figure 2.2. Biogenesis of the chromophore in wild-type GFP.

Chemical synthesis of GFP model chromophore is straight forward. The model compound was prepared via Erlenmeyer azlactone synthesis[79] as shown in Scheme 2.1. Reaction of *N*-acetylglycine, 4-hydroxybenzaldehyde and anhydrous sodium acetate in acetic anhydride provided the corresponding azlactones which was reacted with methylamine in the presence of potassium carbonate to afford the imidazolinone model compound.[78] Libraries of model chromophores of GFP and mutant proteins have been prepared similarly and the structure-property relationships were systematically studied in order to figure out the photophysical parameters that defines the GFP spectroscopic properties.[80,81]

Scheme 2.1. Chemical synthesis of model chromophore of GFP.



The absorption spectrum of wild-type GFP contains two main bands at *ca* 395 and 475 nm, and these are usually assigned to the neutral and anionic forms of the chromophore, respectively. Excitation of either neutral or anionic species results in a similar emission spectrum centered at 508 nm with a high quantum yield (Φ = 0.8).[67,82,83] This is presumably due to the fact that the phenolic oxygen of Tyr66 is more acidic in the excited state than in the ground state; excited-state proton transfer (ESPT) occurs resulting in a common anionic excited state that is responsible for the observed emission spectrum.[84]

The substitution of one or more amino acids within the chromophore or in its immediate proximity has allowed the development of GFP mutants with different absorption and emission properties. The first major improvement was a single point mutation (S65T: Ser65 → Thr) reported by Tsien in 1995.[72] This mutation dramatically improved the spectral characteristics of GFP, resulting in increased fluorescence, photostability and a shift of the major excitation peak to 488 nm with the emission peak kept *ca* 508 nm. Further mutation (F64L: Phe64 → Leu) to this scaffold generated enhanced GFP (EGFP) with improved folding efficiency. Many other mutants have also been made such as blue fluorescent protein (BFP), cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP). BFP derivatives contain the Y66H (Tyr66 → His) mutation which gave blue shifted absorption (383 nm) and emission (448 nm). Substitution of Tyr66 with Trp caused the chromophore to form with an indole component giving CFP which absorbed *ca* 440 nm and emitted *ca* 490 nm.[71] The red-shifted wavelength of the YFP derivatives is achieved by the T203Y (Thr203 → Tyr) mutation and is due to π-stacking interactions between the substituted tyrosine residue and the chromophore.[66] The recent discovery of red fluorescent protein (RFP) from a non-bioluminescent reef coral launched a second revolution of fluorescent proteins due to its significantly red-shifted excitation and emission maxima (λ_{ex} 558 nm, λ_{em} 583 nm, Φ 0.7).[73] The chromophore in RFP DsRed (Figure 2.3.) shares the same 4-hydroxybenzylidene- imidazolinone core with GFP chromophore, but with an acylimino group at the 2-position of the imidazolinone ring. The resulting additional conjugation from the acylimino group in the chromophore is responsible for the red-shifted spectroscopic properties of DsRed compared to those of GFP.

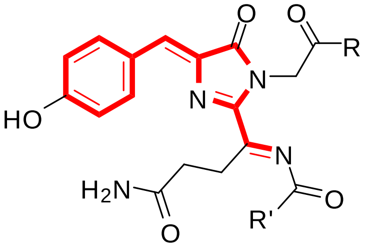


Figure 2.3. Structure of the red fluorescent protein (DsRed) chromophore.

It is now well known that the denatured-GFP and the chromophore isolated from the native protein will totally lose their fluorescence. Molecules representing the chromophores of GFP (**A**[78] and **B**[83]), and point mutants (**C** for Y66F mutant of GFP; **D** for CFP representing the Y66W mutant of GFP; **E** for BFP representing the Y66H mutant of GFP)[78] or naturally occurring analogs (**F** for red fluorescent protein RFP)[85] have been prepared and studied in solution.[86,87] They fluoresce at somewhat shorter wavelengths than their parent proteins and *with extremely low quantum yields.* This striking difference in quantum yields has been much discussed in the literature, but the consensus opinion is relatively simple.[80] Loss of fluorescence energy from the synthetic chromophores in solution is mainly the result of radiationless transfer arising from free rotation about the aryl-alkene bond (blue arrows, Figure 2.4.). Isomerization of the alkene (red arrows) and the polar nature of media relative to the apolar environment within the β-barrel protein structures may also be contributing factors, but the main one is that free rotation parameter. Steric and electronic factors prevent free rotation of the aryl substituents when the chromophores are encapsulated in the proteins. The environment also disfavors *E*:*Z*-isomerization, and provides an apolar medium that is often conducive to fluorescence.

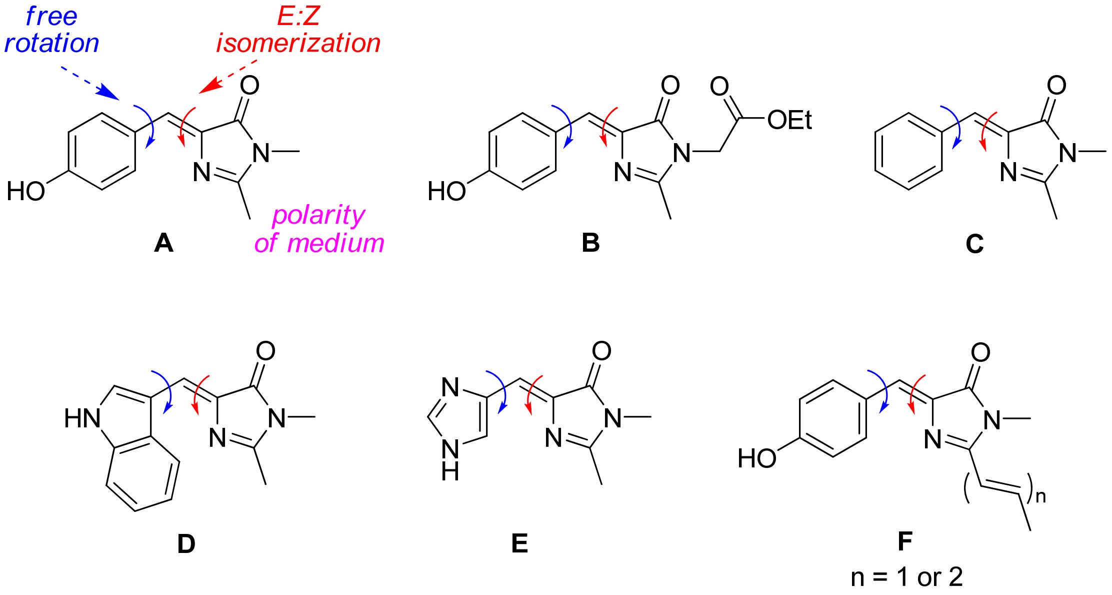


Figure 2.4. Some synthetic analogs of chromophores in fluorescent proteins (FP).

The research discussed in this chapter explores a hypothesis that highly fluorescent analogs of the GFP chromophore could be made by including boron. The conceptual genesis of these analogs **1**, and the particular compounds that in fact have now been made, are represented in Figure 2.5. It was supposed that inclusion of the boron atom would preclude free rotation of the aryl-alkene bond so these chromophore analogs would be highly fluorescent in solution. Thus the specific aims for this work are:

(i) design and syntheses of highly fluorescent GFP-chromophore analogs

(ii) studies of the structure-property relationships

(iii) preparation of water-soluble derivatives for biomolecular labeling

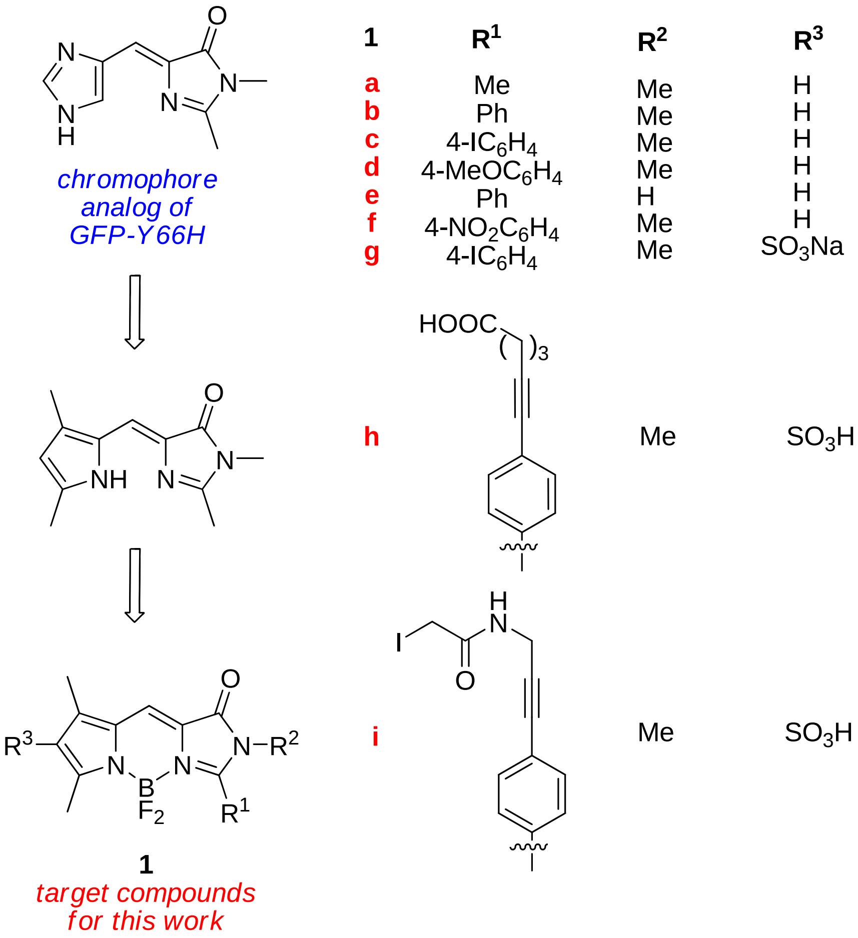

chromophore analog of GFP-Y66H
1
target compounds for this work
1 R1 R2 R3
a Me Me H
b Ph Me H
c 4-IC6H4 Me H
d 4-MeOC6H4 Me H
e Ph H H
f 4-NO2C6H4 Me H
g 4-IC6H4 Me SO3Na
h Me SO3H
i Me SO3H
HOOC
( )3
I
O
H
N
N-R2
R3
N
B
F2
NH

Figure 2.5. Thought process that led to the design of analogs **1** and the derivatives discussed in this chapter.

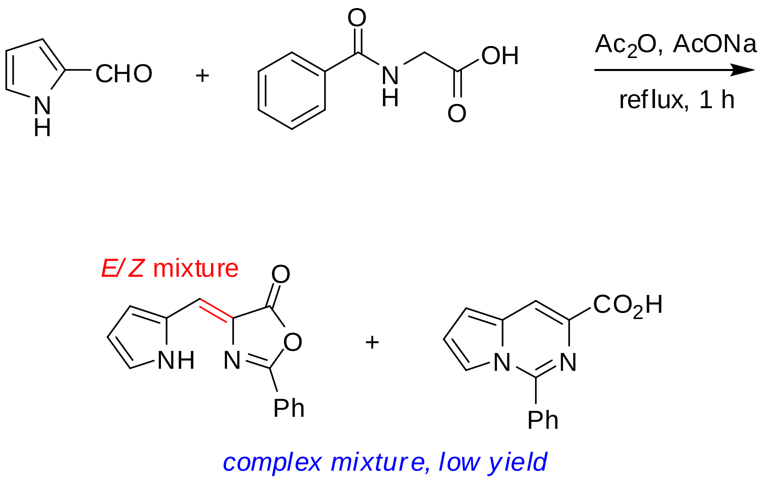
2.2 Results and discussions

2.2.1 Syntheses of the highly fluorescent GFP-chromophore analogs

To the best of our knowledge, nearly[83] all the syntheses of GFP-chromophore analogs are based on condensation of hippuric or aceturic acid derivatives with aldehydes (the "Erlenmeyer azlactone synthesis"[79]).[78,80,81,85,87,88] However, at least one report indicates these conditions do not work well for pyrrole-2-carbaldehyde[89] which gives a mixture of products, and in fact our attempts to use them for similar substrate were also unsuccessful, which led to complex mixtures and very low yield of the desired product (Scheme 2.2.). Consequently, a new strategy was developed that began with the synthesis of an imidazolinone[90], then condensation of this with the pyrrole-aldehyde. That route was successful for compounds **1a** – **e**.

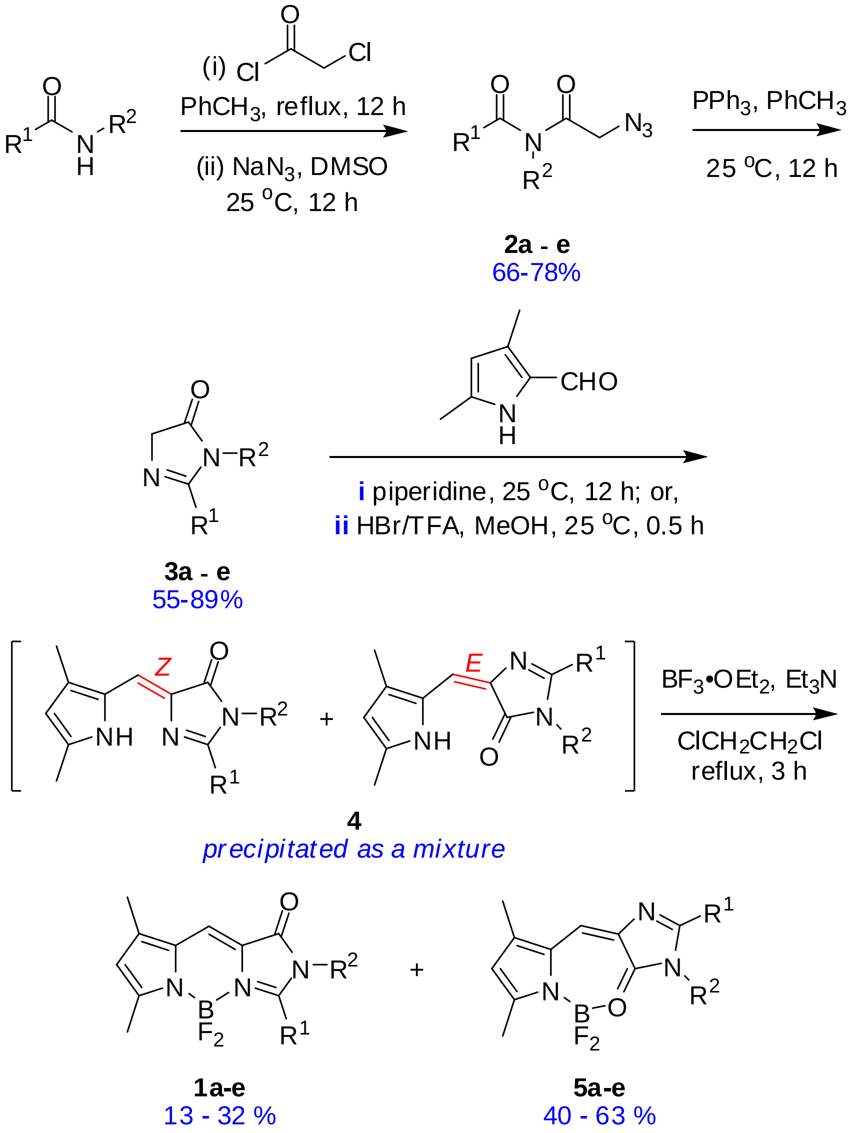
Scheme 2.2. Routes to target compounds **1.** **a** conventional condensation approaches were unsuccessful for pyrrole-2-carboxaldehyde;[89] **b** the strategy developed here involving condensation of imidazolinones.

a



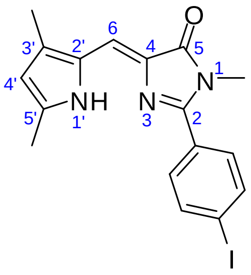
Scheme 2.2. continued.

b

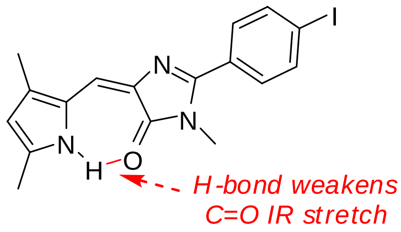

(i) Cl Cl O
PhCH3, reflux, 12 h
(ii) NaN3, DMSO
25 °C, 12 h
R1 N R2 H O
R1 N R2 N3 O O
2a - e
66-78%
PPh3, PhCH3
25 °C, 12 h
CHO N H
i piperidine, 25 °C, 12 h; or,
ii HBr/TFA, MeOH, 25 °C, 0.5 h
3a - e
55-89%
Z
E
NH
N−R2
R1
4
precipitated as a mixture
BF3•OEt2, Et3N
ClCH2CH2Cl
reflux, 3 h
N B F2 N
B−O
1a-e
13 - 32 %
5a-e
40 - 63 %

We initially believed the main challenge in the synthesis described in Scheme 2.2. was that compounds **4** tended to form as mixtures of *E*:*Z*-isomers, but later it was determined that these materials isomerized under the reaction conditions to form **1** and **5**. This conclusion was reached in the following way. Compound **1c** and the corresponding intermediates **4c** were chosen as models for the series. First, evidence was collected for assignment of *E*- or *Z*- stereochemistry for **4c** and related materials.[91,92] Table 2.1. shows selected spectroscopic parameters for the two isomers of **4c**. Coupling between the C^6-H proton and the C^5=O carbonyl carbon was probably the most definitive stereochemical probe: a larger coupling constant was observed for one isomer, which was therefore assigned an *E*-configuration. This assignment was supported by observation of a reduced wavenumber for stretching of the carbonyl group in the putative *E*-isomer, as would be expected from internal H-bonding. Having made this assignment, then chemical shift differences of the protons at H^6, $H^{1'}$, $H^{4'}$ and the carbon at C^5 were noted as possible indicators of *E*- or *Z*- stereochemistry.

Table 2.1. Spectroscopic properties of *Z*- and *E*-**4c** isomers



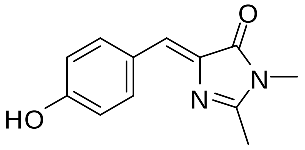
***Z*-4c**



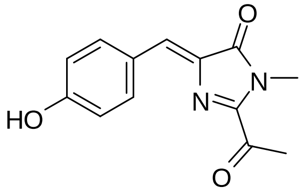
H-bond weakens C=O IR stretch

***E*-4c**

	${}^{3}J_{H6,\,C5}$ (Hz)	$\nu_{C=O}$ (cm^{-1})	δ_{H6} (ppm)	δ_{C5} (ppm)	$\delta_{H1'}$ (ppm)	$\delta_{H4'}$ (ppm)
***Z*-4c**	3.8	1686	7.17	169.8	10.62	5.87
***E*-4c**	8.9	1663	7.33	168.6	13.01	5.99



A



G

Work by others on syntheses of GFP-chromophore analogs (via approaches analogous to that shown in Scheme 2.2.a) leading to compounds like **A**[78] and **G**[88] could have led to isomeric mixtures but apparently that was not a problem for those structures. Thus the *Z*-isomer of compound **A** was shown to be more stable than the *E*-form by 2.1 kcal•mol^{-1} via experimental measurements.[91] Similarly, B3LYP calculations for molecule **G** indicated that the *Z*-isomer was thermodynamically preferred by 1.66 kcal•mol^{-1}.[93] However, the same B3LYP method was applied to compound **4c** in the current work and this led to the wrong conclusion. Specifically, the *E*-isomer was calculated to be more stable both *in vacuo* and in a medium of the same dielectric constant as DMSO (by 1.33 and 0.61 kcal•mol^{-1} at 25 ^{o}C, and 0.82 and 0.60 kcal•mol^{-1} at 100 ^{o}C, respectively), whereas NMR experiments in d^6-DMSO showed that the *Z*-isomer was in fact the thermodynamically preferred one. Details of the calculations are given in the appendix material, and spectra for the thermal isomerization studies are shown in Figure 2.6.a. In fact, the thermal isomerization in d^6-DMSO at 100 °C was quite slow (the reaction approaches equilibrium after 51 h indicative of a very high energy barrier to *E*,*Z*-isomerization). The *E*,*Z*-ratio at equilibrium was *ca* 0.6:1.0 indicating the *Z*-isomer was more stable by 0.41 kcal•mol^{-1} at 100 ^{o}C. Photoisomerization of two different *E:Z* mixtures of **4c** in $CDCl_3$ radiated at 360 nm was also studied, the data (Figure 2.6.b) indicates the *E*-isomer is dominant in the photostationary state.

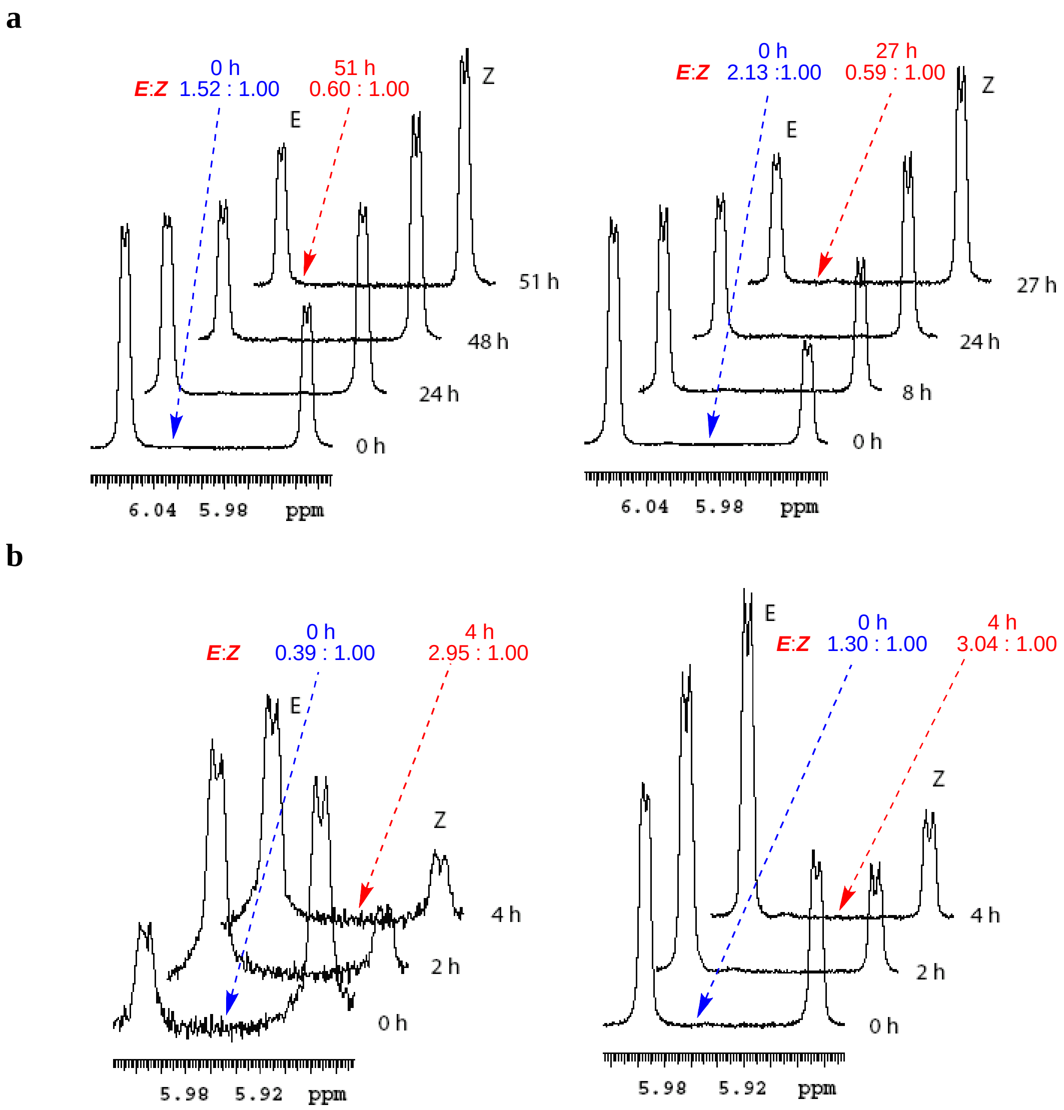
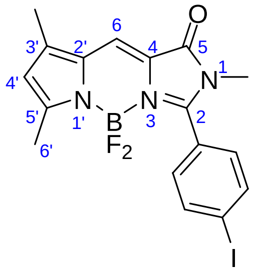


Figure 2.6. Isomerization of *E*,*Z*-mixtures of compound **4c**. **a** thermally in d^6-DMSO at 100 °C; **b**, and photochemically in $CDCl_3$ at 25 °C and 360 nm radiation.

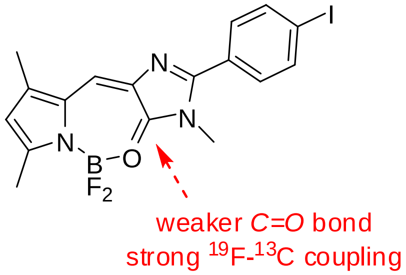
The next step in this project was to make the final products **1** from the intermediates **4**. Two isomeric products could be formed, **1** and **5**, and the spectroscopic strategies used to differentiate them were similar but not identical to those applied to compound **4c**. Coupling between the C^6-$\underline{H}$ proton and the $\underline{C}^5$=O carbonyl carbon is complicated by additional spin-pairings involving the *B*- and *F*-nuclei. Consequently, that particular spectroscopic probe is not so conveniently measured, however the *E*-

isomer is easily identified by the presence of coupling between the ^{19}F to C^{5}=O atoms. Further, the carbonyl stretch for compounds **5** is less than that for **1**; in this case this is because the boron atom is directly coordinated with the carbonyl C=O only for compounds **5**. Single crystal X-ray analyses were also obtained for compound **1c** and **5c**, confirming the assignment in this case (see below). Table 2.2. shows the key spectroscopic parameters for compounds **1c** and **5c**. In retrospect, the magnitude and direction of small chemical shift differences for these two series of compounds show consistent trends (Figure 2.7.) that could be used to make tentative assignments for new compounds in the series.

Table 2.2. Spectroscopic properties of **1c** and **5c**



1c



5c

	$\nu_{C=O}$ (cm^{-1})	δ_{C5} (ppm)	δ_{H6} (ppm)	$\delta_{H4'}$ (ppm)	δ_{F} (ppm)
1c	1705	160.7	7.41	6.01	37.02
5c	1610	158.4 (t, J_{CF} = 6.1Hz)	7.64	6.28	34.20

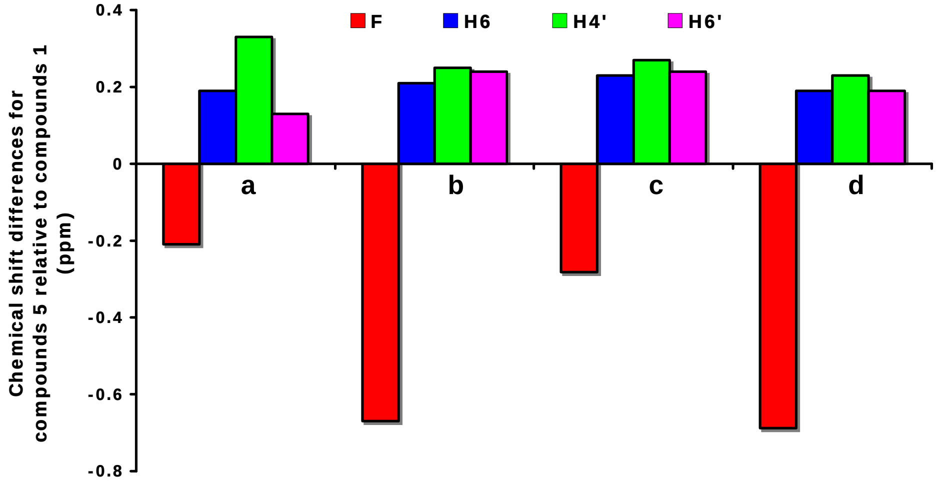


Figure 2.7. Trends in chemical shift differences for compounds **5** relative to compounds **1** illustrate that these can be used to differentiate between them. Here the zero line is the chemical shift values for compounds **1**, and the relative chemical shift difference for the corresponding compound **5** is indicated. The chemical shift differences for the ^{19}F signals shown here are 0.1 x their actual values.

The X-ray structures of molecules **1c** and **5c** (Figure 2.8.) have at least two interesting features, one is relevant to the spectroscopic assignments made above, and the other relates more to their fluorescence properties. First, the C=O bond length for compound **1c** is 1.226(3) Å whereas the corresponding distance in compound **5c** is 1.311(5) Å. The first of these carbonyls is not complexed, so the bond length is shorter, and the bond is presumably stronger than in **5c** where its bond order is reduced by interaction with the BF_2-entity. Second, the pyrrole-derived ring in compound **1c** is almost exactly co-planar with the one originating from an imidazolinone; in fact, the dihedral angle between those two rings is only 0.51°. However, for compound **5c** the corresponding dihedral angle is 9.94°. Electronic spectra of compounds **1** and **5** are described in the next section. Briefly, compounds **1** tend to absorb at longer wavelengths than their isomers **5**, with higher extinction coefficients, and they also fluoresce at longer wavelengths. All these properties are consistent with **1** being the more planar, conjugated structure. Moreover, fluorescence quantum yields for **5** are, without exception, 10 to 100 times less than for the corresponding structures **1**. It may be that one contributing factor to this is rapid equilibration between the two enantiomeric twisted conformations serving as a pathway for radiationless decay.

Interestingly, similar phenomenon has also been observed in Dronpa, a GFP homolog evolved from a *Pectiniidae* coral. Dronpa can be repeatedly and reversibly switched between a fluorescent on- and a non-fluorescent off-state by irradiation with light.[94] Recently, it was found out that the fluorescent on-state of Dronpa is due to the *cis*-deprotonated (anionic) chromophore, whereas the non-fluorescent off-state corresponds to the protonated (neutral) *trans*-isomer.[95] On the basis of structural comparisons of the on- and off-states from NMR studies, Mizuno proposed that in a fluorescent state the chromophore is tethered to the opposite β-barrel by the hydrogen bond between the chromophore phenolate oxygen and the neighboring Ser142 hydroxyl group. This hydrogen bond holds the two rings of the chromophore in a *cis*-configuration. The His193 imidazole ring, which is located below the chromophore phenyl moiety, stabilizes the entire chromophore planar conformation. This rigid structure favors the radiative relaxation process from the first electronic state to the ground state. By contrast, the *trans*-chromophore in the non-fluorescent state lacks both the hydrogen bond with Ser142 and chromophore stacking with His193. These structural changes increase the flexibility not only of the chromophore but also a part of the corresponding β-barrel which might lead to the fluorescence loss and give the off-state.[96,97]

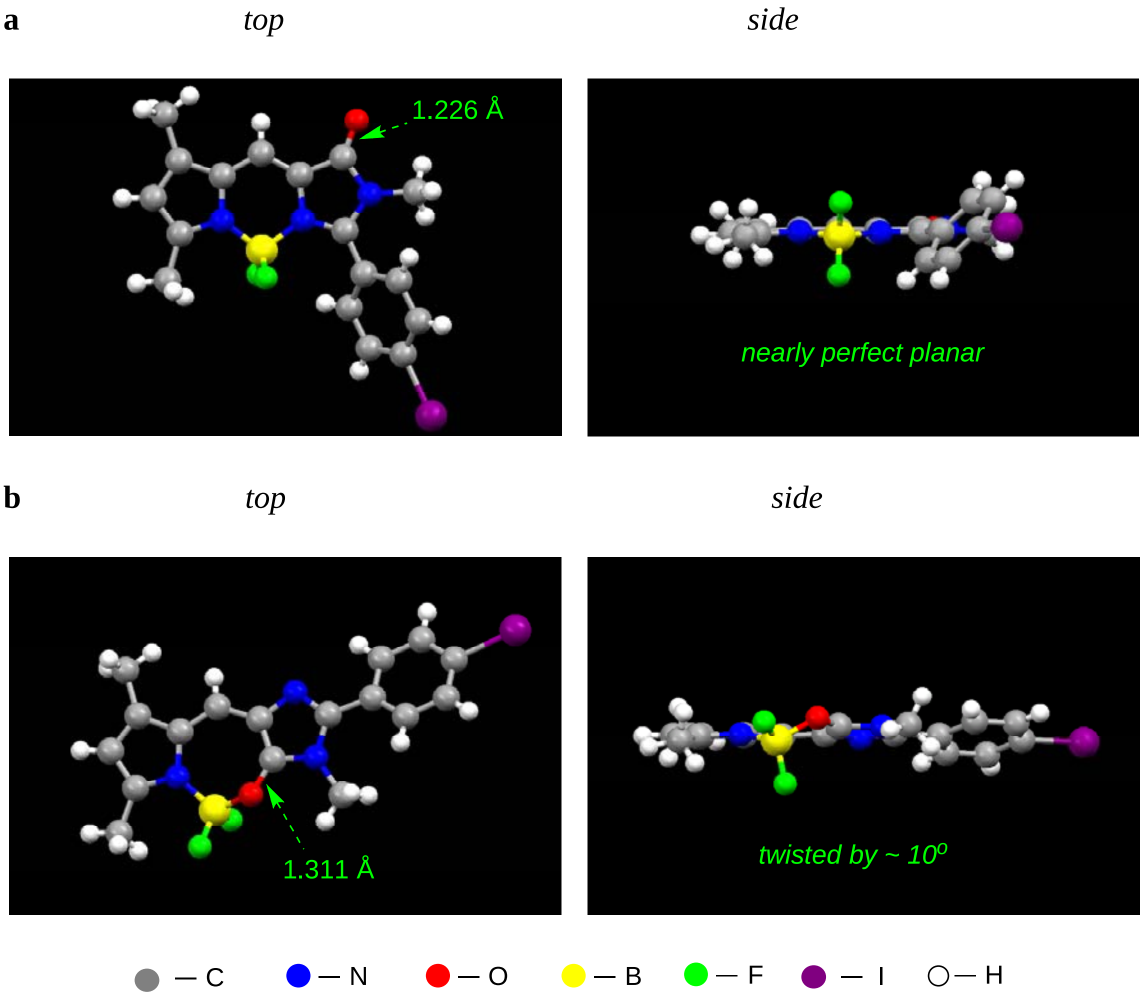
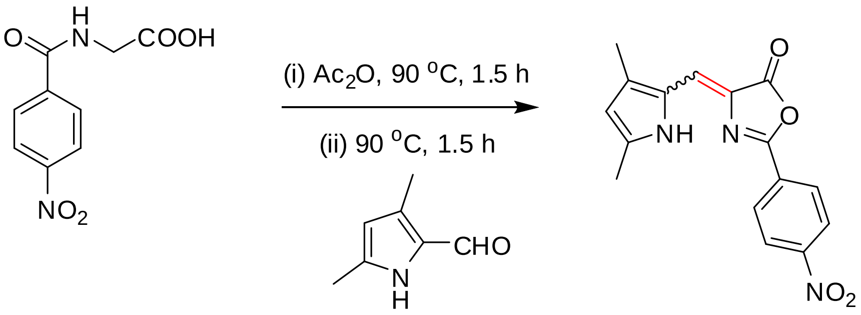


Figure 2.8. Top and side views of structures based on single molecule X-ray diffraction analyses of: (**a**), compound **1c**; and, (**b**) compound **5c**.

In the event, *E*,*Z*-isomerization of compounds **4** apparently does not significantly impact the syntheses of products **1** because the isomers seem to interconvert under the conditions used in the final step, as supported by the following observations. Synthesis of **4c** via condensation mediated by HBr/TFA (conditions ii in Scheme 2.2.b) in fact gave ***E*-4c** almost exclusively. However, when this material was converted to the boron-containing final products, compound **5c** and **1c** were formed in a ratio of approximately 2:1. Similarly, when a 1:1 mixture of ***E*-4c** and ***Z*-4c** was converted to the boron-containing products via the same reaction conditions, **5c** and **1c** were formed in a ratio of approximately 2:1. It appears that formation of compound **1c** from **4c** is a stereorandom process, *ie* little or no stereochemical information in the starting material is retained.

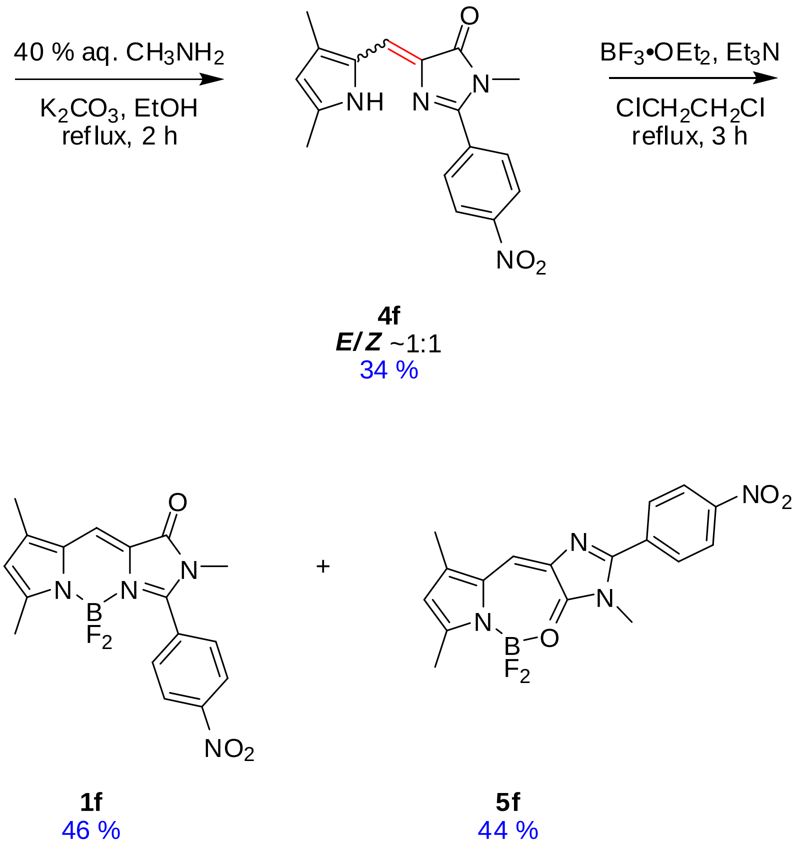
The pathway outlined in Scheme 2.2.b worked well for syntheses of compound **1a**-**e**, but there was a problem when the same protocol was applied to the 4-nitrobenzene-substituted product **1f**. Specifically, the intramolecular "aza-Wittig" closure did not afford the desired azlactone **3f**; instead, a complex mixture of products formed. Consequently, the route described in Scheme 2.3. was devised. Here the azlactone was formed via an intramolecular condensation of the hippuric acid derivative;[79] this intermediate was not isolated, but instead was condensed with 3,5-dimethyl-2-pyrrolaldehyde to give the azlactone-derivative **6**. Conversion of **6** to the corresponding imidazolinone **4f** was achieved via reaction with methylamine.[78] Finally, the BF_2-entity was introduced using the same conditions as outlined for the other compounds in the series in Scheme 2.2.b.

Scheme 2.3. Revised procedure for synthesis of compounds **1f** and **5f**.



6
E*/*Z ~1:1
46 %

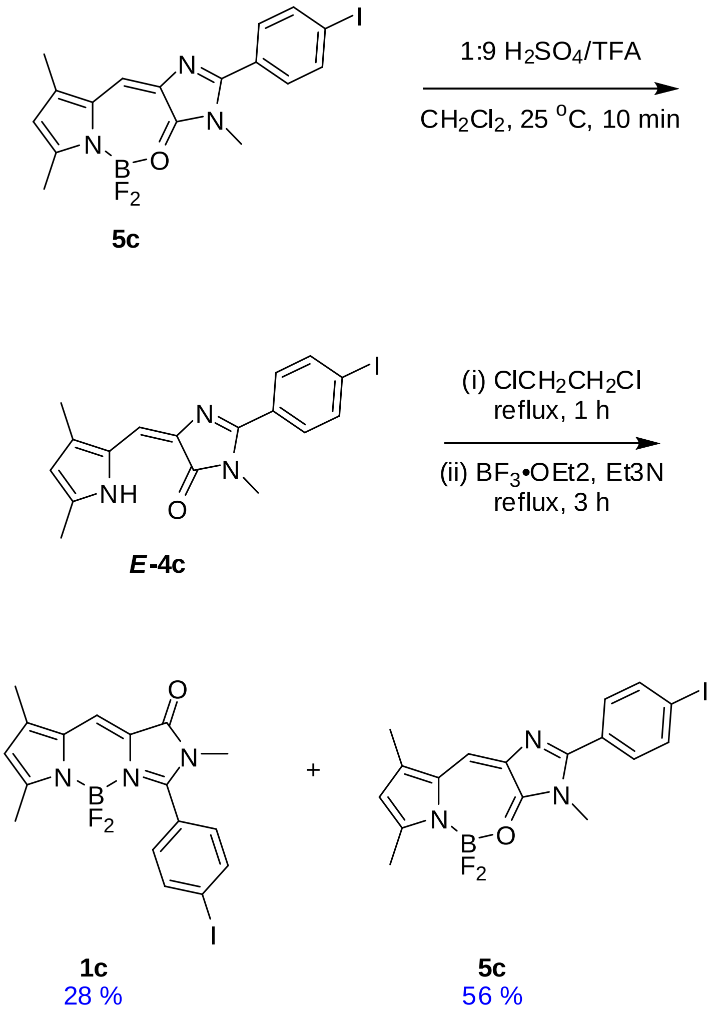
Scheme 2.3. Continued.


40 % aq. CH_3NH_2
K_2CO_3, EtOH
reflux, 2 h
BF_3•OEt_2, Et_3N
$ClCH_2CH_2Cl$
reflux, 3 h
O
NH
N
N
NO_2
4f
E*/*Z ~1:1
34 %
O
N
N
N
B
F_2
NO_2
+
NO_2
N
N
N
B
O
F_2
1f
46 %
5f
44 %

Compounds **1f** and **5f** were crystallized and studied by X-ray diffraction. They show the same features as outlined for the structures of **1c** and **5c**. Thus the carbonyl of **1f** is shorter than that in **5f**, and the molecular shape is almost perfectly flat whereas **5f** is twisted. Further details are given in the appendix material.

For synthetic convenience in preparing compounds **1**, concurrent formation of compounds **5** is a nuisance. Conditions to form compounds **1** with high selectivities were not identified in this study, but it was shown that the less fluorescent material **5** could be stripped of the BF_2-group giving intermediates **4** that could thereby be “recycled” (Scheme 2.4.). Thus, for our most studied compounds, those in series **c**, treatment of **5c** with sulfuric acid gave removal of the BF_2-group to regenerate intermediate **4c**. If this intermediate is not heated above room temperature or radiated with relatively intense and energetic light (*eg* 360 nm) then **4c** will be isolated as a pure *E*-isomer. However, as outlined above, this has no real consequence on the formation of compound **1c** because **5c** is invariably also formed.

Scheme 2.4. Method to recycle the undesired compounds **5** into the target materials **1**.

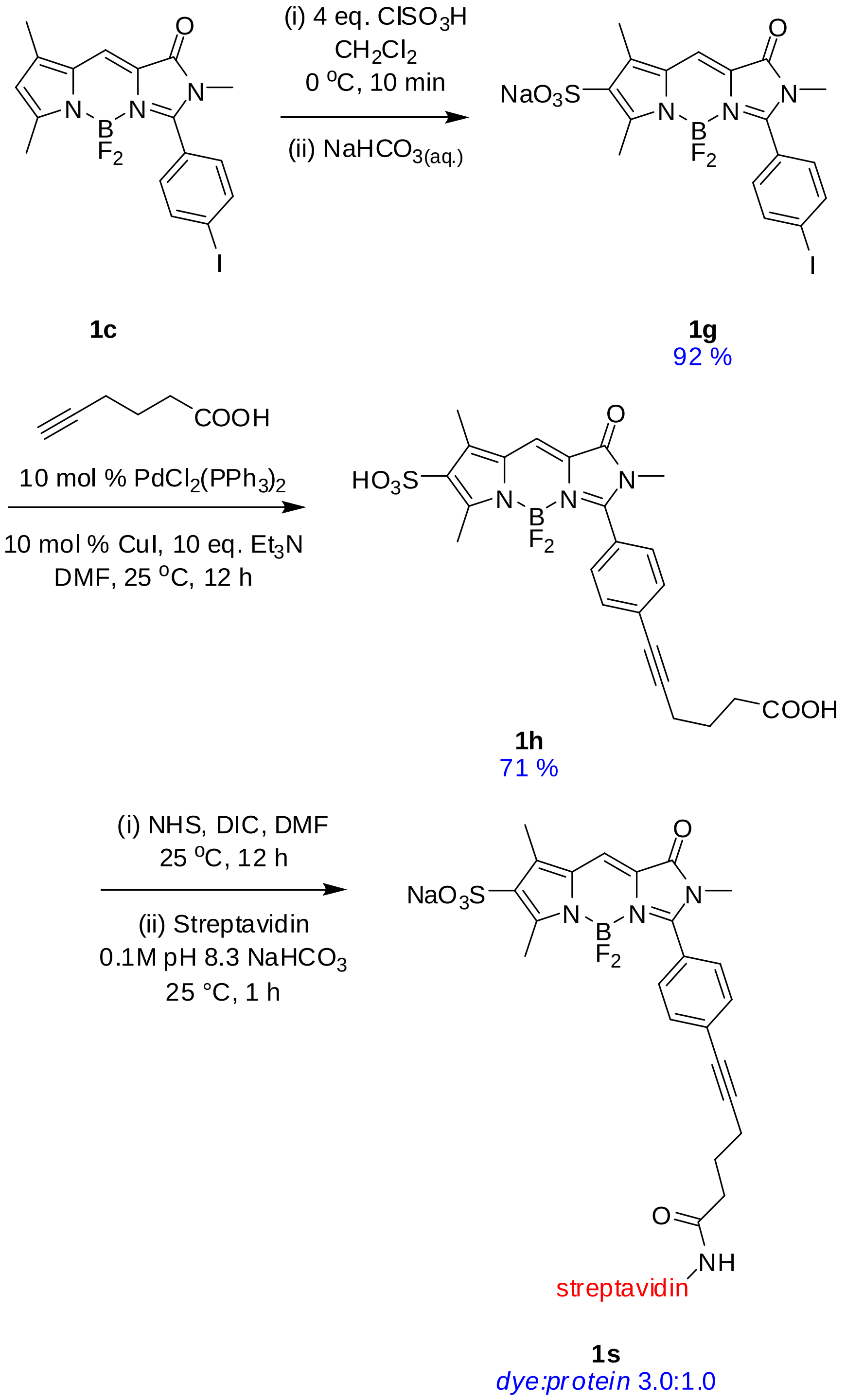


1:9 H_2SO_4/TFA
CH_2Cl_2, 25 °C, 10 min
5c
(i) $ClCH_2CH_2Cl$
reflux, 1 h
(ii) BF_3•OEt2, Et3N
reflux, 3 h
***E*-4c**
+
1c
28 %
5c
56 %

Review of the literature on BODIPY dyes leads to the conclusion that there are an abundance of lypophilic probes of that general type, but relatively few water-soluble ones.[98] Consequently, there was a strong motivation to prepare the first water-soluble variants of dyes **1**. Compound **1c** was chosen as a starting point because this material was conveniently obtained in gram amounts, and because it contains an aryl iodide functionality that can also be manipulated. Sulfonation of compound **1c** was achieved by treatment with excess chlorosulfonic acid at 0 °C (Scheme 2.5.). Practically, isolation of such sulfonic acids can be difficult. Here the crude reaction mixture was quenched with $NaHCO_{3\ (aq)}$ and the dye accumulated in the aqueous phase. The dichloromethane layer was removed, and the sodium salt of the product was extracted back into an organic medium (1:1 $CH_2Cl_2/^{i}PrOH$). After removal of the solvents the crude residue contained over 90% of the desired product. This material was then further purified by flash column chromatography on silica using 10% MeOH in CH_2Cl_2. Overall, the extraction and chromatography procedure is relatively easy because the products are so strongly colored.

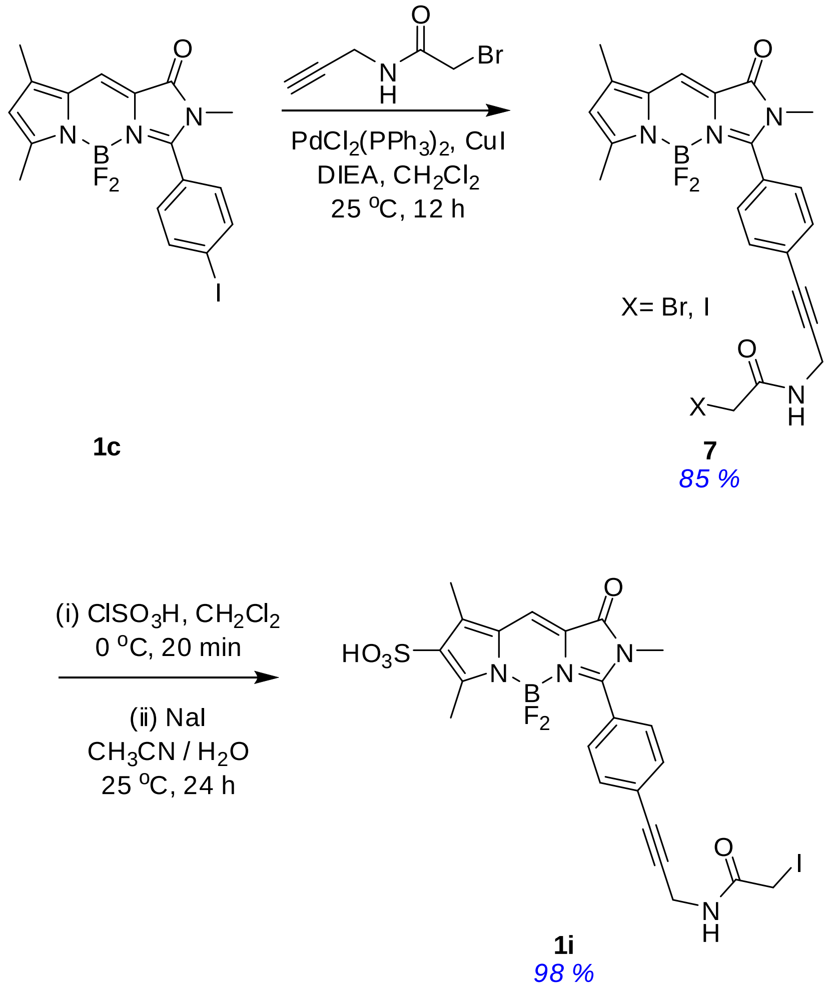
Scheme 2.5. also shows how the sulfonated product **1g** was coupled with 5-hexynoic acid via a Sonogashira reaction.[99] Again, the sulfonated product could be isolated via column chromatography on silica (20% MeOH in CH_2Cl_2). Compounds **1g** and **1h** are soluble in aqueous media, and **1h** contains a carboxylic acid functionality that could be used as a point of attachment to biomolecules. To illustrate this, and to measure the spectroscopic properties of this probe on a protein, **1h** was coupled to streptavidin. Measurement of the dye:protein ratio via UV[100,101] indicated this was almost exactly 3.0:1.0.

Scheme 2.5. Syntheses of the water-soluble probes **1g** and **1h**.


(i) 4 eq. ClSO$_3$H
CH$_2$Cl$_2$
0 °C, 10 min
(ii) NaHCO$_{3(aq.)}$
1c
1g
92 %
COOH
10 mol % PdCl$_2$(PPh$_3$)$_2$
10 mol % CuI, 10 eq. Et$_3$N
DMF, 25 °C, 12 h
1h
71 %
(i) NHS, DIC, DMF
25 °C, 12 h
(ii) Streptavidin
0.1M pH 8.3 NaHCO$_3$
25 °C, 1 h
streptavidin
1s
dye:protein 3.0:1.0

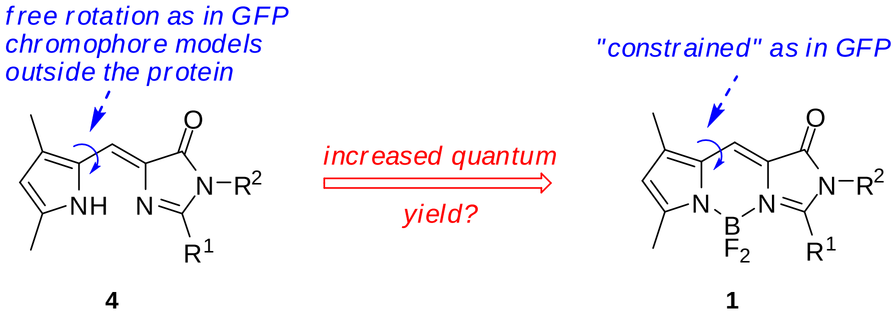
The thiol functional group is not very common in most proteins and can be labeled with high selectivity. Thiol-reactive probes often provide a means of modifying a protein at a defined site. Scheme 2.6. shows how a thiol-reactive analog **1i** was easily prepared from compound **1c**. Sonogashira reaction of **1c** with bromoacetamide derivative gave a mixture of bromo- and iodo- substituted derivatives **7**. After sulfonation and halide exchange, thiol-reactive probe **1i** was obtained in almost quantitative yield.

Scheme 2.6. Synthesis of water-soluble thiol-reactive analog **1i**.


PdCl2(PPh3)2, CuI
DIEA, CH2Cl2
25 °C, 12 h
X= Br, I
1c
7
85 %
(i) ClSO3H, CH2Cl2
0 °C, 20 min
(ii) NaI
CH3CN / H2O
25 °C, 24 h
1i
98 %

2.2.2 UV and fluorescence properties of the GFP-chromophore analogs

The central hypothesis of this work is that the fluorescence quantum yields of "unconstrained" GFP analogs like **4** will be greatly increased in compounds such as **1** that are conformationally locked by a boron atom. This was verified, and illustrative data are shown in Table 2.3.



free rotation as in GFP chromophore models outside the protein
"constrained" as in GFP
increased quantum
yield?
NH
N
N
O
R1
R2
N
B
F2
4
1

The unconstrained molecule ***Z*-4c** has a very low quantum yield (0.0005) compared to the locked analog **1c** (0.86); this is a dramatic illustration of the hypothesized effect. Several other differences in the electronic spectra of these materials were also observed. The UV absorbance and fluorescence emissions of compound **1c** are both red-shifted relative to ***Z*-4c**, and the fluorescent emission from **1c** is sharper (Figure 2.9.). For further comparisons, data for the *trans*-unconstrained intermediate ***E*-4c**, and the isomeric locked product **5c** are also given in Table 2.3. The UV and fluorescence properties of ***E*-4c** are almost identical to those of ***Z*-4c**. The locked compound **5c** has a quantum yield that is approximately 30-fold greater than ***E*-4c**, and its absorbance and emission wavelengths are also red shifted. The quantum yield of **5c** is, however, 86-fold less than **1c**, possibly for the reason implied in the discussion of the X-ray structures of **1c** and **5c**, *ie* rapid conformational equilibria between ring-puckered forms leading to radiationless decay pathways. Other consequences of deviation from planarity for **5c** are that the absorbance is blue-shifted, and the extinction coefficient is less.

Table 2.3. Spectroscopic properties of **4c**, **1c** and **5c** in MeOH

	$\lambda_{\text{max abs}}$ (nm)	ε ($M^{-1}cm^{-1}$)	$\lambda_{\text{max emiss}}$ (nm)	fwhm[a] (nm)	Φ[b]
***Z*-4c**	458	40200	515	69	0.0005
1c	492	57300	532	54	0.86 ± 0.02
***E*-4c**	458	48000	515	73	0.0003
5c	463	42700	530	54	0.01

[a] Fluorescence full width at half maximum peak height (fwhm); [b] Fluorescein in 0.1M NaOH as standard ($\phi = 0.92$).[102]

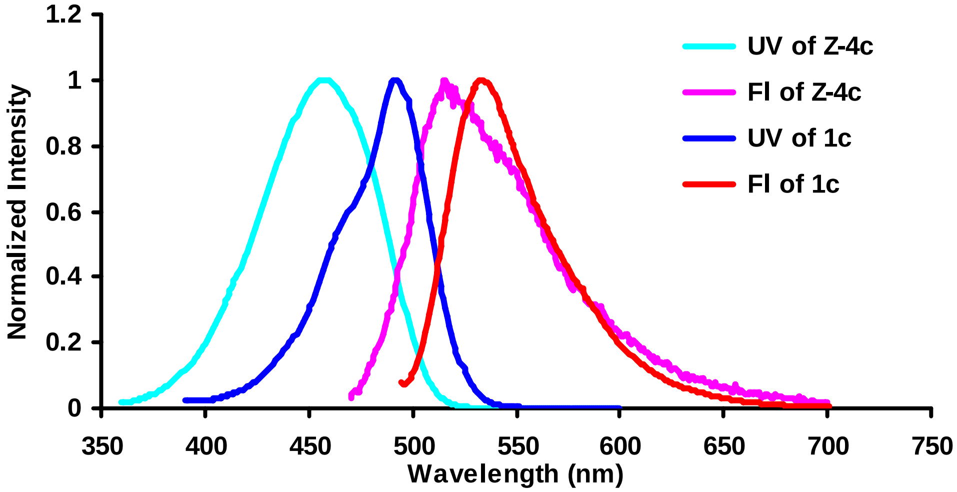


Figure 2.9. Normalized UV absorbance and fluorescence of ***Z*-4c** and **1c** in MeOH, (*ca* 10^{-6} M for UV and 10^{-6} - 10^{-7} M for fluorescence).

Table 2.4. gives the UV absorbance and fluorescence data for all the compounds **1** and **5**. The UV and fluorescence properties of the lypophilic compounds **1a** - **f** and **5a** – **f** are quite similar; Figure 2.10. shows illustrative spectra for compounds **1a** - **f** and **5a** – **f**.

The data in Table 2.4. show that the UV absorbance maxima for all the lipophilic compounds **1a** – **f** in MeOH are very similar ($\lambda_{\text{max abs}}$ = 494 $\pm$ 4 nm) except for compound **1a** which has a methyl- rather than an aryl-substituent on the core. Measurements of fluorescence full width at half-maximum peak heights (fwhm) show that the methyl substituted compound **1a** gives the sharpest fluorescence in the series. Fluorescence emission maxima for compounds **1b** – **e** are in a 10 nm range ($\lambda_{\text{max emis}}$ = 526 $\pm$ 5 nm). The outliers in terms of fluorescence are **1a** (because it does not have an aryl substituent) and the nitroaryl-substituted compound **1f**. Compound **1f**, fluorescing at 598 nm, has a much larger Stoke's shift than the others in the series, but a much lower quantum yield. The anomalously low quantum yield is most probably due to photoinduced electron transfer from the excited state of the fluorescent core to the LUMO of the nitroaryl group (*ie* d-PeT). When hexane was used as a medium in place of methanol then the quantum yield of the nitroaryl-substituted compound **1f** was only about 33 % less than **1a** – **e** in MeOH. This implies the d-PeT quenching mechanism is not dominant in apolar media, reflecting changes of the oxidation potentials in this molecule. The absorption and fluorescence maxima of **1a** – **e** were shifted by up to 20 nm to the red when a less polar media (dichloromethane) was used (see appendix for an expanded version of Table 2.4.).

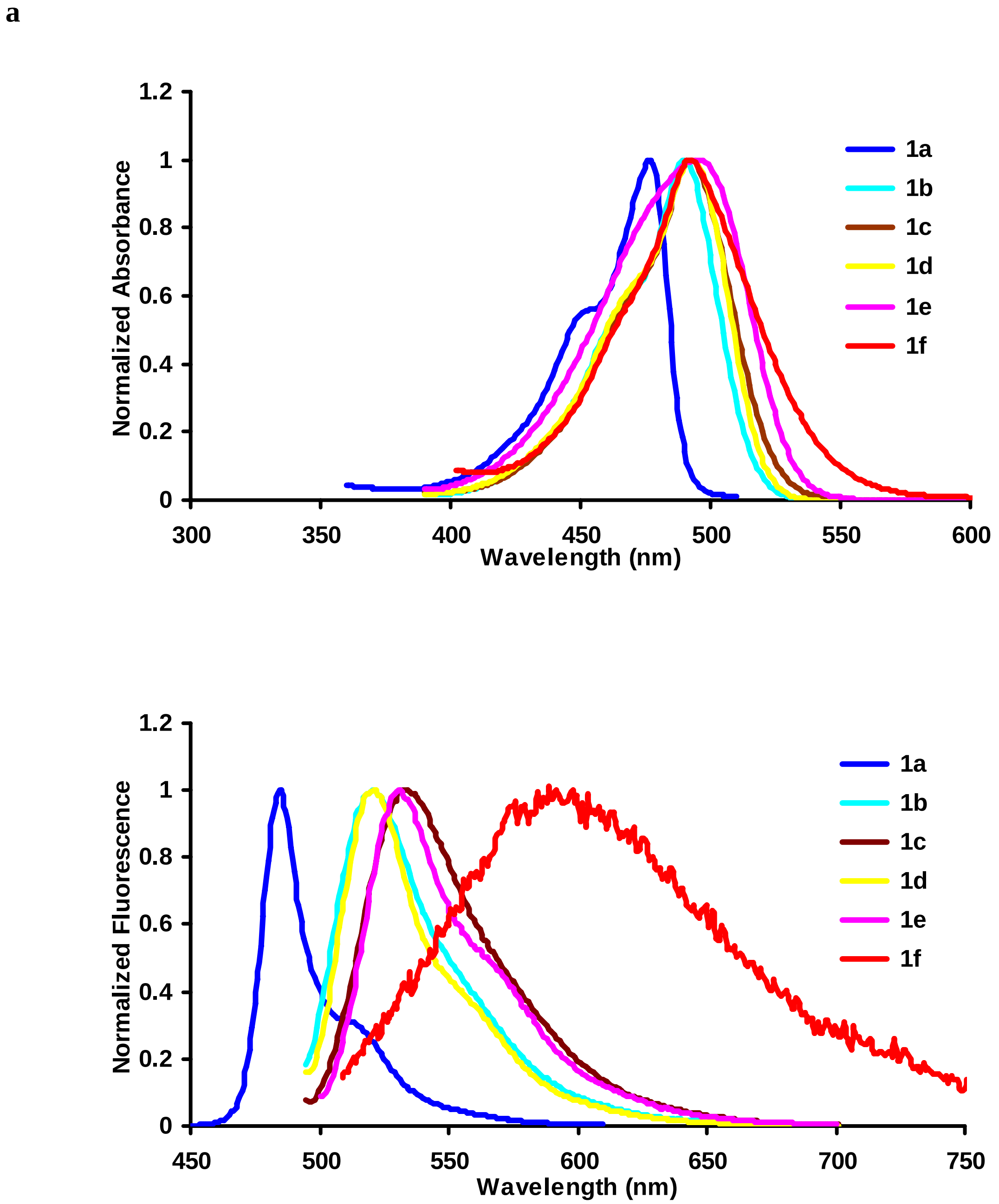
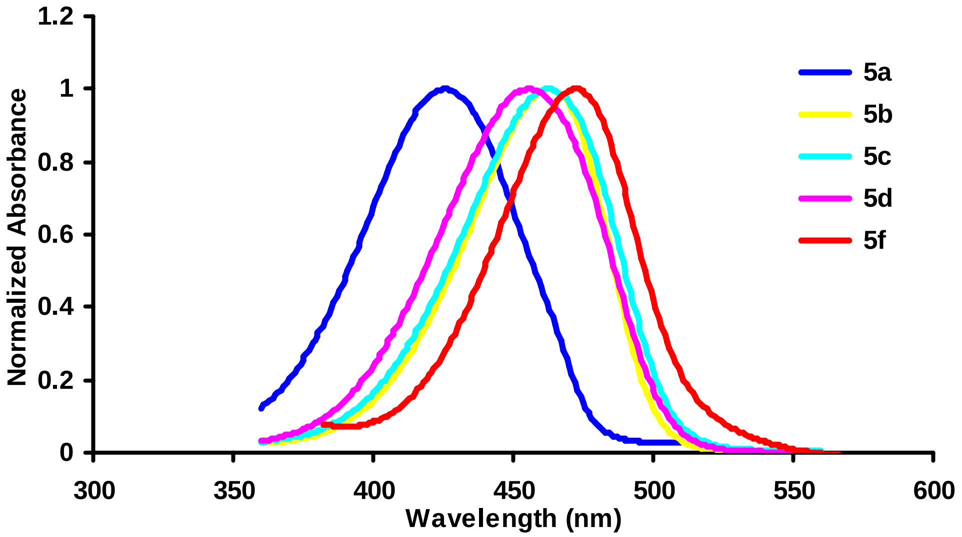


Figure 2.10. UV absorbance and fluorescence spectra for: (**a**) GFP-chromophore analogs **1** in MeOH; and, (**b**) GFP-chromophore analogs **5** in MeOH. Throughout: 10^{-6} M for UV and 10^{-7} M for fluorescence.

b



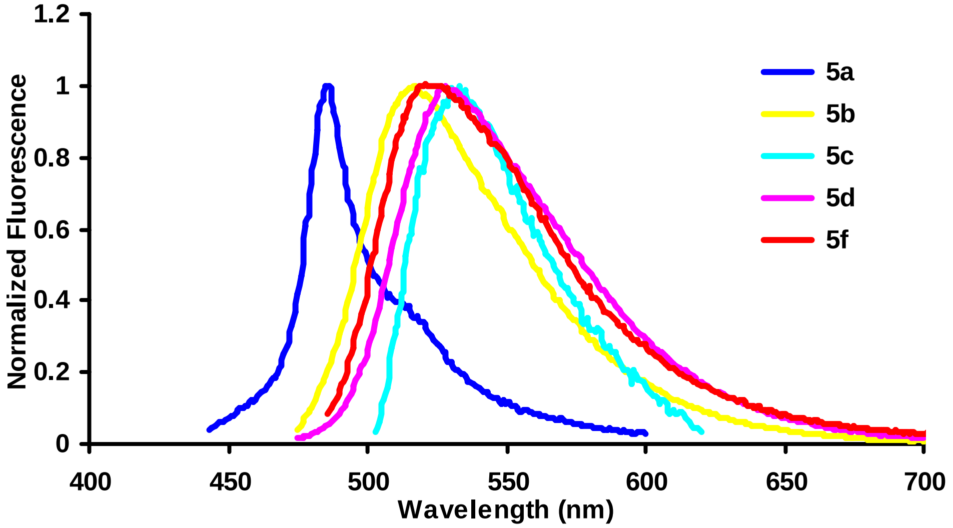


Figure 2.10. Continued.

Table 2.4. Spectroscopic properties of GFP-chromophore analogs

	$\lambda_{\text{max abs}}$ (nm)	ε ($M^{-1}cm^{-1}$)	$\lambda_{\text{max emiss}}$ (nm)	fwhm (nm)	Φ^{a}	solvent
1a	477	57300	485	19	0.89 ± 0.01	MeOH
5a	426	37000	485	25	0.0007	MeOH
1b	490	58700	521	46	0.87 ± 0.01	MeOH
5b	463	50500	517	64	0.006	MeOH
1c	492	57300	532	54	0.86 ± 0.02	MeOH
5c	463	42700	530	54	0.01	MeOH
1d	493	51300	521	38	0.85 ± 0.03	MeOH
5d	455	38600	528	69	0.03	MeOH
1e	495	37900	531	48	0.80 ± 0.01	MeOH
1f	492	44000	598	120	0.0004	MeOH
1f	498	n.d.	587	53	0.53 ± 0.01	Hexanes
5f	473	35500	527	73	0.005	MeOH
5f	486	n.d.	545	97	0.05	Hexanes
1g	488	48100	531	53	0.88 ± 0.01	MeOH
1g	481	46500	518	50	0.87 ± 0.01	Phos 7.4[b]
1h	488	35000	538	58	0.84 ± 0.01	MeOH
1h	482	34800	526	56	0.82 ± 0.01	Phos 7.4[b]
1i	489	26200	542	61	0.87 ± 0.01	MeOH
1i	483	26900	528	59	0.87 ± 0.01	Phos 7.4[b]
1s	482	n.d.	529	56	n.d.	Phos 7.4[b]

[a] Fluorescein in 0.1M NaOH as standard (ϕ=0.92); [b] 0.1M lithium phosphate buffer (pH=7.4).

Table 2.4. also gives data for the water-soluble dyes **1g**, **1h**, **1i** and **1s**. The absorbance and emission maxima for **1g**, **1h**, and **1i** were slightly blue-shifted in phosphate buffer relative to the same dyes (and to **1c**) in MeOH. Quantum yields for **1g** **1h** and **1i** were also high in phosphate buffers, just as in MeOH. The spectra of compound **1s** (Figure 2.11.), *ie* **1h** on streptavidin, was not significantly different to the parent dye; in other words, conjugation to this particular protein does not have significant effect on the fluorescent properties.

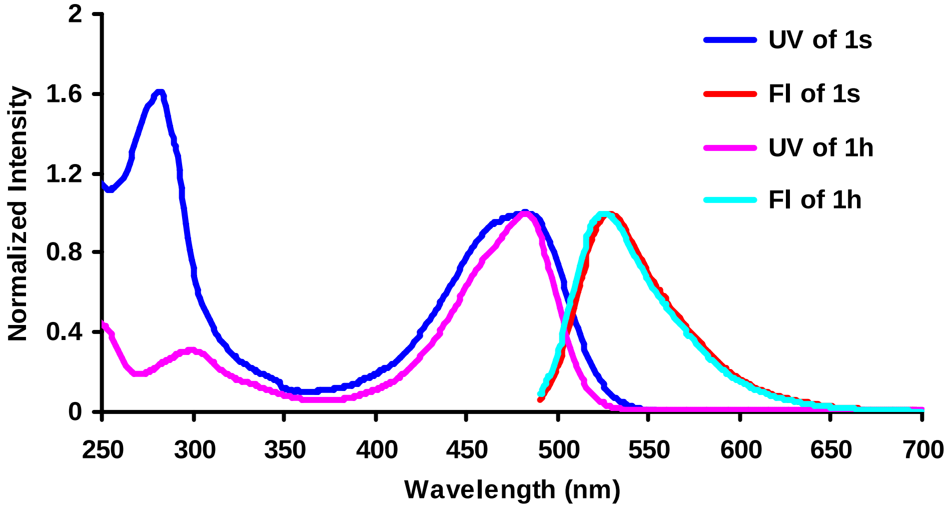


Figure 2.11. Normalized UV absorbance and fluorescence of **1h** and the labeled streptavidin **1s** in 0.1 M lithium phosphate buffer, pH 7.4.

2.3 Conclusions

Compounds **1** in this chapter are conceptual hybrids of GFP chromophore analogs and BODIPY dyes. The data presented prove that addition of the BF_2-entities to the open chain intermediates **4** greatly increases the quantum yields across a range of compounds. Overall, this provides more evidence for the assertion that GFP chromophores within the protein have enhanced quantum yields due to conformational locking within the β-barrel cocoon.

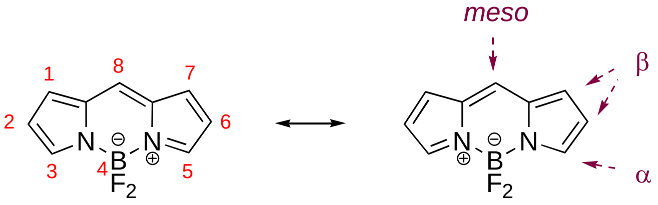
The fluorescence properties of compounds **1** tend to be very similar to BODIPY dyes. Both compounds are neutral, uncharged, highly fluorescent molecules that absorb with high extinction coefficients and emit in around 520 – 530 nm (unless high conjugating substituents are added). BODIPY dyes have many applications in chemistry and biology. Probes **1** could presumably be used in the same ways, and there may be some advantages to doing this, in terms of intellectual property at the very least. Further, the data presented here show that dyes **1** can be sulfonated and modified to include a point of attachment to biomolecules. The water-soluble modifications to **1** retain their fluorescent properties in aqueous media and when attached to a model protein (streptavidin).

CHAPTER III

A NEW SYNTHESIS OF SYMMETRIC BORAINDACENE (BODIPY) DYES*

3.1 Introduction

4,4-Difluoro-4-bora-3a,4a-diaza-*s*-indacene, or BODIPY® (hereafter abbreviated to BODIPY) dyes are extremely fluorescent materials and have become the preferred fluorophores in new fluorescent probes that have found numerous applications in biochemistry and molecular biology. This is due to their superior characteristics, such as (i) their relatively high absorption coefficients and fluorescence quantum yields; (ii) narrow emission bandwidth and therefore higher peak intensity and good signal/noise ratio; (iii) insensitivity to the polarity and pH of their environment; (iv) greater chemical and photochemical stabilities; (v) optical properties can be tuned by small modifications to their structures; (vi) relatively short synthesis to the simple derivatives.[98] The first member of BODIPY family was reported by Treibs and Kreuzer in 1968,[103] although relatively little attention was given to the discovery until the end of the 1980s.[104] Then, the potential use of this dye for biological labeling was recognized. Tons of patents and papers were published based on BODIPY dyes in the past two decades. Today, BODIPY dyes have become one of the most popular dyes in biological studies.



BODIPY dyes
small, intensely fluorescent systems

---

*Reprinted in part with permission from "A New Synthesis of Boraindacene (BODIPY) Dyes", Liangxing Wu, Kevin Burgess, *Chem. Commun.*, **2008**, *40*, 4933-4935. Copyright 2009 Royal Society of Chemistry.

Synthesis of BODIPY is based on the well-known pyrrole condensation reaction which was originally developed for the synthesis of certain porphyrins. Symmetric BODIPYs with *meso*-substituents are formed via condensation of a highly electrophilic carbonyl compound such as acetyl chloride, aldehyde or acid anhydride with two pyrrole units.[98,104]

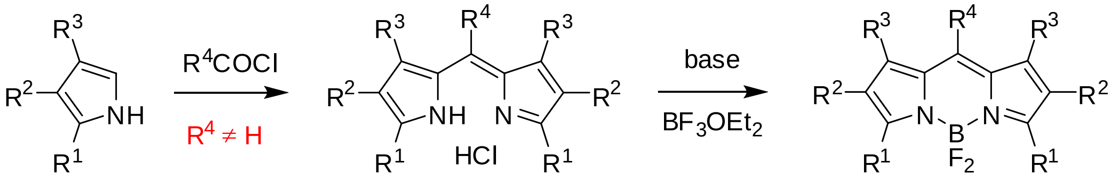
8-Substituted BODIPY dyes (*ie* ones with substituents in the *meso* position) tend to be relatively easy to prepare via condensation of acyl chlorides with pyrroles (Scheme 3.1.a). These transformations involve unstable dipyrromethene hydrochloride salt intermediates. The intermediate dipyrromethene hydrochlorides are easier to handle and purify as C-substitution increases but, even so, these are not generally isolated in syntheses of BODIPY dyes.

Other activated carboxylic acid derivatives could be used in place of acid chlorides. In the particular case of acid anhydrides, this concept has been reduced to practice. Scheme 3.1.b shows how a BODIPY derivative was prepared from glutaric anhydride.[105] An attractive feature of this chemistry is that a free carboxylic acid is produced, and this may later be used to attach the probe to target molecules.

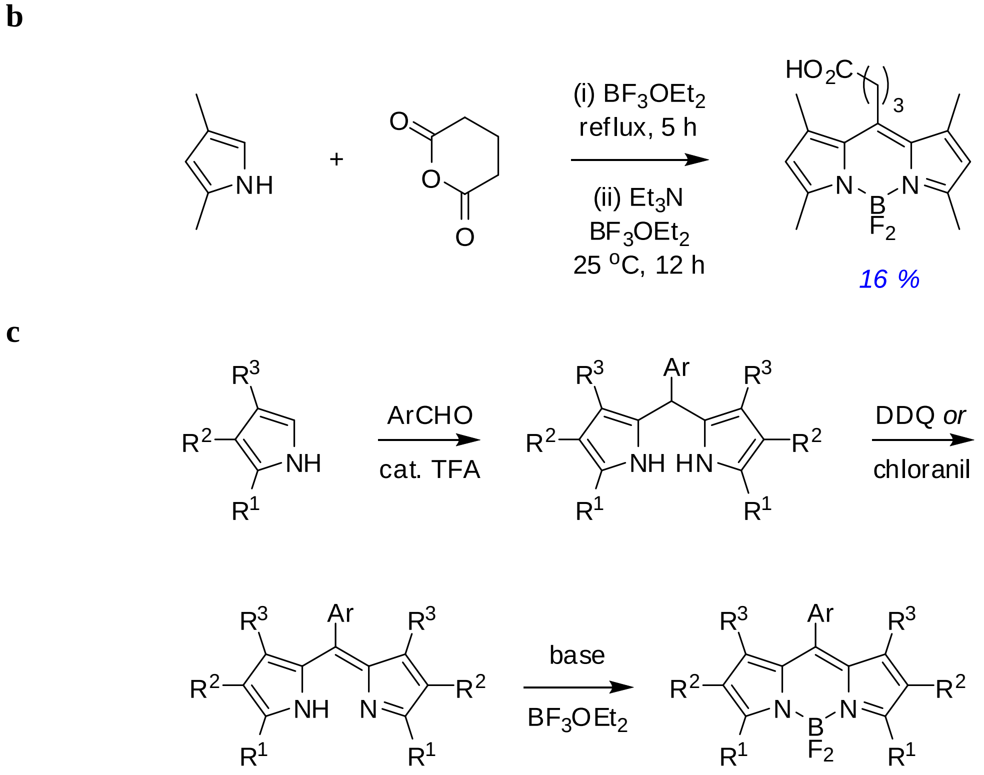
Syntheses similar to those depicted in Scheme 3.1.a and 3.1.b but which use aromatic aldehydes (to the best of our knowledge, aliphatic aldehydes have not been reported in this reaction) require oxidation steps. The reagents for these oxidations can introduce experimental complications. Thus in Scheme 3.1.c the oxidant used was *p*-chloranil or DDQ and eventually the byproducts from this had to be removed (in fact, this was done after complexation with the boron).

Scheme 3.1. Typical syntheses of symmetric BODIPYs with *meso*-substituent from (**a**) acyl chloride; (**b**) anhydride; (**c**) aldehyde.

a

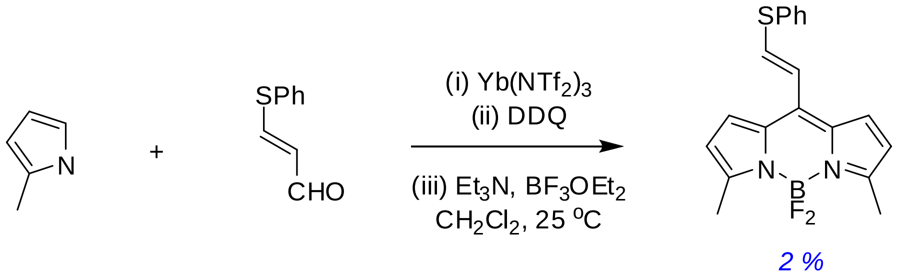


Scheme 3.1. Continued.


b
NH
+
O
O
O
(i) BF3OEt2
reflux, 5 h
(ii) Et3N
BF3OEt2
25 °C, 12 h
HO2C
3
N
B
N
F2
16 %
c
R3
R2
NH
R1
ArCHO
cat. TFA
Ar
NH HN
DDQ or
chloranil
N
base
BF3OEt2
F2

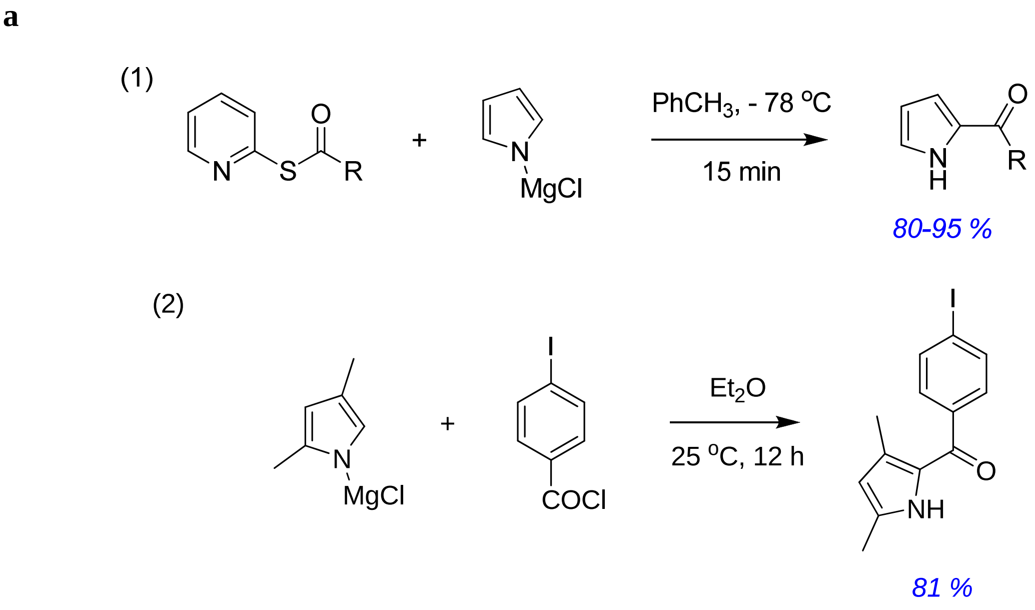
α,β-Unsubstituted BODIPYs (*eg* $R^1 = R^2 = R^3 = H$ in Scheme 3.1.c) can also be prepared from aldehydes using neat conditions.[106] The aldehyde is dissolved in excess pyrrole at room temperature, and the dipyrromethane intermediate (the reduced form of the dipyrromethene) was formed and isolated. The BODIPY dye was obtained after oxidation with DDQ and complexation with boron. Acid chlorides probably would be too reactive to use with pyrrole (since it is unsubstituted and more reactive) so this aldehyde-based approach is the method of choice.

It is unusual to use any aldehydes that are not aromatic to prepare BODIPY derivatives. In this sense the vinylic thioether probe as shown in the following reaction was exceptional. A catalytic amount of ytterbium (III) trifluoromethane sulfonamide was used to mediate the condensation process, the intermediate dipyrromethane was oxidized with DDQ, and complexation with boron trifluoride gave the product, though in very poor yield.[107]


SPh
+
CHO
(i) Yb(NTf$_2$)$_3$
(ii) DDQ
(iii) Et$_3$N, BF$_3$OEt$_2$
CH$_2$Cl$_2$, 25 °C
N
B
F$_2$
2 %

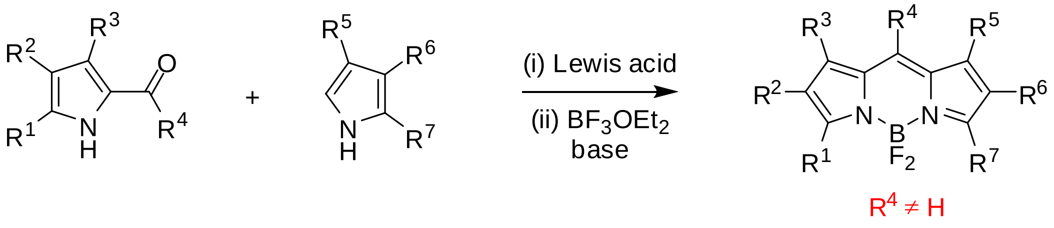
Condensations of pyrroles with acid chlorides or with benzaldehyde derivatives, as outlined above, are direct and convenient methods to access symmetrically substituted BODIPY dyes. However, another approach is required to form unsymmetrically substituted ones. Generally, this is achieved via preparations of ketopyrrole intermediates, followed by a Lewis acid mediated condensation of these with another pyrrole fragment. Schemes 3.2.a shows reactions of magnesium anions of pyrroles with activated carboxylic acid derivatives to give the corresponding 2-ketopyrroles.[108,109] In part b, one such ketopyrrole is condensed with another pyrrole unit to give a BODIPY framework. This method is particularly valuable for unsymmetrical ones.

Scheme 3.2. (**a**) Two different methods for production of ketopyrroles from magnesium derivatives of pyrrole, and (**b**) application of these starting materials in the production of unsymmetrical BODIPYs.


a
(1)
N
S
O
R
+
N
MgCl
PhCH$_3$, - 78 °C
15 min
N
H
O
R
80-95 %
(2)
N
MgCl
+
I
COCl
Et$_2$O
25 °C, 12 h
I
O
NH
81 %

Scheme 3.2. Continued.

b


R2
R3
O
R1
N H
R4
+
R5
R6
N H
R7
(i) Lewis acid
(ii) BF3OEt2
base
R3
R4
R5
R2
R6
N
B
N
R1
F2
R7
R4 ≠ H

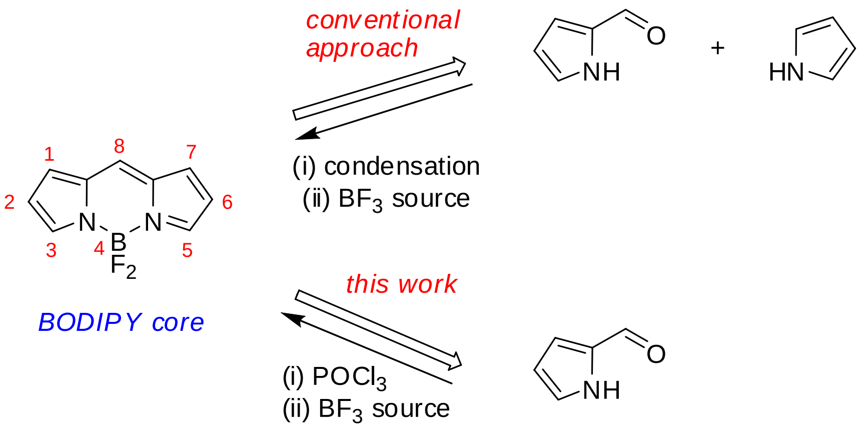
The syntheses of BODIPYs without *meso*-substituents (*eg* R^4 = H in Scheme 3.2.b) are usually achieved by Lewis acid mediated condensation of pyrrole-2-CHO with another pyrrole molecule that is not substituted at the 2-position. The conditions are very similar to the ones shown in Scheme 3.2.b. Only difference is that, in order to synthesize meso-unsubstituted BODIPYs, a pyrrole-2-CHO is used in place of the ketopyrrole. Many simple BODIPY derivatives are prepared using this method.[98]

This work describes a serendipitous discovery that symmetric BODIPY derivatives without *meso*-substituents can be prepared from the pyrrole-2-CHO component alone. Further the yield can be superior for this new approach. Thus the specific aims for this project are:

(i) optimization of conditions for the new reaction

(ii) investigation of the mechanism and kinetics of the new transformation

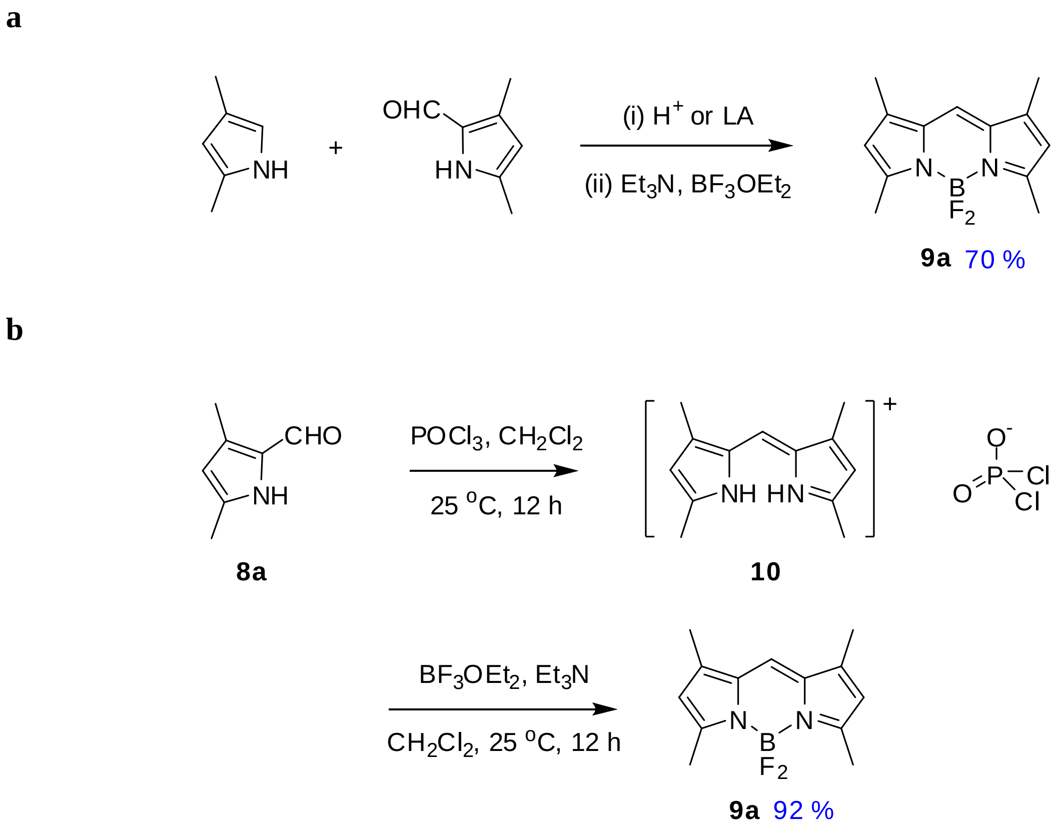
(iii) preparation of BODIPY derivatives using this new approach


conventional approach
NH
O
+
HN
(i) condensation
(ii) BF3 source
1
8
7
2
6
N
B
N
3
4
5
F2
BODIPY core
this work
(i) POCl3
(ii) BF3 source
NH
O

3.2 A new approach to symmetric BODIPYs

When 3,5-dimethylpyrrole-2-carboxaldehyde **8a** in CH_2Cl_2 was treated with $POCl_3$ as indicated in Scheme 3.3.b, a yellow solution resulted, and evolution of gas was observed. ^1H, ^{13}C, ^{31}P NMR and MS analyses of the crude reaction mixture showed that it contained the known[110-112] dipyrromethenium cation **10** and the dichlorophosphate counterion. This intermediate **10** was converted to the BODIPY product **9a** without isolation by adding BF_3•OEt_2 to the reaction mixture.

Scheme 3.3. (**a**) Conventional synthesis of tetramethyl-BODIPY **9a**; and, (**b**) the new approach described here.


a
NH
+
OHC
HN
(i) H+ or LA
(ii) Et3N, BF3OEt2
N
B
N
F2
9a 70 %
b
CHO
NH
POCl3, CH2Cl2
25 °C, 12 h
NH HN
+
O-
O=P-Cl
Cl
8a
10
BF3OEt2, Et3N
CH2Cl2, 25 °C, 12 h
N
B
N
F2
9a 92 %

The reaction occurred efficiently with 3,5-dimethylpyrrole-2-carboxaldehyde **8a** (0.2 M in CH_2Cl_2) using only 1.2 equivalents of $POCl_3$ at room temperature. Almost complete conversion of the starting pyrrole-2-CHO **8a** to the dipyrromethene intermediate **10** was observed within 30 min and no obvious by-product was detected by NMR. After installing the boron center, the desired product tetramethyl BODIPY **9a** was obtained in 92 % yield which was higher than the conventional approach yield (70 %). The reaction with **8a** was also carried out on 1-3 grams scale and the product **9a** was isolated with similar yields by recrystallization. Thus no further optimization of the new reaction condition was performed.

Scheme 3.4. Synthesis of pyrrole-2-aldehyde **8**.

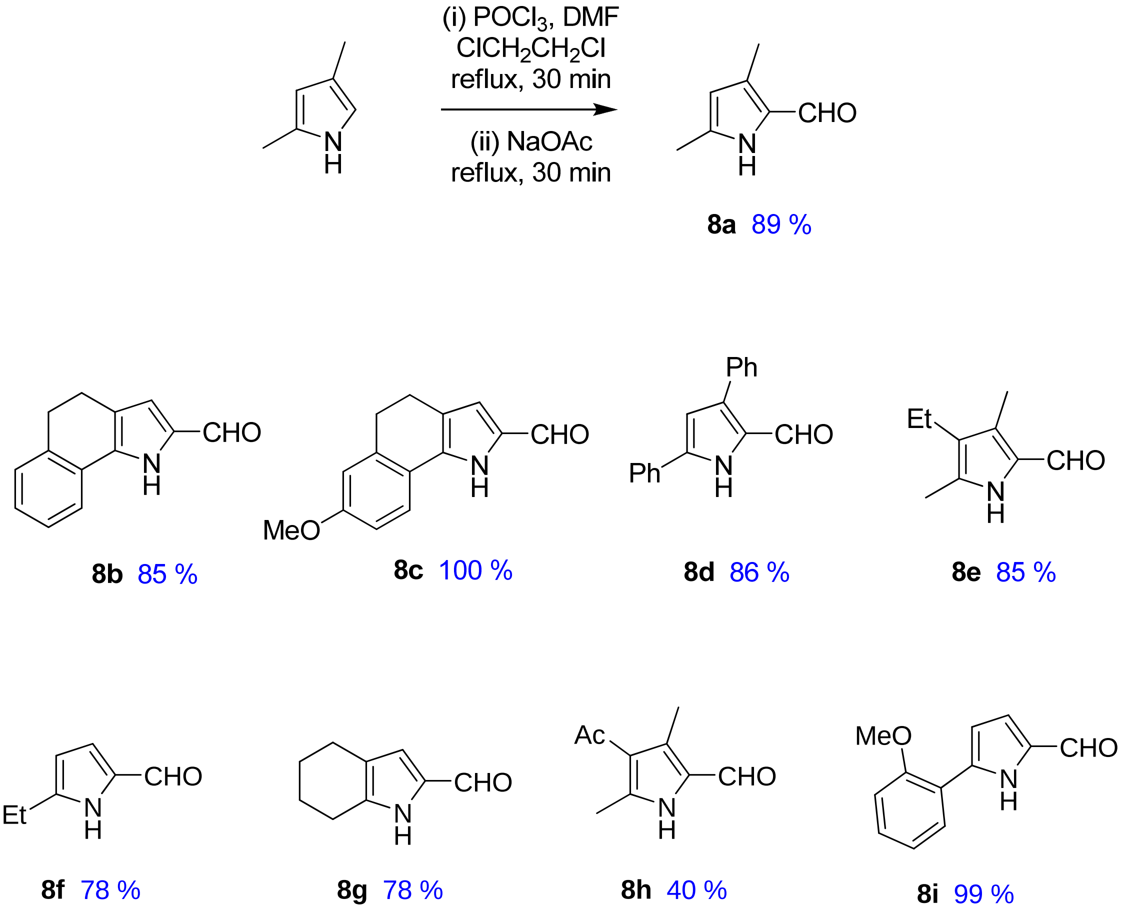
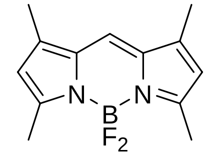
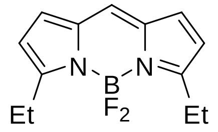
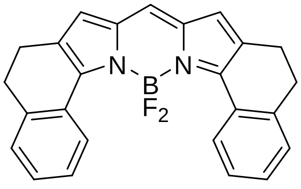
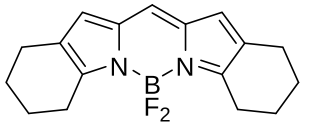
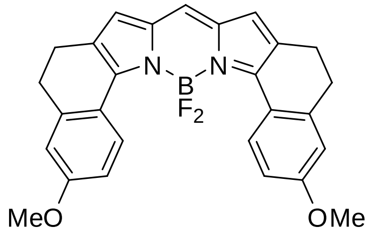
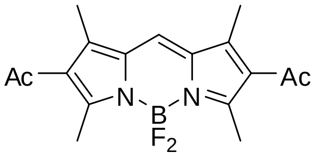
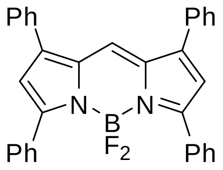
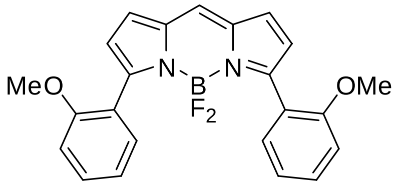
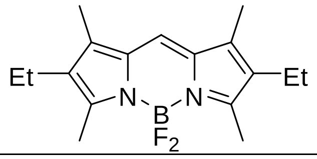

(i) POCl3, DMF
ClCH2CH2Cl
reflux, 30 min
(ii) NaOAc
reflux, 30 min
N
H
CHO
8a 89 %
8b 85 %
MeO
8c 100 %
Ph
Ph
8d 86 %
Et
8e 85 %
Et
8f 78 %
8g 78 %
Ac
8h 40 %
MeO
8i 99 %

Table 3.1. Synthesis of symmetric BODIPYs **9**

S.M	BODIPY product **9**	Yield (%)	S.M	BODIPY product **9**	Yield (%)
8a		92	**8f**		28
8b		91	**8g**		22
8c		85	**8h**		0^b
8d		21	**8i**		0^b
8e		75^a			

[a] Literature reported yield was 45 %;[113,114] [b] complex mitures; trace amount of desired product.

The pyrrole-2-aldehyde substrates needed for the new reaction were prepared via Vilsmeyer-Haack reaction from the corresponding pyrroles in good yields (Scheme 3.4.). The standard new reaction condition (0.1-0.2 M substrate in CH_2Cl_2, 1.2 equivalents of $POCl_3$, room temperature, overnight; then addition of BF_3OEt_2) was used to prepare series of BODIPY derivatives **9** from the pyrrole-2-CHO substrates **8** (Table 3.1.). Various substituents were tolerated in the synthesis. Product yields were substrate dependent and tended to be good except where there were steric issues from a C^3-

substituent (compare **9b** and **9d**), where the pyrrole was less electron rich (compare **9e** and **9h**), or for reasons which cannot be explained simply (**9f**, **9g** and **9i**). Further optimization of the reaction condition with respect to each substrate might improve the product yield.

Table 3.2. Photophysical properties of BODIPYs **9** in CH_2Cl_2

	$\lambda_{max\ abs}$ (nm)	ε ($M^{-1}cm^{-1}$)	$\lambda_{max\ emiss}$ (nm)	fwhm[a] (nm)	Φ_{fl}
9a	505	83200	516	-	0.80[b]
9b	629	161500	637	23	0.78 ± 0.03[c]
9c	651	149100	662	27	0.71 ± 0.04[c]
9d	564	78000	593	-	1.00[d]
9e	531	87000	538	21	0.86 ± 0.03[e]
9f	512	110200	516	17	0.96 ± 0.02[e]
9g	547	92000	553	20	0.90 ± 0.02[e]

[a] Full width at half maximum height of fluorescence (fwhm); [b] data were obtained from reference;[115] [c] Nile blue (Φ 0.27 in EtOH)[116] was used as a standard; [d] data were obtained from reference[117]; [e] Rhodamine 6G (Φ 0.94 in EtOH)[118] was used as a standard.

Characteristically, all the new BODIPY dyes emit bright, sharp fluorescent signals (Table 3.2. and Figure 3.1.). Substituents at 3 and 5 positions of BODIPY shift the absorption and emission to the red (compare **9a**, **9f** and **9e**). Aryl substituents dramatically shift the absorption and emission maxima to the longer wavelengths (compare **9a** and **9d**). Constrained derivatives gave even more red-shift (compare **9e** and **9g**, **9d** and **9b**, **9c**).

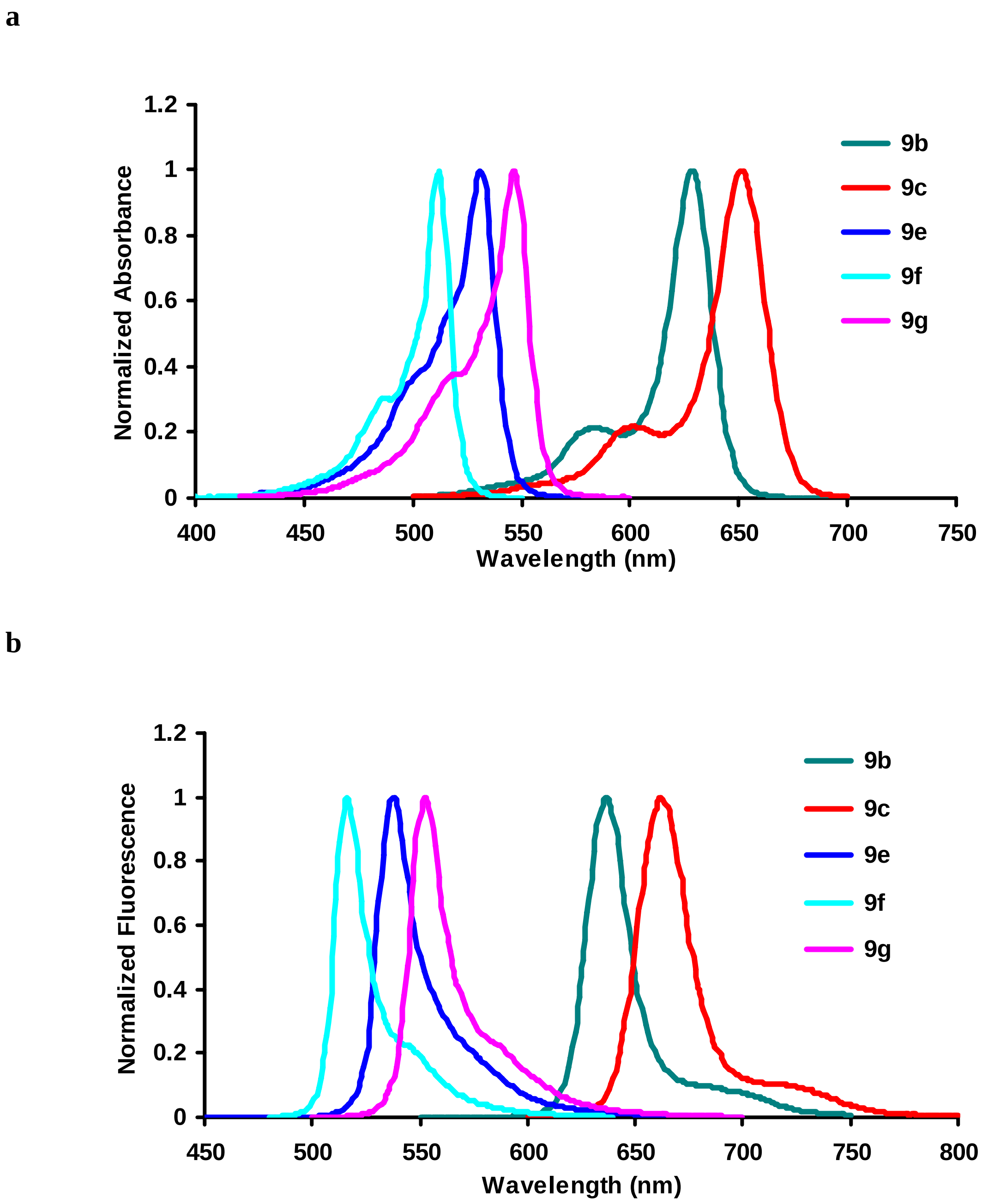


Figure 3.1. Spectra of BODIPYs **9** in CH_2Cl_2. **a** absorbance (10^{-6} M); **b** fluorescence (10^{-7} M, excited at $\lambda_{\text{max abs}}$).

3.3 Kinetic and mechanistic studies of the new reaction

3,5-dimethylpyrrole-2-carboxaldehyde **8a** was chosen as the substrate to study the kinetics and mechanism of the new reaction. The self-condensation process from **8a** to **10** as shown in Scheme 3.1.b (but in $CDCl_3$, at 0.2 M of substrate) was followed by ^1H, ^{13}C, and ^{31}P NMR over 10 min intervals at room temperature. More than 95 % conversion was observed in less than 20 min, and no by-products were observed by NMR (see Appendix Information). Gas evolution occurred, possibly due to loss of CO and HCl from the system.

The reaction shown in Scheme 3.1.b was also followed via continual UV measurements. This experiment required a considerably more dilute solution (6.7×10^{-5} M). Excess phosphorus oxychloride was used in this experiment, so the concentration of that reagent was effectively invariant throughout the transformation. Absorbance peaks for phosphorus oxychloride, the pyrrole-2-carboxaldehyde **8a** and the product **10** were well resolved under these conditions (Figure 3.2.a). As the reaction progressed the concentration of the pyrrole-2-CHO **8a** declined and the dipyrromethene **10** was formed. No intermediates or byproducts were detected assuming no coincident UV absorption characteristics (Figure 3.2.b).

Kinetic data extracted from the UV experiment are plotted in Figure 3.2.c. Over the first 20 min of the transformation it is second order with respect to the pyrrole-2-CHO **8a** ($6.43\ L\bullet mol^{-1}\bullet s^{-1}$). This analysis assumes that there were no intermediates having coincident UV absorption characteristics.

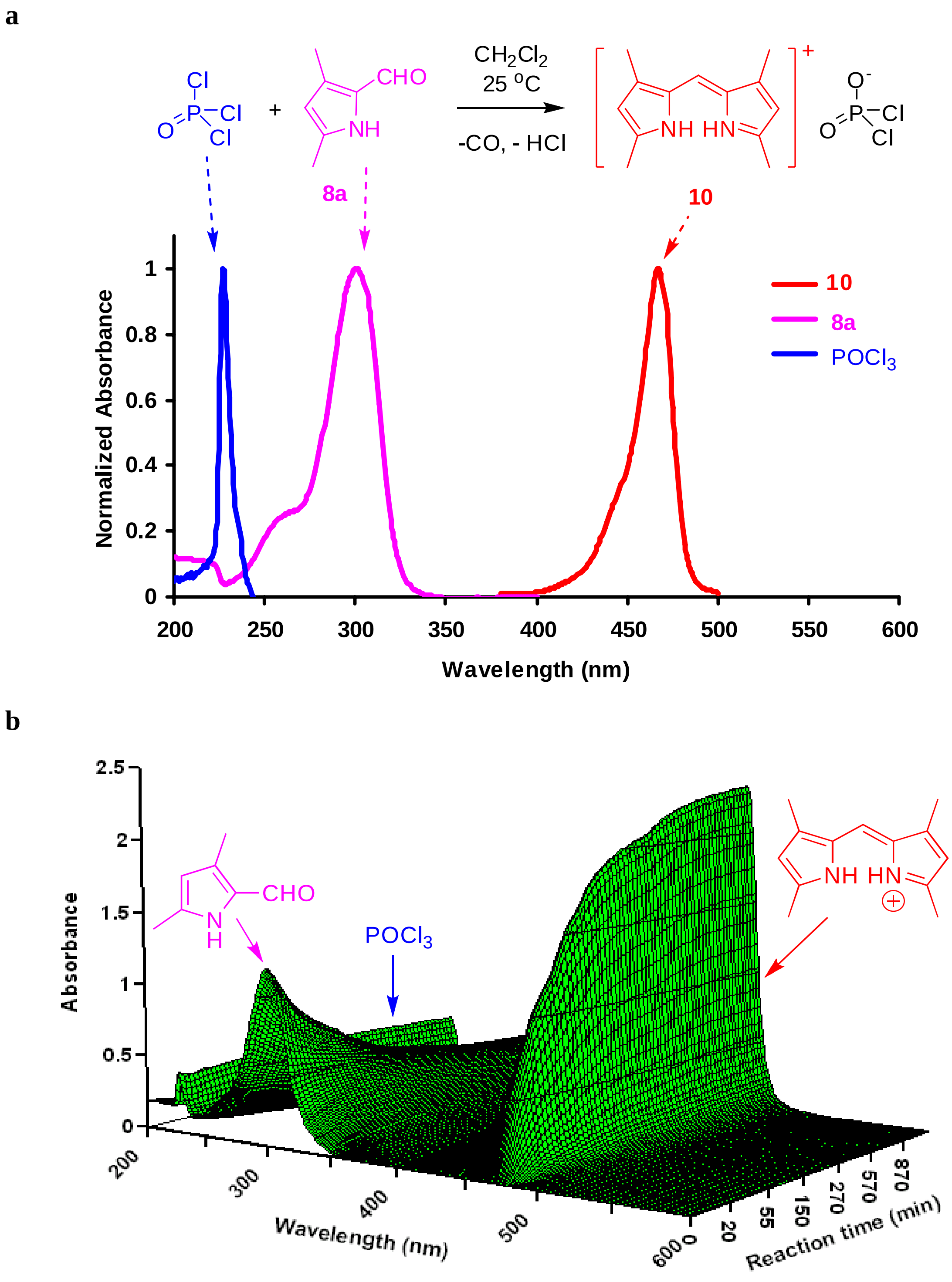


Figure 3.2. UV study of the self-condensation reaction in CH$_2$Cl$_2$ at 25 °C, and 6.7 x 10^{-5} M of pyrrole-2-CHO **8a**. (**a**) Reaction components and their UV spectra; (**b**) UV profile for the reaction; (**c**) plot used to deduce the rate of the reaction.

c

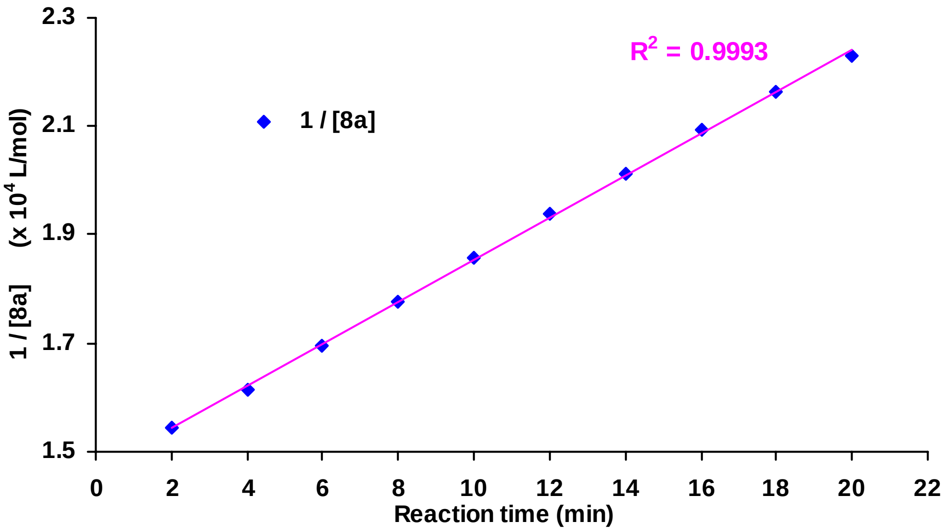


Figure 3.2. Continued.

The data collected on this reaction type was quite extensive, but still no clearly favored mechanism emerges. All the evidence indicates that phosphorus oxychloride seems to be intimately involved in this reaction; when this reagent is replaced with another Lewis acid (specifically, TFA or BF_3•OEt_2) the dipyrromethenium cation **10** did not form. One possibility is that $POCl_3$ reacts with pyrrole-2-carboxyaldehyde to give a vinylogous Vilsmeyer-Hack reagent (Figure 3.3.a, where alternatively the chloride could be a *P*-based leaving group). Formation of the product then relies on electrophilic attack of this reagent on the C^2-atom of another pyrrol-2-carboxaldehyde molecule. However, it is somewhat surprising if this is the regioselectivity of the attack in that step for some of the substrates. Another possiblity is that in the presence of $POCl_3$, pyrrole-2-aldehyde deformylated to generate the free pyrrole which condensated with another molecule of pyrrole-2-aldehyde to give the product (Figure 3.3.b). However, this transformation is quite unusual and it is hard to draw a plausible mechanism for the deformylation step.

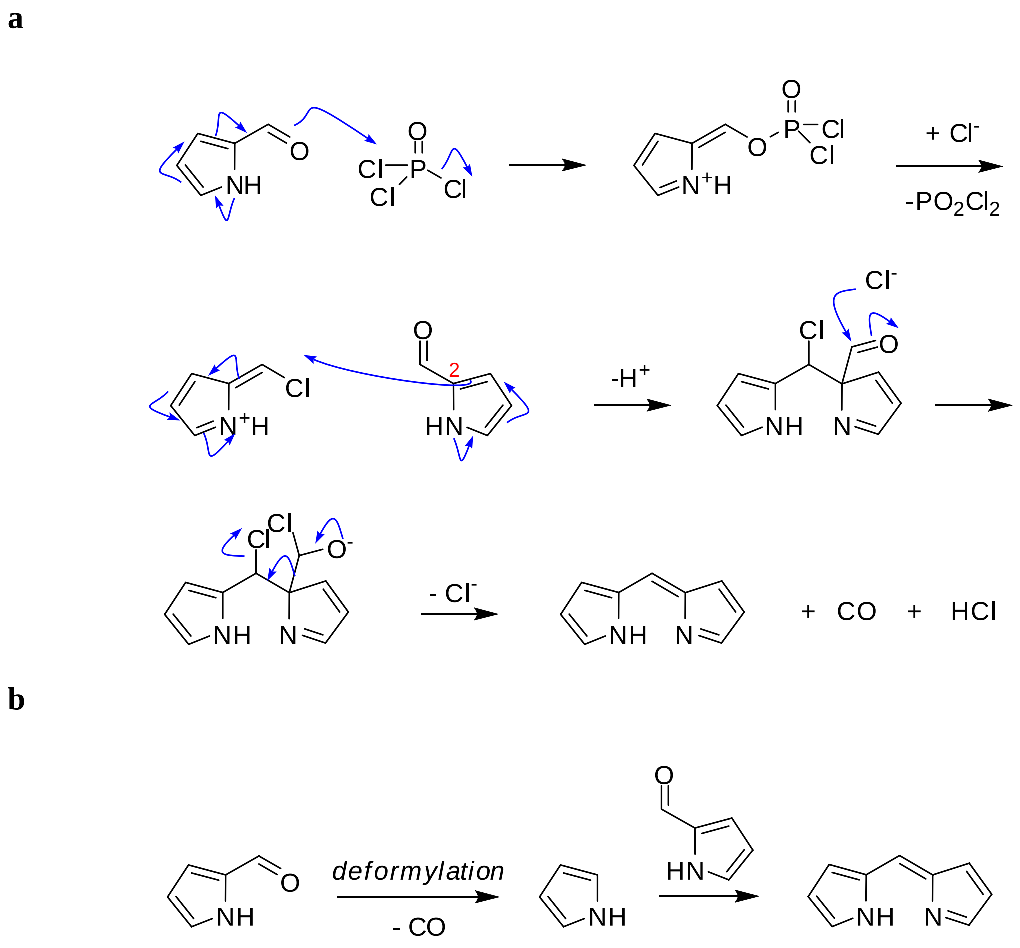


Figure 3.3. (**a**) and (**b**) Two possible mechanisms for the formation of dipyrromethene.

3.4 Conclusion

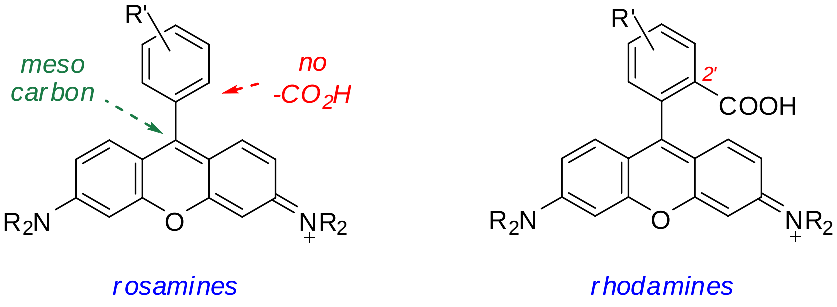
In summary, a novel synthesis of symmetric BODIPY dyes was discovered. The new reaction condition was optimized and gave superior yield of tetramethyl-BODIPY (92 %) than the conventional synthesis (70 %). Systematic investigation on the kinetics and mechanism for the new reaction was carried out. Several possible mechanisms were proposed for the transformation. The new condition was applied on various pyrrole-2-CHO substrates with a diversity of alkyl and aryl substituents and afforded the desired product in moderate to good yields. A few NIR BODIPY derivatives were also prepared using this new approach.

CHAPTER IV

ROSAMINES WITH CYCLIC AMINE SUBSTITUENTS: SYNTHESES, SPECTROSCOPIC PROPERTIES AND APPLICATIONS*

4.1 Introduction

Rhodamine and rosamine dyes are highly fluorescent xanthylium chromophores that have been widely used in many different fields of research, from the lasing medium in dye lasers to fluorescent markers in biological studies.[119] The structural difference between rhodamines and rosamines is that rhodamines carry an additional 2′-carboxylic acid which constrains the rotation of the meso-phenyl substituents and also offers a handle for attachment to biomolecules.



Noelting and Dziewonsky first reported the preparation of rhodamine in 1905.[120] Many rhodamine derivatives have been synthesized since then. Among them, Rhodamine B, rhodamine 6G, rhodamine 110, Texas Red and Alex Fluor dyes (Figure 4.1.) are the most commonly used ones. Their fluorescence characteristics are highly dependent on the substituents and ring conjugation. The *meso*-substituents are not significantly conjugated to the xanthylium core so that they do not have much effect on

---

*Reprinted in part with permission from "Synthesis and Spectroscopic Properties of Rosamines with Cyclic Amine Substituents", Liangxing Wu, Kevin Burgess, *J. Org. Chem.*, **2008**, *73*, 8711-8718. Copyright 2009 American Chemical Society.

the emission wavelength. But the nature and flexibility of the *meso*-substituents affect the fluorescence quantum yields and stability of the rhodamine derivatives. Substituents on the nitrogen will red-shift the absorption and emission. For example, rhodamine 110 absorbs at 502 nm and emits at 527 nm while the ethyl-substituted derivative (rhodamine B) absorbs *ca* 550 nm and emits *ca* 570 nm.[119] The additional conjugation of the xanthylium π-system, as well as formation of rigid rings in rhodamines can shift their absorption and emission maxima to longer wavelengths.[121] The rigidity shifts the emission of Texas Red to 615 nm and additional conjugation in Alexa594 also red-shifts the fluorescence to 617 nm.[122]

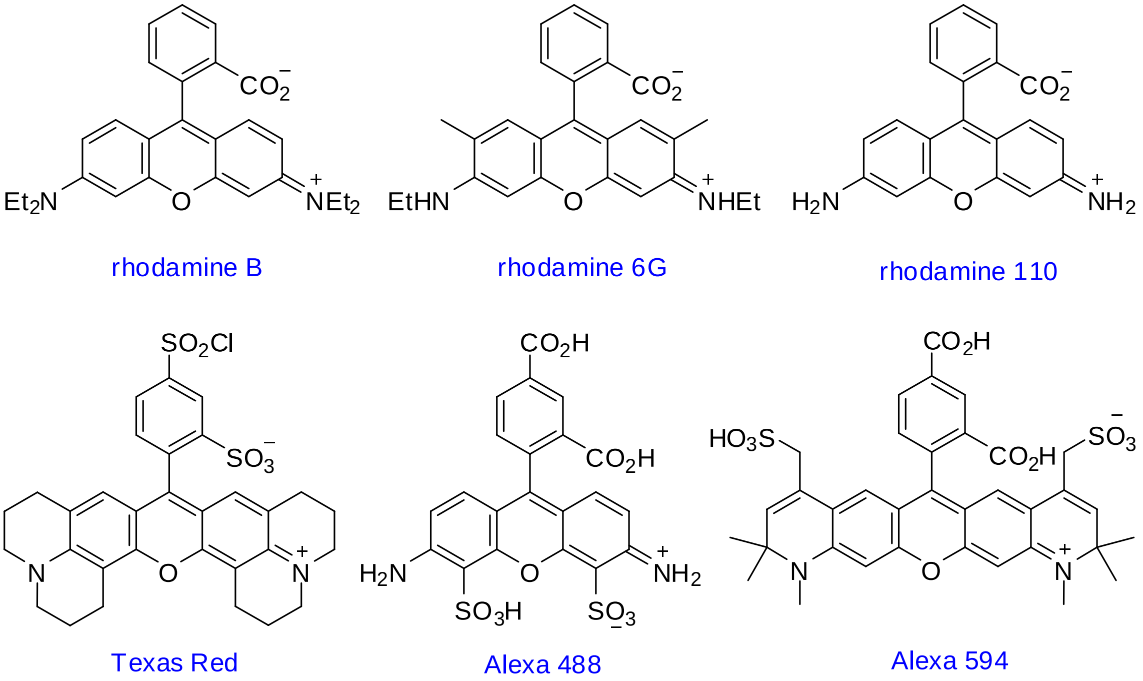
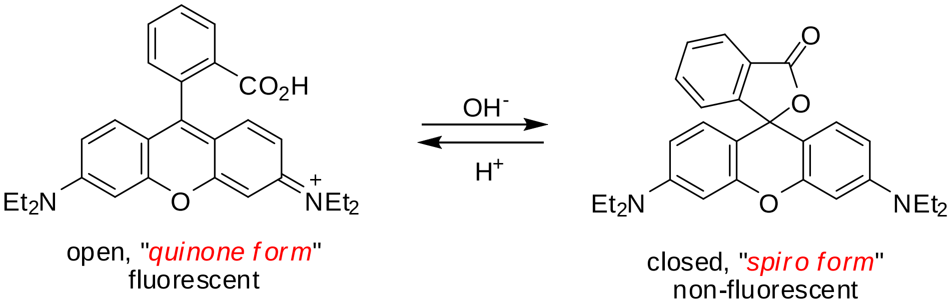

rhodamine B
rhodamine 6G
rhodamine 110
Texas Red
Alexa 488
Alexa 594

Figure 4.1. Structures of some common rhodamines.

Rhodamine dyes are highly favorable fluorophores because of their excellent photophysical properties, such as high extinction coefficient, excellent quantum yields (often near 1), great photostability and relatively long emission wavelength (> 550nm) compared to fluorescein.[119] Additionally, their fluorescence spectra are pH-independent over a broad range (low to neutral pH). In contrast to fluorescein, which exists in a non-fluorescent spiro form at acidic pH, rhodamines exist in a fluorescent quinine form at neutral and acidic pH, and a non-fluorescent spiro form only at strong basic condition.[123]



Due to its remarkably high photostability, high quantum yield and low cost, rhodamine 6G is often used as a laser dye and as a fluorescence quantum yield standard.[119] Rhodamine B is frequently used in biology as a staining fluorescent dye, sometimes in combination with auramine O, as the auramine-rhodamine stain (AR). The AR is a histological technique used to visualize acid-fast bacilli notably species in the Mycobacterium genus using fluorescence microscopy.[124] Acid-fast organisms display a reddish-yellow fluorescence. Rhodamine B can also be used as a layer dye and fluorescence quantum yield standard. Rhodamine 110 especially its sulfonated analog Alexa 488 is used extensively for biological labeling in place of fluorescein since they absorb and emit at similar wavelengths but are far more photostable. In fact, Alexa Fluor dyes are probably the most used rhodamine derivatives in biochemistry and cell biology. Alexa dyes are highly water-soluble which facilitate the applications under physiological conditions. Furthermore, they give exceptionally bright and photostable conjugates with biomolecules. The fluorescence emission even of dyes with the highest possible quantum yield and great water-solubility, such as fluorescein and sulforhodamine (Texas Red), is considerably quenched on conjugation to biological molecules, particularly to proteins. However, conjugation does not affect the optical properties of Alexa dyes. Actually, diverse protein conjugated of Alexa dyes tend to be brighter and more resistant to photobleaching than those of any alternative dyes that have similar spectral properties.

In addition, Alexas dyes are insensitive to pH in the 4-10 range. All the above characteristics facilitate the wide use of Alexa dyes in biological studies.[122]

Rhodamine derivatives can also be used as substrates for the multidrug resistance (MDR) transporter assays.[125,126] Multidrug resistance (MDR), mediated by the plasma membrane protein P-glycoprotein[127] (Pgp), has been identified as a major impediment to successful chemotherapy for a variety of cancers.[128] Pgp pumps a variety of structurally and mechanistically unrelated chemicals and chemotherapeutic agents from tumor cells, resulting in treatment failures.[129,130] The mechanism by which Pgp overexpression is induced during exposure to chemotherapeutics is still not fully understood.[131,132] Attempts have been made to alleviate the function of Pgp using modulators such as verapamil, quinidine, cuclosporin A, PSC833, and others.[133-135] These modulators bind to the drug-transport sites on the protein and impede the binding and subsequent efflux of agents of interest, thus allowing their accumulation. Studies of the efflux function of Pgp have traditionally use Pgp transport molecules that are either intrinsically fluorescent or tagged with a fluorescent marker. Among the best transport molecules for Pgp are the xanthylium dyes rhodamine 123 (the methyl ester of rhodamine 110), rhodamine 6G and tetramethylrosamine.[136-138] Such studies usually monitor the efflux of a specific molecule, such as rhodamine 123, from cells by measuring the decrease in intracellular fluorescence.[125,131] These dyes have also been used as fluorescent markers to determine the efficacy of Pgp modulators.[139]

Rhodamine 123 has also been extensively used as a mitochondrial stain for studying cellular functions.[140] Mitochondria are the main energy generators that maintain cell life and essential cell functions.[141] As a powerhouse, mitochondria generate energy through oxidative phosphorylation where oxidation of respiratory substrates is coupled to the synthesis of ATP under aerobic conditions.[142] This process involves a sequence of electron transfers from respiratory substrates to oxygen, concurrent with proton translocation from the mitochondrial inner compartment to the intermembrane space through a series of respiratory chain complexes located on the inner membrane. The electrochemical proton gradient thus formed, also designated as the proton motive force, is the driving force for ATP synthesis through the back flow of protons through the ATP synthase complex. Importantly, this mitochondrial transmembrane proton-motive force which results in a negative potential inside the mitochondrial matrix selectively accumulates lipophilic cations such as rhodamine 123 which are membrane-permeable compounds with cationic characteristics.[143,144]

High concentrations of lipophilic cations in mitochondria often results in cell death by decreasing cellular ATP production,[145] rendering mitochondria a unique target for cellular toxicity.[143] Numerous studies have shown that the mitochondrial membrane potentials of carcinoma cells are higher than in normal epithelial cells and that the accumulation and retention of lipophilic cations correlated with the mitochondrial membrane potential. This increase in mitochondrial membrane potential in carcinoma cells which leads to selective accumulation of toxic lipophilic cations provides a rationale for selective chemotherapy of cancer cells. Rhodamine 123 was the first example of lipophilic cation to exhibit selective anti-tumour activity.[146-148] In *in vitro* experiments, this compound markedly induced cell death in 9/9 of carcinoma cell types while 6/6 of non-tumorigenic epithelial cell types remained unaffected when tested at similar concentrations.[149-151] Rhodamine 123 also displays selective toxicity in certain carcinoma cells *in vivo* and prolongs the survival of mice bearing experimentally induced Ehrlichs ascites tumor or MB49 bladder carcinoma, but not that of mice bearing L1210 leukemia or B16 melanoma.[152] The moderate selectivity and systemic toxicity limit the usefulness of rhodamine 123.

In spite of their wide applications, there are not many available rhodamine or rosamine derivatives. There is considerable potential for modulating the chemical and biochemical properties of rosamines by varying the substituents, but the scope of such research is limited by synthetic constraints. The first syntheses of rosamines center around condensation reactions of the type shown in Figure 4.2.a.[121,153-155] Such approaches tend to involve high temperatures. They can be facilitated via microwave heating,[156] but even in those cases the conversion to product is not very efficient. Other work deals with modification of the 2-carboxylate group of rhodamines (not shown),[123,157] but the parent systems are also made via condensation reactions at elevated temperatures. An alternative approach to syntheses of rosamine dyes is to add an organo-lithium or -magnesium compound to the carbonyl group of a xanthone, then protonate to cause loss of water and generate the product[158] (Figure 4.2.b). That approach has been applied extensively by Detty and co-workers to make tetramethylrosamine derivatives (and especially *S*-, *Se*-, and *Te*-analogs)[159-162] for their photodynamic and other medicinal properties.[125,126,163-165] Similarly, attack of aryl Grignards on solid-supported xanthones was used in a supported parallel synthesis approach to a library of rosamine derivatives.[166] However, the preparation of rosamine derivatives using the addition approach has been limited by the paucity of synthetic routes to 3, 6-diaminesubstituted-xanthones. The reported approaches to 3,6-diamino-xanthone usually require many step synthesis and only lead to low yield of the desired products.[166-172] Some routes only work for certain substrates and are not general[173-178] (Figure 4.3.).

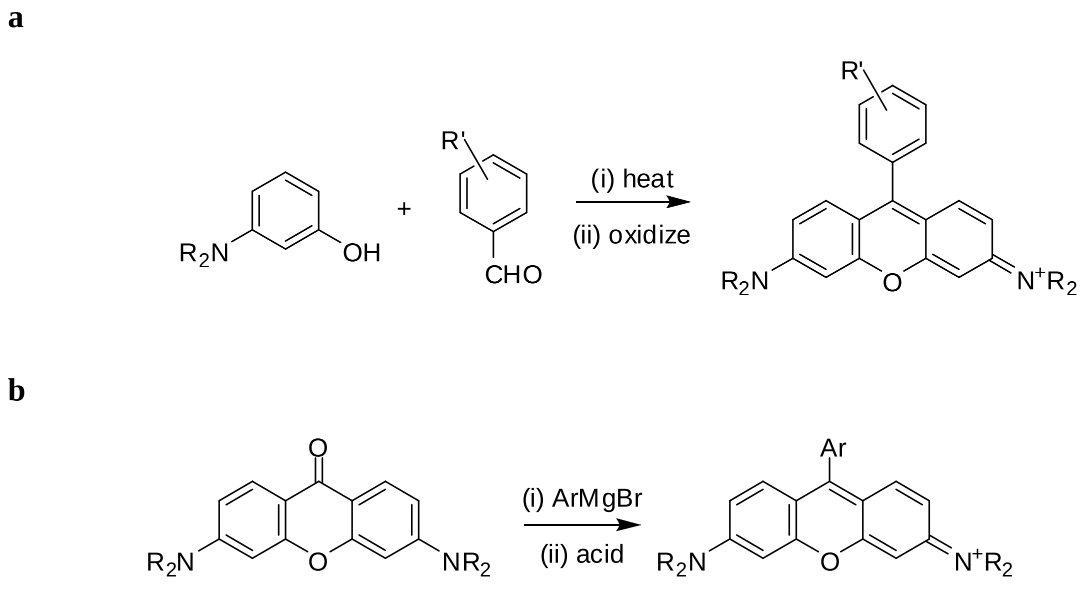


Figure 4.2. Syntheses of rosamine dyes via: (**a**) the conventional condensation approach; and, (**b**) nucleophilic additions to 3,6-diamino-xanthones.

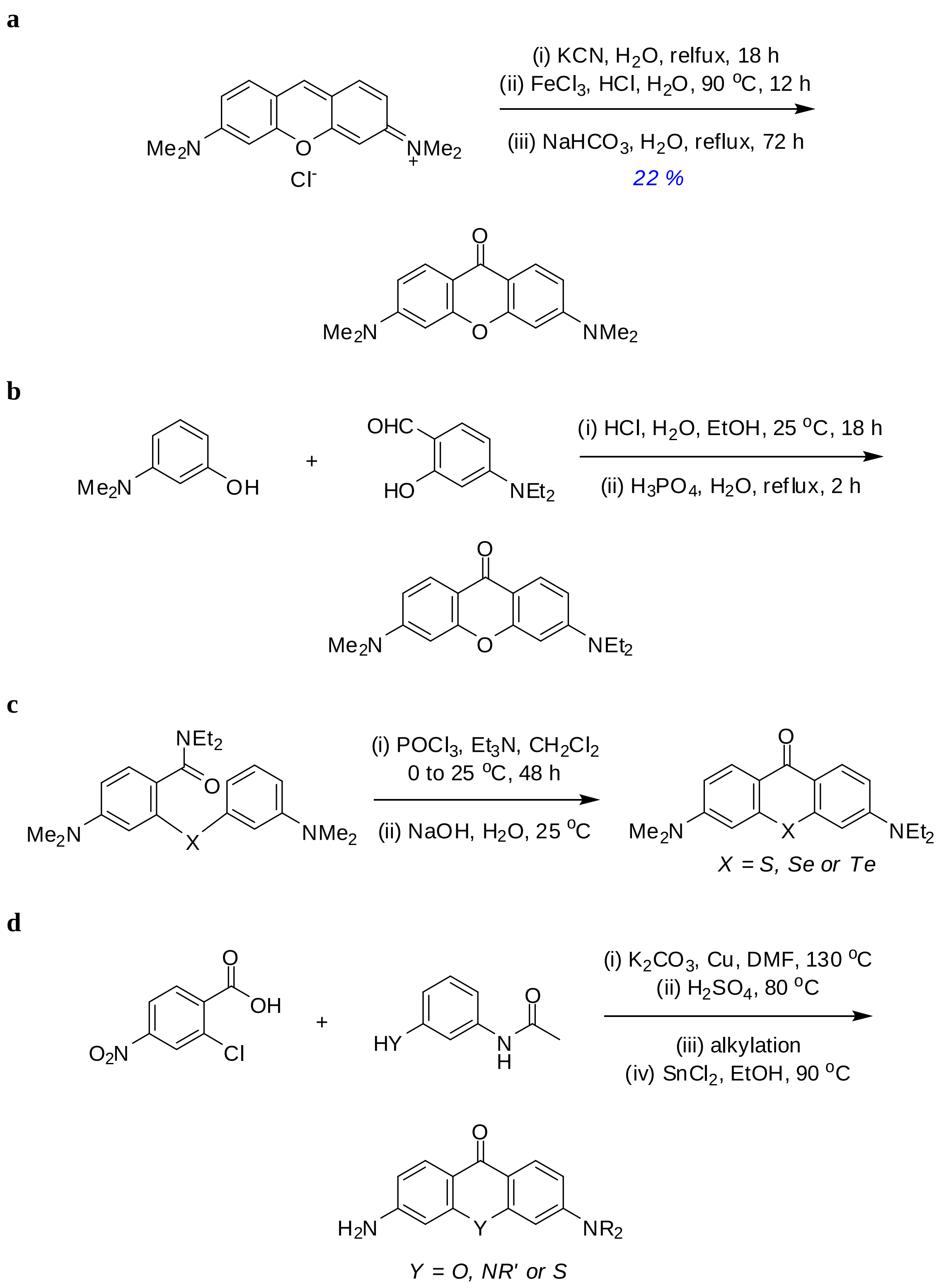

a
(i) KCN, H2O, relfux, 18 h
(ii) FeCl3, HCl, H2O, 90 °C, 12 h
(iii) NaHCO3, H2O, reflux, 72 h
22 %
Me2N
O
NMe2
+
Cl-
O
Me2N
O
NMe2
b
Me2N
OH
+
OHC
HO
NEt2
(i) HCl, H2O, EtOH, 25 °C, 18 h
(ii) H3PO4, H2O, reflux, 2 h
O
Me2N
O
NEt2
c
NEt2
O
Me2N
X
NMe2
(i) POCl3, Et3N, CH2Cl2
0 to 25 °C, 48 h
(ii) NaOH, H2O, 25 °C
O
Me2N
X
NEt2
X = S, Se or Te
d
O
OH
O2N
Cl
+
HY
O
N
H
(i) K2CO3, Cu, DMF, 130 °C
(ii) H2SO4, 80 °C
(iii) alkylation
(iv) SnCl2, EtOH, 90 °C
O
H2N
Y
NR2
Y = O, NR' or S

Figure 4.3. Reported syntheses of 3,6-diaminesubstituted-xanthones.

For several years, work in our laboratory has focused on a solution phase approach to the rosamine dyes. This chapter shows the efficient synthesis of rosamine cores **11** – **15** with a variety of *meso*-substituents **a** - **i**. The specific aims for this work are:

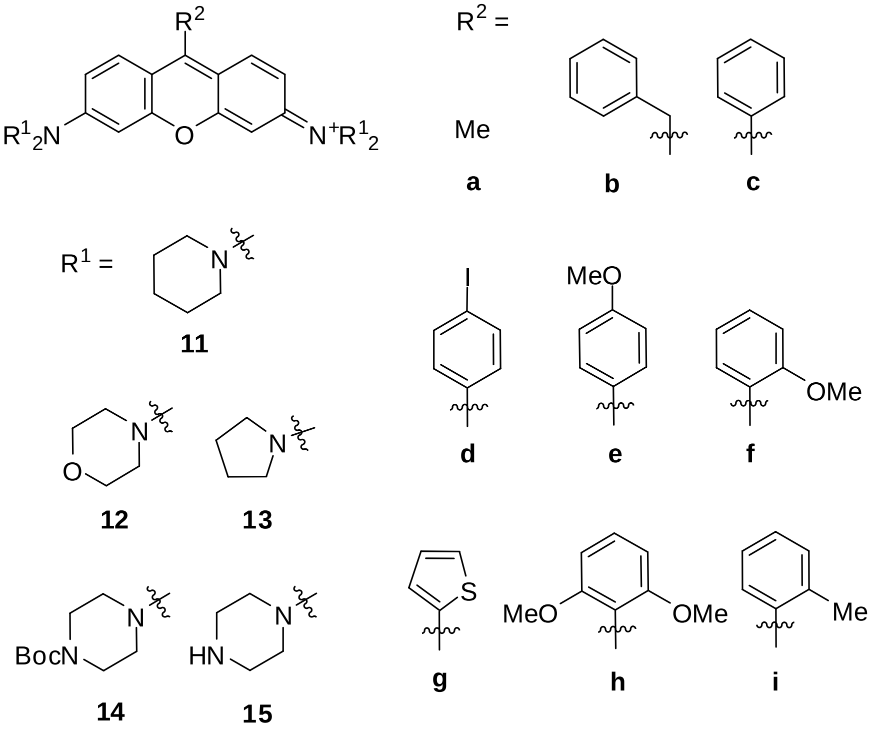
(i) development of an efficient synthesis of 3, 6-diamino-xanthone

(ii) preparation of rosamines from the synthesized aminoxanthones

(iii) functionalization to prepare water-soluble derivatives

(iv) photophysical property studies and applications

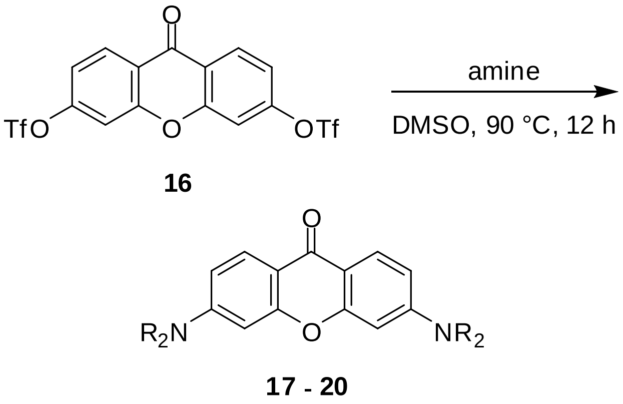
(v) biological activity test for the new rosamines

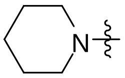
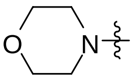
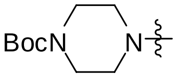
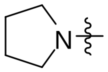

R2
R12N
O
N+R12
R2 =
Me
a
b
c
R1 =
N
11
I
MeO
OMe
d
e
f
O
N
N
12
13
S
MeO
OMe
Me
g
h
i
BocN
N
HN
N
14
15

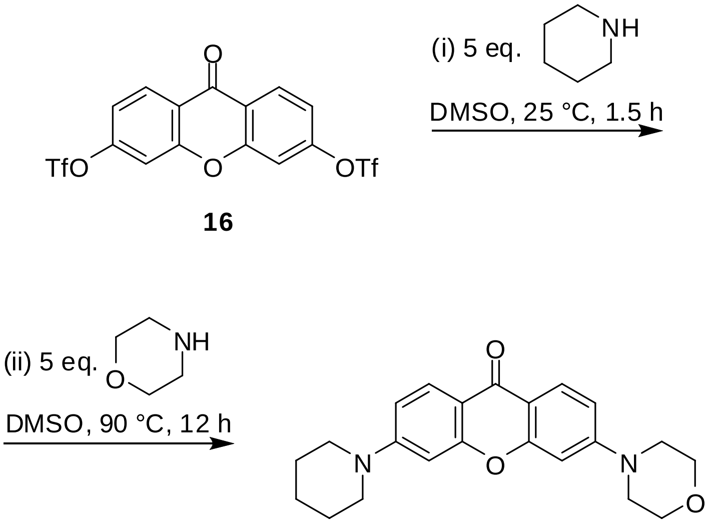
4.2 Syntheses of aminoxanthones and rosamines

The ditriflate **16** was easily prepared on a multigram scale without chromatography via a slight modification of a known procedure.[179,180] This reacted with piperidine, morpholine, Boc-piperazine, and pyrrolidine to give the symmetrical products **17** – **20** as indicated in Table 4.1. Stepwise addition to give an unsymmetrical product was also explored briefly, and this proved to be feasible (reaction 4.1), presumably because the product of the first addition (a vinylogous amide) is less reactive than the starting material.

Table 4.1. Amination of the ditriflyl xanthone **16**



product	NR_2	yield (%)
17		89
18		93
19		60
20		18



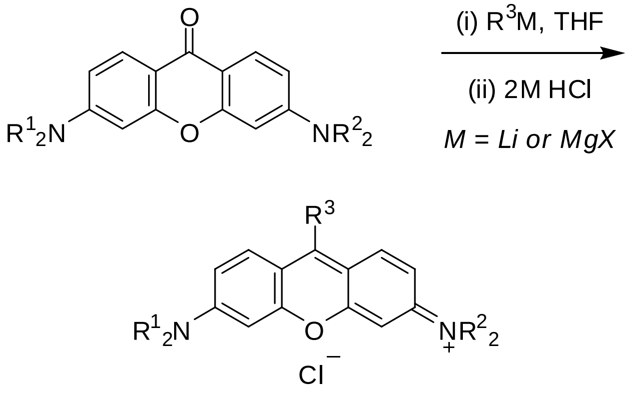
16

21 84 %

reaction 4.1

Ketones **17** to **21** are not exceptionally reactive, they are vinylogous ureas, but they combined efficiently with organolithium and Grignard reagents to give tertiary alcohols. These adducts were not isolated; instead they were treated with aqueous acid to give the desired products as indicated in Table 4.2. All the compounds were isolated in tens to hundreds of milligram amounts via flash chromatography. However, in three cases (**11d**, **11h** and **14d**) several grams of product were isolated via recrystallization of crude materials obtained on a larger scale. The benzenoid compounds shown in Table 4.2 have methyl-, iodo-, and one or two methoxy substituents on the *meso*-aryl group. Steric effects associated with *ortho*-substitution did not seem to have any significant adverse effects on the syntheses of any compounds in this series.

Table 4.2. Synthesis of rosamine derivatives


O
R1 2N
NR2 2
(i) R3M, THF
(ii) 2M HCl
M = Li or MgX
R3
R1 2N
O
NR2 2
+
Cl−

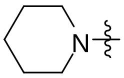
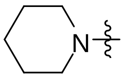
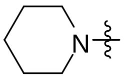
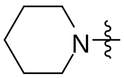
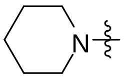
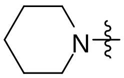
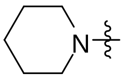
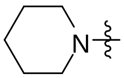
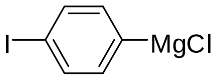
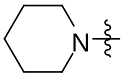
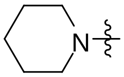
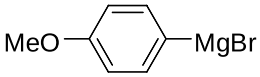
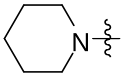
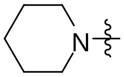
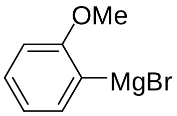
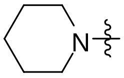
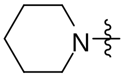
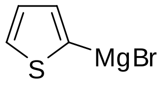
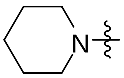
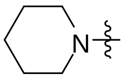
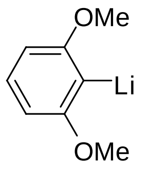
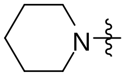
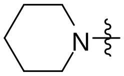
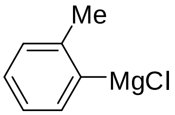
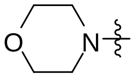
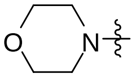
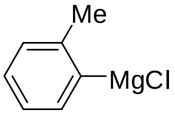
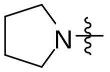
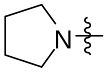
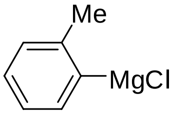
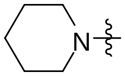
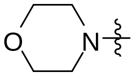
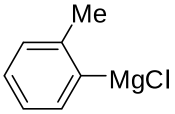
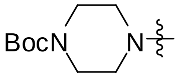
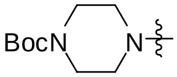
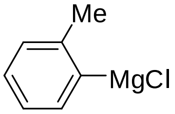
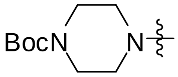
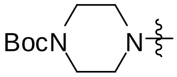
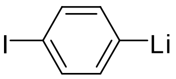
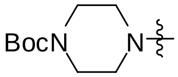
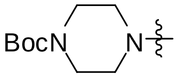
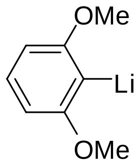
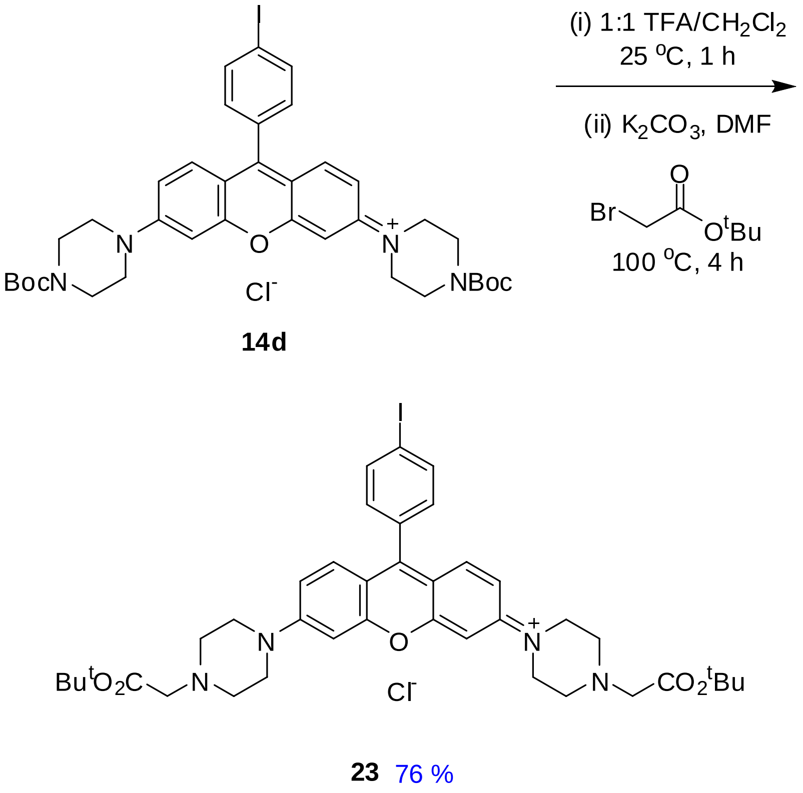
entry	NR^1_2	NR^2_2	R^3M	product	yield (%)
1			CH_3MgBr	**11a**	89
2			$PhCH_2MgCl$	**11b**	70
3			PhMgBr	**11c**	100
4				**11d**	90
5				**11e**	100
6				**11f**	100
7				**11g**	74
8				**11h**	96

Table 4.2. Continued

entry	NR^1_2	NR^2_2	R^3M	product	yield (%)
9				**11i**	99
10				**12i**	98
11				**13i**	53
12				**22i**	98
13				**14i**	90
14				**14d**	68
15				**14h**	94

None of the compounds shown in Table 4.2. are water-soluble, but several modifications were tested to demonstrate that water soluble derivatives are accessible from some of the products. Reaction 4.2, for instance, illustrates how the piperazine derivative **14d** could be unmasked to give a bisammonium salt. That salt was not isolated but aqueous solution of it was strongly fluorescent, and it was neutralized then reacted with *tert*-butyl bromoacetate to give the diester **23**. Treatment of diester **23** with trifluoroacetic acid gave a product that was water-soluble and fluorescent but difficult to isolate. The inference of these experiments is that the intermediate ammonium salt and the diacid derived from it could be prepared via deprotection of their precursors then used to couple to biomolecules.



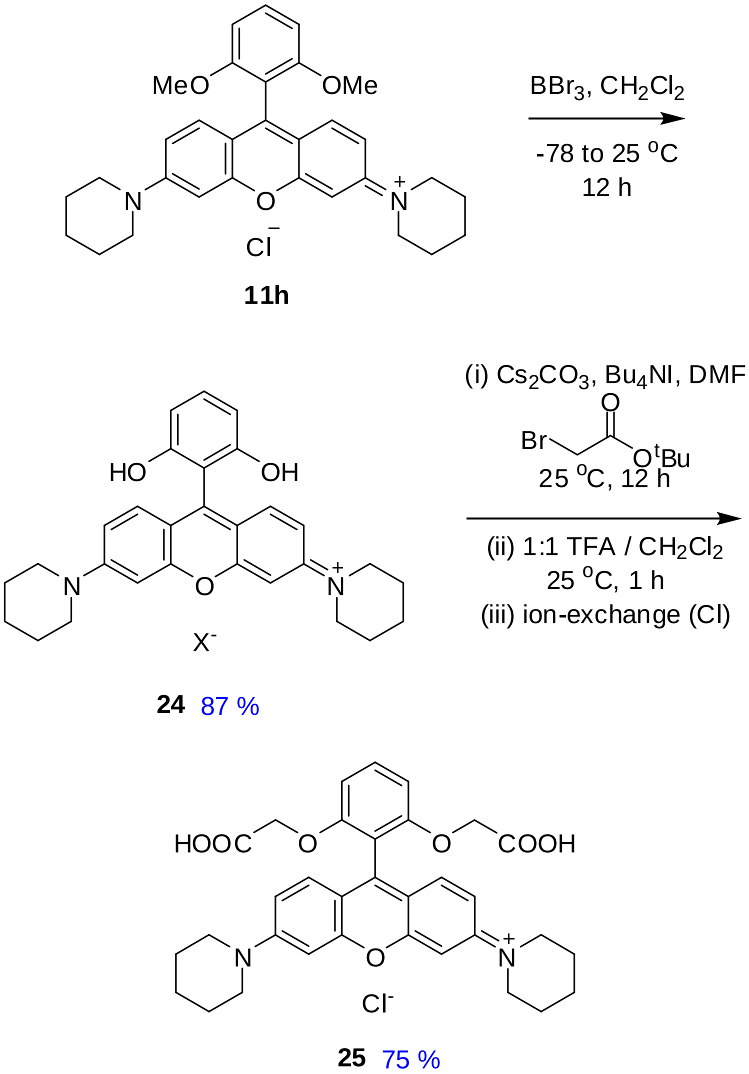
14d

23 76 %

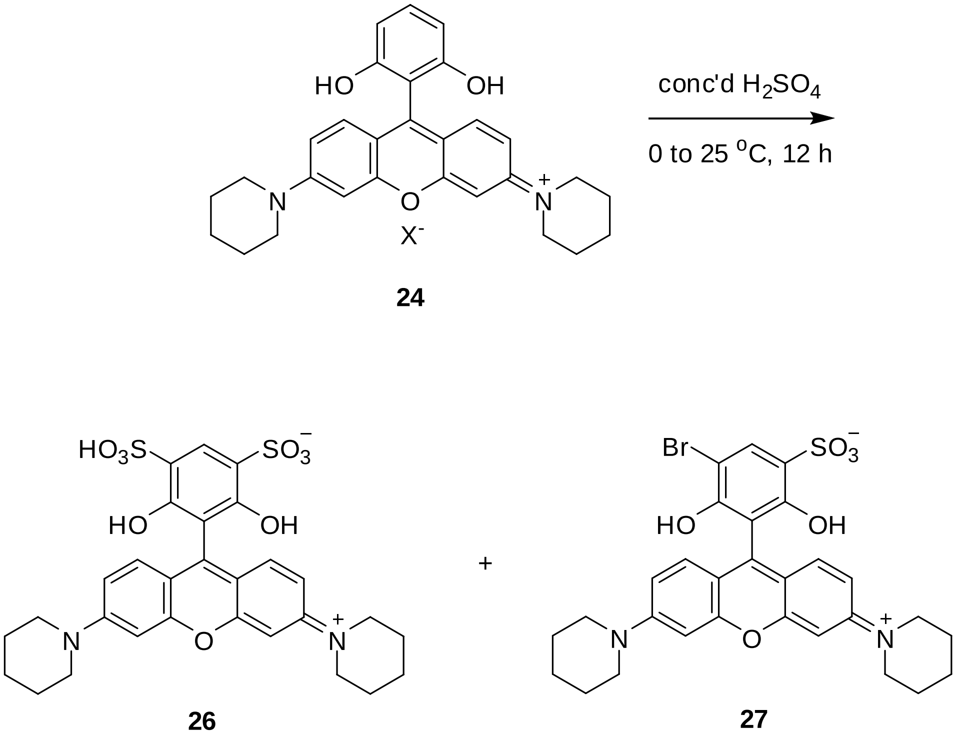
reaction 4.2

The dimethoxy compound **11h** was also modified to give a water soluble analog. Scheme 4.1. outlines how this was achieved via a demethylation procedure to give the phenolate **24**. That derivative has no functional group for direct attachment to a biomolecule, and in any case it is only slightly water-soluble. Consequently, the two phenolate groups were alkylated with *tert*-butyl bromoacetate, then the esters were deprotected to give the diacid **25**. Diacid **25** is a polar, water-soluble compound that was conveniently isolated by washing away residual starting materials using acetonitrile/ether. This derivative was conjugated to avidin to explore how its spectroscopic properties are modulated by coupling to this protein (*see below*).

Scheme 4.1. Synthesis of water-soluble rosamine **25**.


MeO
OMe
N
O
N
+
Cl−
11h
BBr3, CH2Cl2
-78 to 25 °C
12 h
(i) Cs2CO3, Bu4NI, DMF
O
Br
O^tBu
25 °C, 12 h
(ii) 1:1 TFA / CH2Cl2
25 °C, 1 h
(iii) ion-exchange (Cl)
HO
OH
N
O
N
+
X-
24 87 %
HOOC
O
O
COOH
N
O
N
+
Cl-
25 75 %

Interestingly, rosamine **24** was easily sulfonated in concentrated sulfuric acid to give a water-soluble derivative **26** along with an unexpected product **27** which was mono-brominated and sulfonated product probably arising from the mixture Cl/Br counterion (reaction 4.3). Both **26** and **27** were unambiguously assigned by NMR, MS and X-ray diffraction analysis (Figure 4.4.).



24

26

27

reaction 4.3

The X-ray structure shows that in rosamine **26** the two nitrogens are essentially flat and coplanar with the xanthene core structure. The piperidine rings adopt perfect chair conformations and the meso-phenyl substituent is almost perpendicular to the xanthene core structure (dihedral angle: 71.90 °) indicating that the meso-substituents are not conjugated with the rosamine cores. Similar conformations were also observed in rosamine **27**. Details of the X-ray structure parameters are listed in the Appendix information.

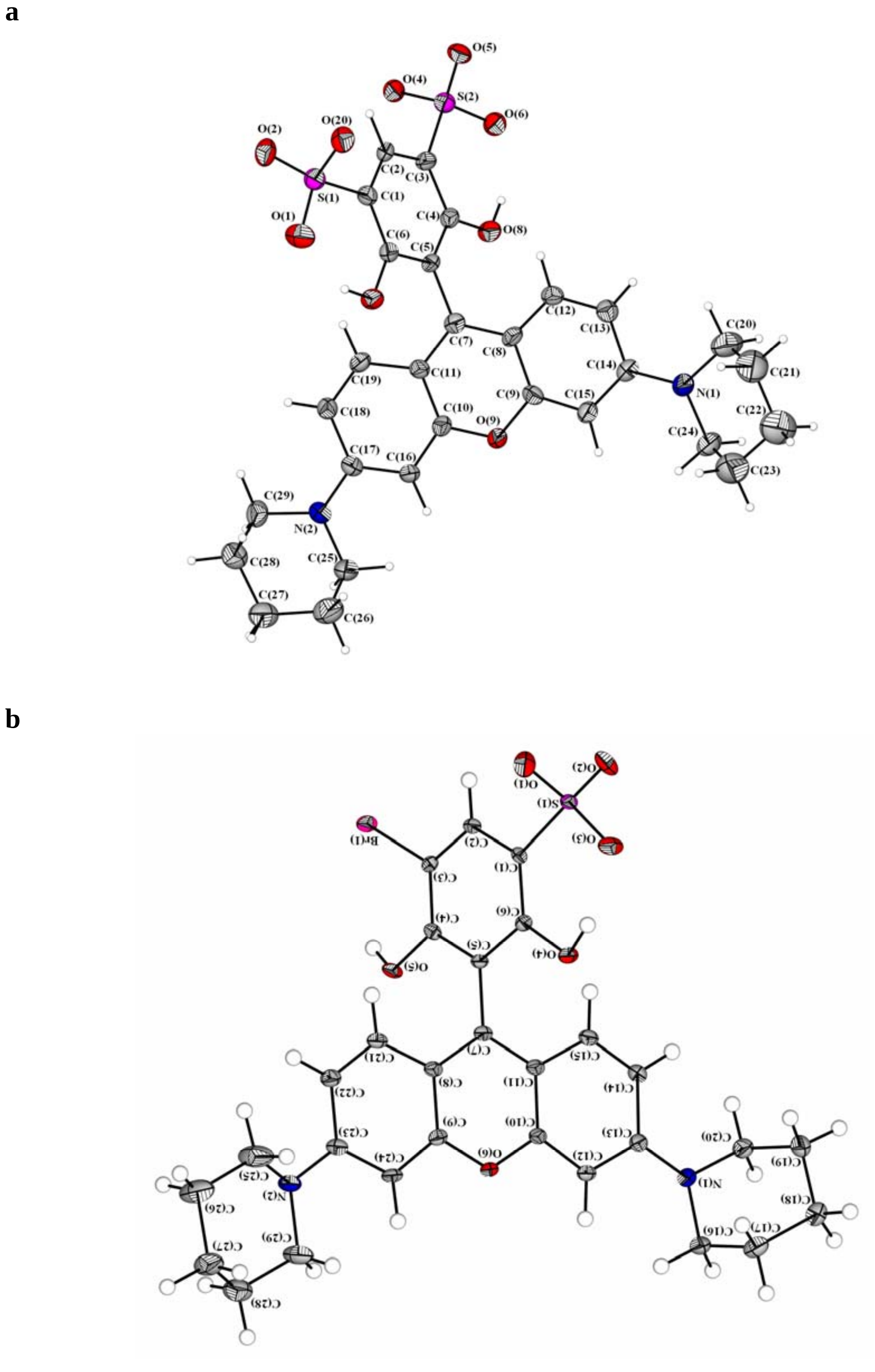
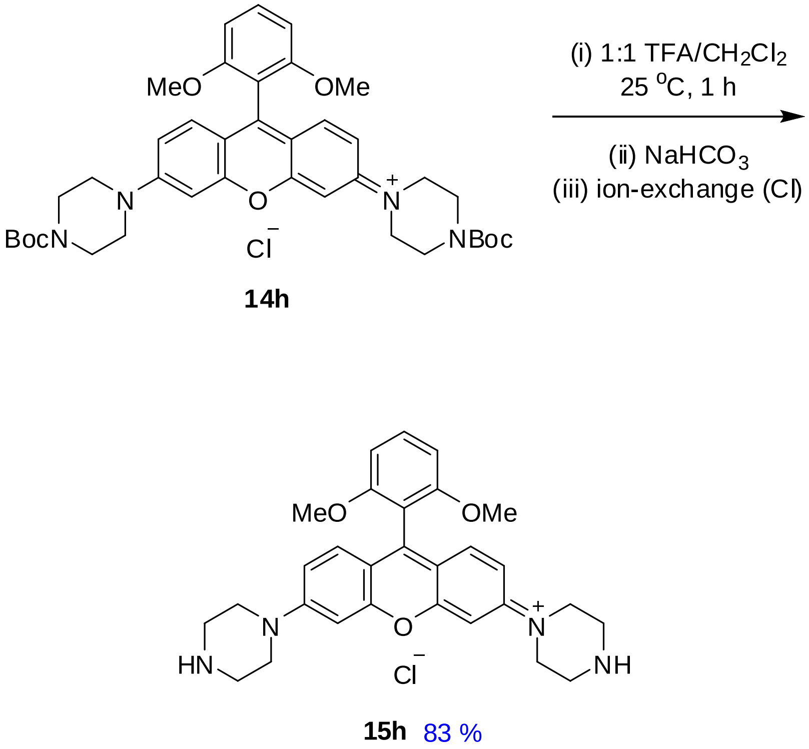

a
O(5)
O(4)
S(2)
O(6)
O(2)
O(20)
C(2)
C(3)
S(1)
C(1)
O(1)
C(4)
O(8)
C(6)
C(5)
C(12)
C(13)
C(20)
C(7)
C(8)
C(14)
C(21)
C(19)
C(11)
N(1)
C(9)
C(15)
C(22)
C(18)
C(10)
O(9)
C(24)
C(17)
C(16)
C(23)
C(29)
N(2)
C(28)
C(25)
C(27)
C(26)
b
O(1)
O(2)
S(1)
Br(1)
C(2)
O(3)
C(3)
C(1)
C(4)
C(6)
O(5)
C(5)
O(4)
C(21)
C(7)
C(15)
C(22)
C(8)
C(11)
C(14)
C(9)
C(10)
C(23)
C(13)
C(20)
C(25)
C(24)
O(6)
C(12)
C(19)
N(2)
N(1)
C(26)
C(18)
C(29)
C(16)
C(17)
C(27)
C(28)

Figure 4.4. X-ray structures for rosamines (**a**) **26** and (**b**) **27**.

Another modification to give a water-soluble probe is shown in reaction 4.4. Here the two *tert*-butyloxycarbonyl groups of rosamine **14h** were removed with acid to give, after neutralization and ion exchange, the bisamine **15h**. Compound **15h** dissolves easily in water; its spectroscopic properties in buffers were explored extensively (see below).


MeO
OMe
BocN
N
O
N
NBoc
Cl
14h
(i) 1:1 TFA/CH2Cl2
25 °C, 1 h
(ii) NaHCO3
(iii) ion-exchange (Cl)
MeO
OMe
HN
N
O
N
NH
Cl
15h 83 %

reaction 4.4

4.3 Spectroscopic properties of rosamine derivatives

Table 4.3. shows UV-Vis absorption wavelength maxima, molar extinction coefficients, fluorescence emission maxima, fluorescence peak broadness (measured by full width at half maximum height, fwhm), and quantum yields for the target compounds. These data were collected for all the compounds when dichloromethane was used as solvent. However, we observed that the *meso* aliphatic compounds **11a** and **11b** tend to be less stable in polar media.

The UV and fluorescence properties of all the rosamine derivatives are quite similar. They all absorb in the range of 555-570 nm except for the *meso*-thiophene derivative **11g** which gives red-shifted absorption and fluorescence maxima (see next paragraph for detailed discussion), indicating that the *meso*-aryl groups are not significantly conjugated to the xanthene core, presumably because steric factors involving these substituents make it energetically unfavorable for the molecules to become planar. The fluorescence maxima fall in the range of 580 ± 10 nm. The emission peak broadness values in this series range from 31 – 44 nm (CH_2Cl_2); these values are comparable with a typical rhodamine, rhodamine B (36 nm in CH_2Cl_2).

Several trends emerge from the spectroscopic data recorded for the compounds dissolved in dichloromethane. Relative to the other compounds, the *meso*-methyl derivative **11a** has slightly blue-shifted absorption and fluorescence maxima, but the extinction coefficient and the quantum yield are comparable with the rosamines having aromatic *meso*-substituents. Curiously, the benzyl compound **11b** is spectroscopically more similar to the *meso*-phenyl compound **11c** than to the methyl derivative **11a**. Compound **11g** is most red-shifted in terms of the wavelengths for absorption maxima and the fluorescence emission maxima. However, the quantum yield for this compound was about a third of that observed for other members in this series; this observation might be attributed to partial fluorescence quenching via electron transfer from the *S*-containing heterocycle.[181] Quantum yields for the Boc-piperazine derivatives were also low relative to other members of the series. The very best quantum yield observed was for the pyrrolidine derivative **13i**.

Differences between the piperidine derivatives **11** and the pyrrolidine derivative **13i** are accentuated in a more polar medium (ethanol, second section of Table 4.3.). For example, the fluorescence quantum yield of piperidine derivatives **11i** varies from 0.89 in CH_2Cl_2 to 0.28 in EtOH, while pyrrolidine derivative **13i** has very high quantum yields in both solvents (1.00 in CH_2Cl_2 and 0.95 in EtOH). To the best of our knowledge, this work is the first to uncover a relationship between the ring size of the peripheral amines in *rosamines* and their fluorescent quantum yields, but others have commented on similar effects in the context of 4-amino-1,8-napthalimide[182] and 4-aminophthalimide derivatives.[183] The origin of these effects is thought to be that smaller-ring cyclic amine substituents are less inclined to invert at nitrogen atom in the excited state, so this pathway for non-radiative loss is disfavored and the overall quantum yields are higher.[184] More polar solvents favor intramolecular charge transfer from cyclic amines where the *N*-atom more closely approximates a planar arrangement as a result of facile interconversion between ring forms. The interconversion is more rapid for piperidine derivatives than the pyrrolidine-based compounds. Overall, the quantum yield data for the compounds in ethanol are, in all the cases measured, less than that in dichloromethane. This is a somewhat unexpected finding since other rosamines are known to be very highly fluorescent even in alcohols; for instance, tetramethyl rosamine, has a quantum yield of 0.84 in MeOH.[101,172]

Quantum yields of 4,4-difluoro-4-bora-3a,4a-diaza-*s*-indacene (BODIPY) dyes with *meso*-aryl substituents are thought to be increased if the aryl group has an *ortho*-substituent, or if rotation of the aryl is otherwise restricted.[98] Similarly, fluorescein and derivatives with 2-substituted *meso*-aryl groups are far more fluorescent than the corresponding *meso*-phenyl compound.[185,186] However, it is surprising that in the series outlined in Table 4.3., unlike the analogous BODIPY dyes, rosamines with *ortho*-substituted *meso*-aryl groups (**11i**: 2-Me, **11f**: 2-OMe, **11h**: 2,6-OMe) have similar quantum yields to the phenyl derivative (**11c**). The rotation of *meso*-substituents did not reduce the rosamine quantum yields.

Table 4.3. Spectral properties of rosamines

	$\lambda_{max\ abs}$ (nm)	ε_{max}	λ_{fluor} (nm)	fwhm (nm)[a]	Φ_f[b]
			CH_2Cl_2		
11a	555	101900	570	35	0.91 ± 0.05
11b	566	95900	582	34	0.87 ± 0.01
11c	565	106000	585	37	0.89 ± 0.01
11d	570	115400	590	38	0.85 ± 0.01
11e	562	116700	582	35	0.91 ± 0.02
11f	566	118000	585	37	0.92 ± 0.03
11g	582	107600	603	41	0.33 ± 0.01[c]
11h	565	87300	586	36	0.86 ± 0.04
11i	565	123600	582	34	0.89 ± 0.01
12i	554	106900	575	37	0.80 ± 0.01
13i	558	127400	573	31	1.00 ± 0.09
22i	559	111000	579	36	0.59 ± 0.01
14i	555	120800	577	40	0.47 ± 0.01
14d	561	77900	585	42	0.41 ± 0.01
14h	555	109300	580	40	0.49 ± 0.02
23	564	62800	586	44	0.007
			EtOH		
11c	564	86700	588	39	0.30 ± 0.01
11d	568	70700	592	40	0.28 ± 0.01
11e	561	100400	582	37	0.31 ± 0.01
11f	564	115600	586	37	0.36 ± 0.01
11g	581	56500	606	43	0.10 ± 0.01[c]
11i	564	119500	584	36	0.28 ± 0.01
12i	556	98600	582	38	0.16 ± 0.01

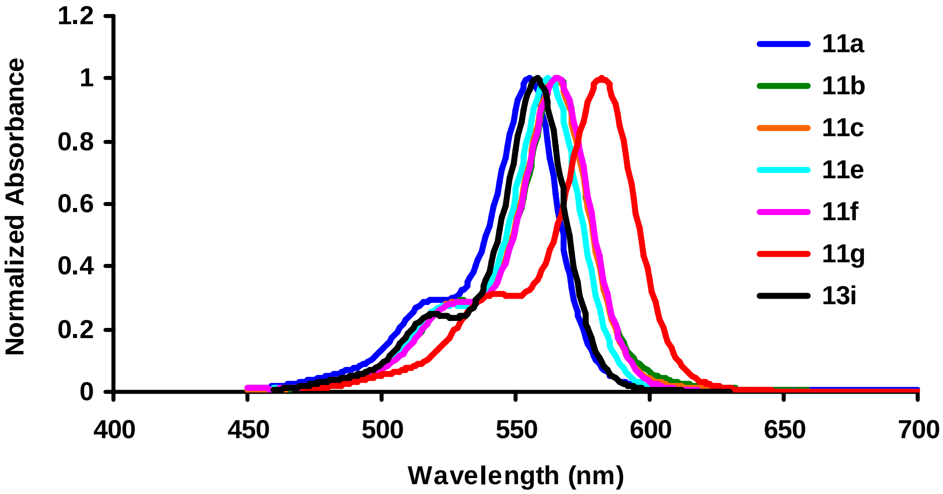
Table 4.3. Continued

	$\lambda_{\text{max abs}}$ (nm)	ε_{max}	λ_{fluor} (nm)	fwhm (nm)[a]	Φ_f[b]
			EtOH		
13i	558	120500	576	33	0.95 ± 0.01
22i	560	97000	582	38	0.14 ± 0.01
14i	556	108000	580	39	0.19 ± 0.01
24	557	71000	580	37	0.26 ± 0.01
25	557	84800	580	38	0.49 ± 0.02
15h	561	81000	582	45	0.008
26	562	82400	589	37	0.28 ± 0.01
27	553	117200	592	38	0.02
			0.1 M Phosphate buffer, pH = 7.4		
15h	542	71500	568	41	0.14 ± 0.02
25	567	76500	594	42	0.13 ± 0.01
26	564	73600	602	44	0.009

[a] Full width at half-maximum height: a measure of the sharpness of the fluorescence peaks; [b] Average of three measurements and Rhodamine 101 was used as a standard (Φ = 1.0 in EtOH)[187]; [c] Cresyl Violet (Φ = 0.54 in MeOH)[188] was used as a standard.

The only non-symmetrical compound (with respect to a plane through the *meso*-substituent) is the piperidine/morpholine derivative **22i**. This has a lower quantum yield than similar symmetrical compounds (*ie* **11i** and **12i**). The water-soluble derivatives **25** and **15h** had quantum yields of approximately 0.13 in 0.1 M phosphate buffer (pH 7.4). It is somewhat unexpected that bis-sulfonated rosamine derivative **26** is almost non-fluorescent in aqueous buffer, while a similar sulforhodamine (Texas Red) is highly fluorescent in water. It is also surprising that sulfonated/brominated derivative **27** lost its fluorescence even in EtOH. All these data suggested that the cyclic amine substituents except for pyrrolidine as described in this work somehow affected the fluorescence properties.

a



b

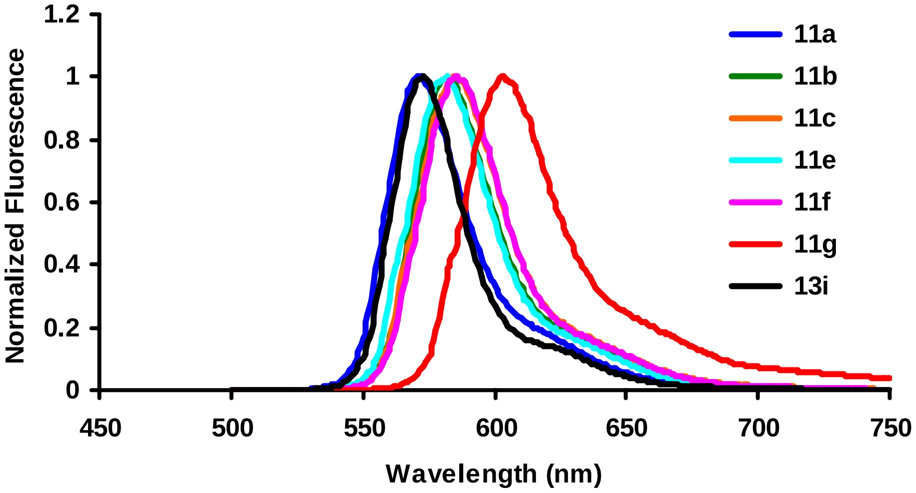


Figure 4.5. Normalized (**a**) UV-Vis absorption (10^{-6} M); and, (**b**) fluorescence spectra (10^{-7} M, excited at $\lambda_{\text{max abs}}$), for representative rosamines in CH_2Cl_2.

Some illustrative UV-Vis absorption and fluorescence emission spectra for molecules of Table 4.2. are shown in Figure 4.5. These data show that the *meso*-thiophene derivative **11g** is red-shifted, both in terms of absorption and fluorescence, relative to the other compounds. Absorption and emissions from all the other probes are clustered in a relatively narrow wavelength range, with the *meso*-methyl derivative **11a** at the blue end of the scale.

The piperidine derivative **11i** was chosen as a typical dye to study solvent-effects on absorption maxima, fluorescence emission maxima, and quantum yields (Table 4.4.). It emerged that the fluorescence efficiency of the rosamines is quite dependent on the solvent used. Two of those with small dipole moments (CH_2Cl_2 and $CHCl_3$, the compounds are not soluble in aliphatic hydrocarbons) gave optimal fluorescence quantum yields. No significant differences were observed in the UV-Vis absorption maxima for these compounds. Similarly, the Stokes shifts observed are all within the range 14 – 24 nm.

Table 4.4. Spectral properties of rosamine **11i** in different solvents

Solvent			Spectral Properties		
	Dielectric Constant	Dipole Moment	$\lambda_{\text{max abs}}$ (nm)	λ_{fluor} (nm)	$\Phi_f{}^a$
$CHCl_3$	4.8	1.04	566	580	0.95 ± 0.02
EtOAc	6.1	1.78	567	586	0.19 ± 0.01
THF	7.5	1.75	567	588	0.32 ± 0.01
CH_2Cl_2	8.9	1.60	565	582	0.89 ± 0.01
Acetone	21.0	2.88	564	587	0.15 ± 0.01
EtOH	24.3	1.69	564	584	0.28 ± 0.01
MeOH	33.0	1.70	562	582	0.19 ± 0.01
CH_3CN	36.6	3.92	562	586	0.14 ± 0.01
DMF	38.3	3.82	570	594	0.20 ± 0.01

[a] Rhodamine 101 ($\Phi = 1.0$ in EtOH) was used as a standard.

Data presented in Table 4.4. indicate the intensity of fluorescence emission from **11i** can be used as a probe for solvent polarity. This assertion was confirmed via series of measurements in a dichloromethane/methanol mixed solvent system (Figure 4.6.). The intensity of the fluorescence of this compound (at 1.4 x 10^{-7} M) increased nearly linearly with the proportion of dichloromethane (see Appendix Information).

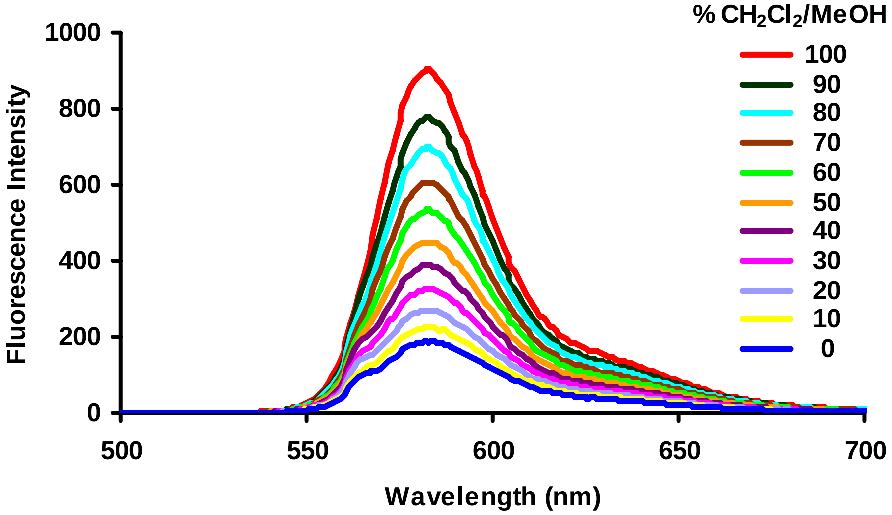
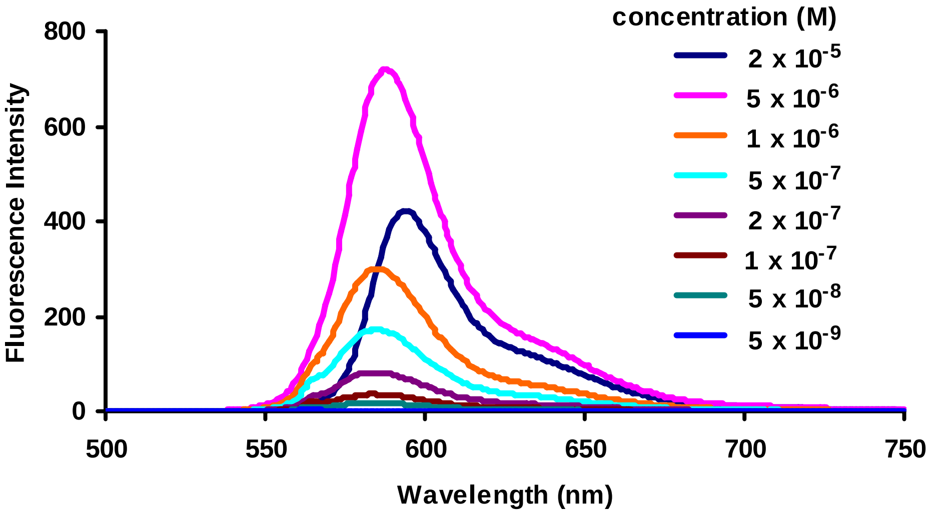


Figure 4.6. Fluorescence intensity (excited at 562 nm) of **11i** at 1.4 x 10^{-7} M in a CH_2Cl_2/MeOH mixed solvent system. The shoulder observed at around 560 nm is an artifact caused by the excitation source which was necessarily close in wavelength.

The fact that the molecules were most fluorescent in two less polar solvents could be due to aggregation in polar media. This possibility was tested by measuring fluorescence wavelength maxima for **11i** as a function of concentration in ethanol (a representative polar solvent; Figure 4.7.a). In fact, the wavelength for emission maxima did not change significantly over a concentration range of 5 x 10^{-9} to 5 x 10^{-6} M. This implies aggregation is not significant at those concentrations. Experiments with the non-ionic detergent Triton-X 100 gave no appreciable change in the fluorescence spectra of the dye in this concentration range (Figure 4.7.b), further supporting the assertion that **11i** does not aggregate. At a relatively high concentration of the dye, 2 x 10^{-5} M, there was a 7 nm red shift in the emission maximum for the fluorescence spectrum, indicating aggregation may only be a problem under such conditions.

a



b

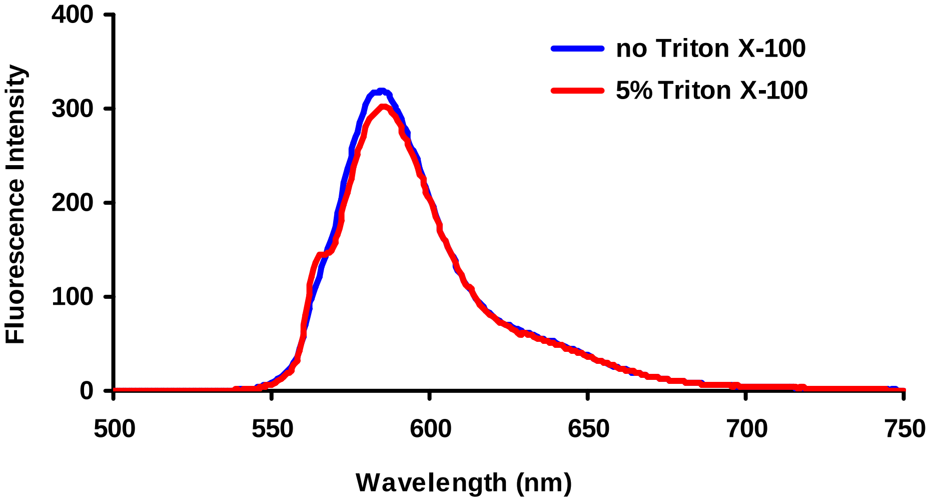


Figure 4.7. (**a**) Fluorescence spectra of **11i** in EtOH at different concentrations (excited at 564 nm); (**b**) effect of triton X-100 on the fluorescence of **11i** in EtOH (10^{-6} M, ex. 564nm).

Predictably, the quantum yields of the piperazine derivatives **14**, **15h**, and **23** are somewhat dependent on the nitrogen atoms at the periphery of the molecule. Derivatives for which this nitrogen is protected with a *tert*-butyloxycarbonyl group have moderate quantum yields. However, the free secondary amine of compound **15h**, for example, almost completely quenches the fluorescence unless it is protonated; this is because photoinduced electron transfer (PET) mechanisms may be possible for the amine.[181] Similarly, compound **23** in the free base form is almost completely non-fluorescent in dichloromethane, but solutions of this dye become strongly fluorescent upon protonation. The photophysical properties of rosamine **15h** are pH dependent. The absorption and fluorescence maxima of dye **15h** slightly shifted to the red with increasing pH from 2 to 12 (Table 4.5.). Figure 4.8. shows how the fluorescence of **15h** drops dramatically when the pH is raised above *ca* 6 and these fluorescence changes are reversible.

Table 4.5. UV absorption and fluorescence maxima of **15h** in phosphate solutions

pH	$\lambda_{max\ abs}$ **(nm)**	$\lambda_{max\ emiss}$ **(nm)**	**pH**	$\lambda_{max\ abs}$ **(nm)**	$\lambda_{max\ emiss}$ **(nm)**
2.00	538	565	8.12	551	571
3.00	538	565	8.43	555	574
4.08	538	565	9.09	559	582
5.18	539	565	10.22	560	584
6.04	539	565	10.67	560	584
7.08	541	567	11.13	560	584
7.55	544	568	11.98	560	584

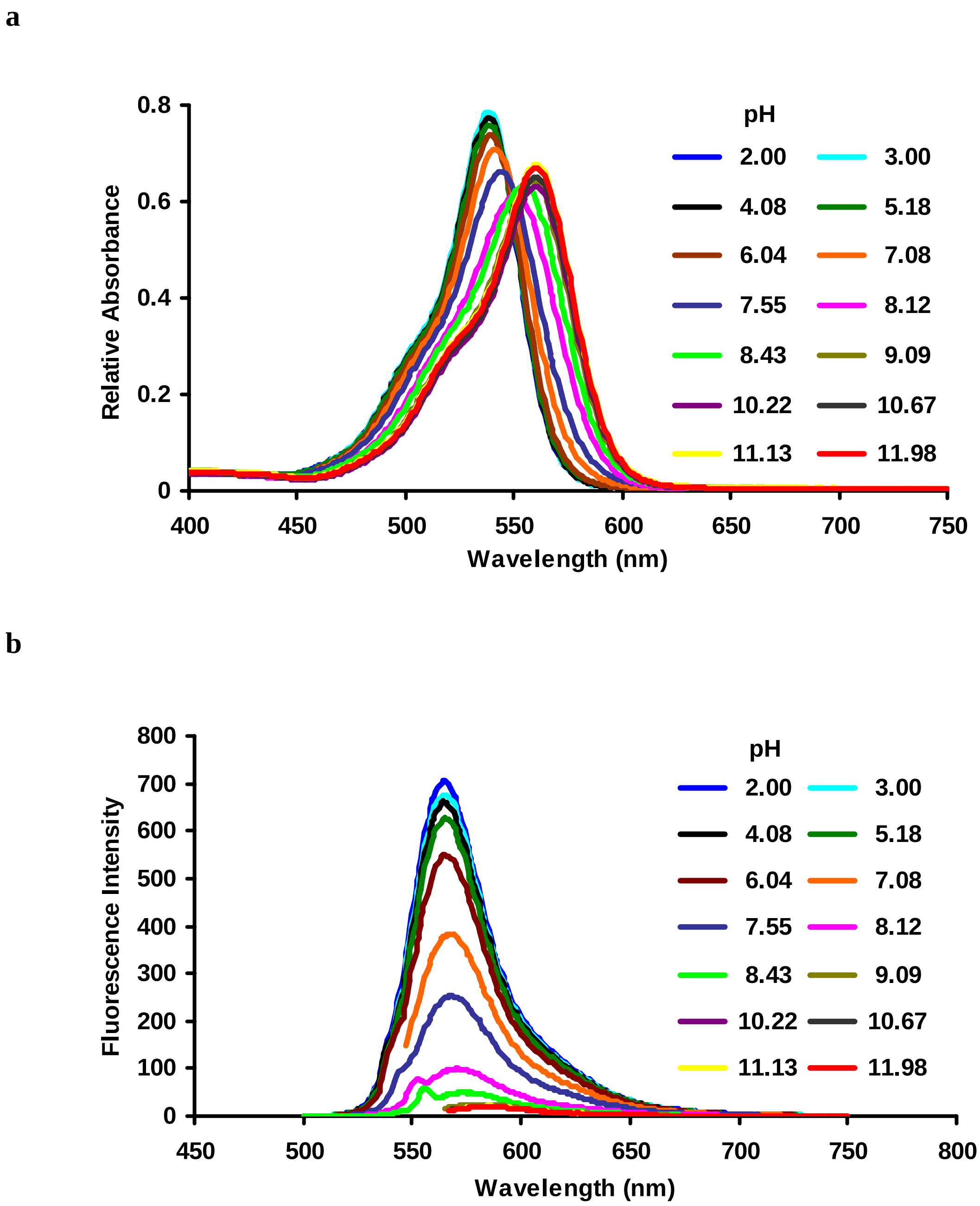
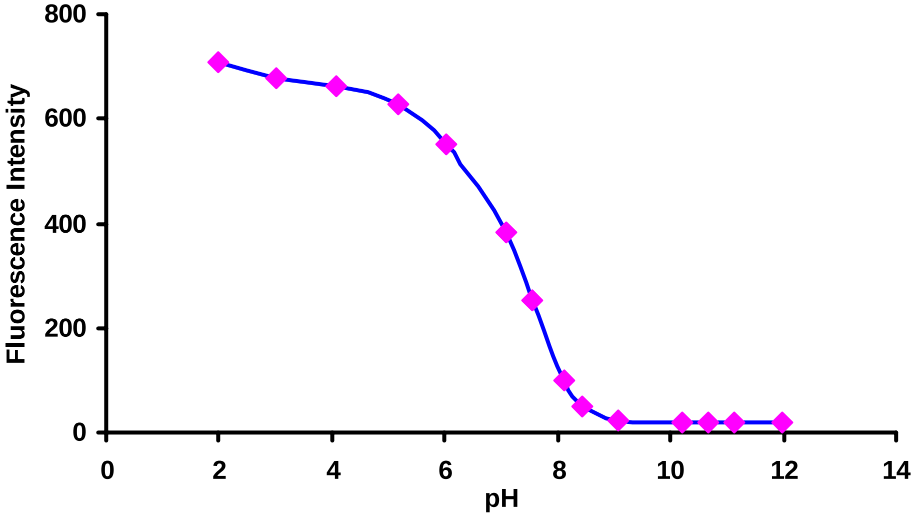


Figure 4.8. Dependence of electronic spectra of **15h** on pH. **a** Absorption (9.8 x 10^{-6} M); **b** fluorescence (5.7 x 10^{-7} M; excitated at $\lambda_{\text{max abs}}$); **c** Fluorescence intensities ($\lambda_{\text{max emiss}}$, excited at $\lambda_{\text{max abs}}$, 5.7 x 10^{-7} M) vs pH in phosphate solutions (50 mM, H_3PO_4+NaOH); **d** reversibility (excitation at $\lambda_{\text{max abs}}$ 538 nm and detection at $\lambda_{\text{max emiss}}$ 565 nm for pH 2.9; excitation at $\lambda_{\text{max abs}}$ 560nm and detection at $\lambda_{\text{max emiss}}$ 584 nm for pH 11.3: throughout the dye was used at 5.7 x 10^{-7} M and the pH was adjusted with 2M HCl or NaOH).

c



d

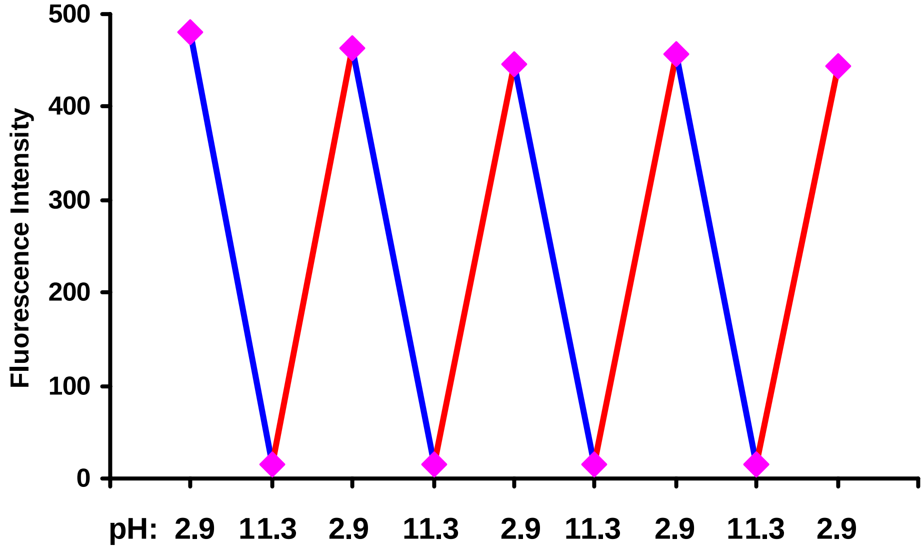


Figure 4.8. Continued.

Finally, the water soluble probe **25** was conjugated with avidin via activation of the dye-carboxylic acids (*N*-hydroxysuccinimide and *N,N'*-diisopropylcarbodiimide in DMF) then addition of the activated probe to the protein in 0.1 M $NaHCO_3$ at pH 8.3. The dye:protein ratio was calculated to be 0.8 by UV-Vis when three equivalents of the dye was used; this corresponds to 27 % labeling efficiency. When 5 eq. of dye was used then a dye protein ratio of 1.4 (28 % labeling efficiency) was observed. The latter sample was used to obtain the spectral data shown in Figure 4.9. The UV-Vis absorption and fluorescence emission maxima of **25** on avidin were observed to be within a few nm of the dye in phosphate buffer. The quantum yield of the protein conjugate was 0.06 in phosphate buffer compared with 0.13 for the free dye **25**.

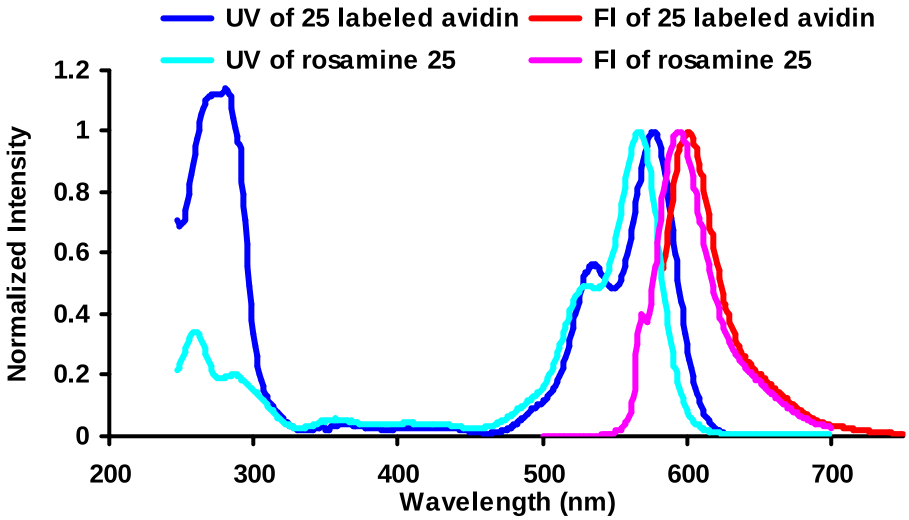
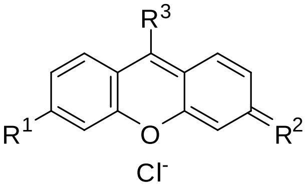


Figure 4.9. UV-Vis absorption and fluorescence spectra of rosamine **25** and **25** labeled avidin in 0.1 M phosphate buffer, pH 7.4.

4.4 Biological activity test

The *in vitro* antiproliferative activity of rosamine derivatives (Table 4.6.) against a promyelocytic leukemia cell-line, HL60 was determined using a 24 h endpoint MTT assay.[189] Results were expressed as IC_{50}: the concentration of compound (in μM) that inhibits proliferation rate by 50 % as compared to control untreated cells. From the assay, compounds **11**, **13i**, **14i**, **22i** demonstrated their anti-tumour activity with IC_{50} values in the sub-micromolar range. Compound **11g** which had a thiofuran and the *para*-iodo aryl substituted **11d** showed the highest activity among the analogues (IC_{50} of 0.10 and 0.09 μM respectively). In contrast, compounds **11a**, **12i**, **24**, **25**, and **28** displayed moderate to poor activity from single-digit micro-molar IC_{50} values to undeterminable IC_{50} up to 100 μM.

Table 4.6. The structure-activity relationship (SAR) and *in vitro* cytotoxicity of rosamine analogues in HL60 cells



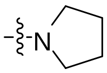
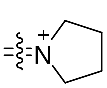
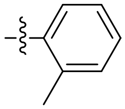
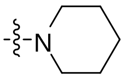
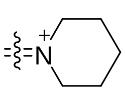
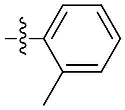
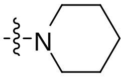
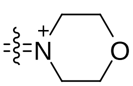
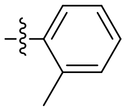
Compd	R_1	R_2	R_3	IC_{50} (μM) ± S.D.^{a} in HL60
13i				0.72 ± 0.09
11i				0.76 ± 0.03
22i				0.35 ± 0.01

Table 4.6. Continued

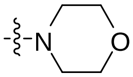
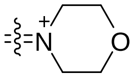
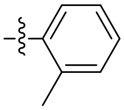
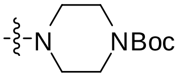
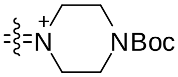
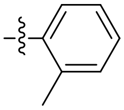
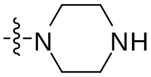
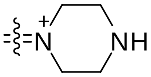
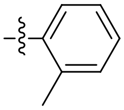
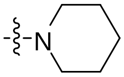
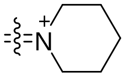
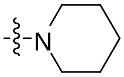
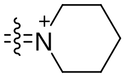
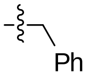
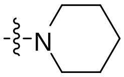
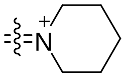
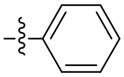
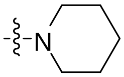
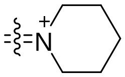
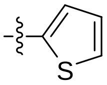
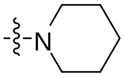
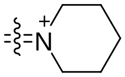
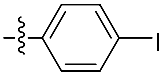
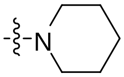
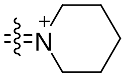
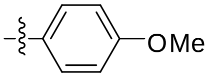
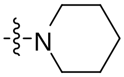
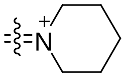
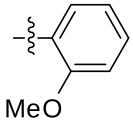
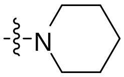
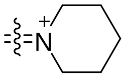
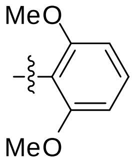
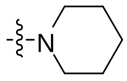
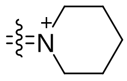
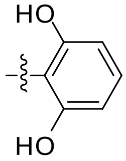
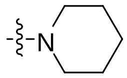
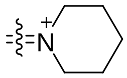
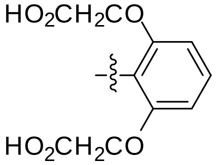
Compd	R_1	R_2	R_3	IC_{50} (μM) ± S.D.[a] in HL60
12i				8.27 ± 2.16
14i				0.62 ± 0.07
28				53.3 ± 0.30
11a			CH_3	3.86 ± 1.46
11b				0.82 ± 0.04
11c				0.47 ± 0.13
11g				0.10 ± 0.04
11d				0.09 ± 0.01
11e				0.66 ± 0.17
11f				0.95 ± 0.01
11h				0.25 ± 0.22

Table 4.6. Continued

Compd	R_1	R_2	R_3	IC_{50} (μM) ± S.D.^{a} in HL60
24				2.91 ± 1.86
25				>100

^{a}Mean IC_{50} and standard deviation of triplicate determination assessed *in vitro* at 24 h post-treatment using MTT assay.

The influence of the cyclic amine substituents on the anti-proliferative activity of the compounds was evident from studying compounds with 2-methyl phenyl at the *meso* position (Table 4.6. and Figure 4.10.). Regardless of the size of the ring, the derivatives containing hydrophobic cyclic amines from pyrrolidine (**13i**), piperidine (**11i**) to Boc-piperazine (**14i**) exhibited moderate anti-proliferative activity with IC_{50} values of 0.62-0.76 μM. On the other hand, the derivatives with cyclic amines that contain exposed oxygen or NH isosteres as in the case of compounds **12i** and **28** had 10- to 50-fold higher IC_{50} of 8.27 or 53.3 μM respectively. The unsymmetrical rosamine **22i**, which had a combination of piperidine and morpholine substituents interestingly had the lowest IC_{50} value among compounds with 2-methyl phenyl at the *meso* position, alluding to the possible importance of an amphilic structure with contrasting hydrophobic and hydrophilic halves.

For the effect of *meso*-substitution on anti-proliferative activity of rosamines, compounds that have symmetrical piperidine-substituted dye cores with a variety of *meso*-substituents were studied (Table 4.6. and Figure 4.10.). Similar to compounds **12i** and **28**, the derivatives with hydrophilic substituents such as the phenolic **24** and the carboxylic **25** had higher IC_{50} values than the unsubstituted *meso*-aryl **11c**. Having an aryl substituent at the meso position, whether directly (**11c-h**) or through an alkyl spacer (**11b**), was important for anti-proliferative activity and was convincingly demonstrated in the lower activity observed in compound **11a** which had only a simple methyl substituent at the *meso* position. Among the aryl substituted compounds, the thiofuran (**11g**) and the *para*-iodo aryl (**11d**) structures had the lowest IC_{50} values compared to a simple phenyl-substituted compound **11c**, while 4-methoxy aryl (**11e**), mono-2-methoxy (**11f**) and di-2-methoxy (**11h**) aryl substitutions did not confer additional activity.

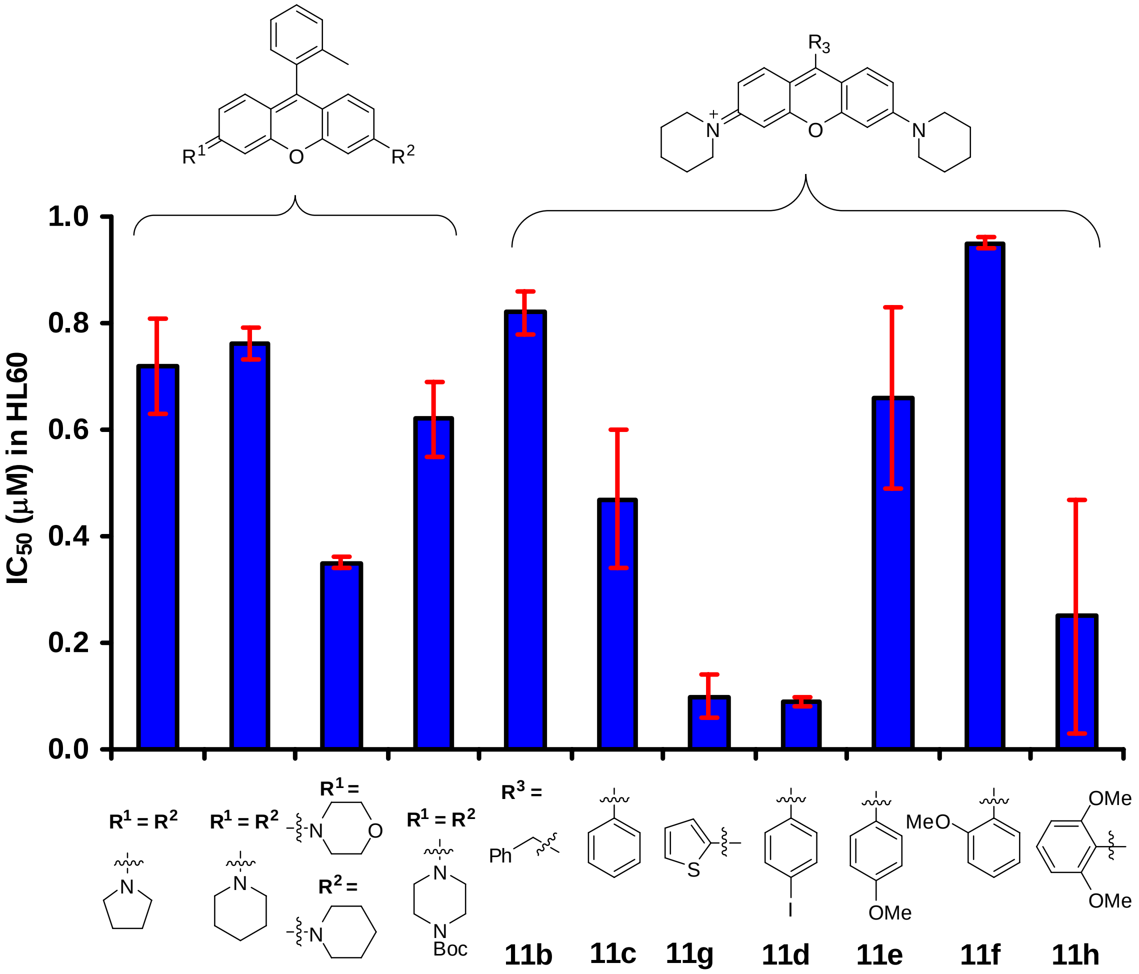


Figure 4.10. Structure-activity relationship of rosamine derivatives in HL60 cells.

The *in vitro* antiproliferative activity of the most active compounds **11d** and **11g** were assessed against a panel of cell-lines derived from human solid tumors including colon cancer, breast cancer, oral squamous cell carcinoma and nasopharyngeal carcinoma (Table 4.7.). A 48 h assay endpoint which is more typical of cytotoxicity studies was used. The anti-proliferative activity of rhodamine 123 (Rh123) was also simultaneously determined for comparison. Both **11d** and **11g** exhibited at least 10-fold lower IC_{50} values compared to Rh123. In addition, the thiofuran-substituted compound (**11g**) consistently showed between 1.5-fold to 4-fold lower IC_{50} values across all four types of solid tumors compared to compound **11d**.

Table 4.7. Cytotoxic effects of rhodamine analogues on carcinoma and immortalized normal human epithelial cell types

Cell line	Tissue Origin	IC_{50} (μM) ± S.D.[a]		
		11d	**11g**	**Rh123**
HCT116	Colon	0.39 ± 0.11	0.15 ± 0.06	7.92 ± 0.95
MCF-7	Breast	0.39 ± 0.22	0.27 ± 0.16	5.61 ± 0.61
HSC2	Oral	0.25 ± 0.12	0.12 ± 0.09	4.48 ± 2.23
OKF6[b]	Oral	0.41 ± 0.07	0.25 ± 0.10	9.84 ± 3.46
HK1	Nasopharyngeal	0.42 ± 0.06	0.09 ± 0.01	5.86 ± 0.15
NP69[b]	Nasopharyngeal	0.51 ± 0.16	0.33 ± 0.20	6.28 ± 0.21

[a] Mean IC_{50} and standard deviation of triplicate determination assessed *in vitro* at 48 h post-treatment using MTT assay. [b] Immortalized normal human epithelial cells

To investigate whether rosamines **11d** and **11g** have greater anti-proliferative effects on cancer cells compared to normal cells, two immortalized epithelial cell-lines from oral (OKF6) and nasopharyngeal (NP69) origin were also included in the study. Gratifyingly, both **11d** and **11g** were more cytotoxic towards the cancer cell-lines than the immortalized normal cell-lines, as demonstrated in the 1.25-fold to 3-fold higher IC_{50} values in the normal compared to the cancer cell-lines.

Even though both compound **11d** and **11g** have equally potent anti-proliferative activity with similar IC_{50} values in the low sub-micromolar range, the fluorescence quantum yield of **11g** is low, at a value that is approximately 3-fold lower compared to **11d** (0.10 vs 0.28 in ethanol). Therefore, only the intracellular localization of compound **11d** in HSC2 cells was examined via confocal microscopy using dual staining techniques (Figure 4.11.).

Co-staining images and topographic profiles of cells containing compound **10d** and a mitochondria-specific dye Rh123 revealed an almost identical overlap, suggesting that compound **11d** localized particularly well in mitochondria (Figure 4.11.a and b). In comparison, compound **11d** displayed only partial co-localization with ER and lysosomes, according to the confocal images and topographic profiles of compound **11d** with ER-Tracker (Figure 4.11.c and d) and with LysoTracker (Figure 4.11.e and f) respectively. Staining of the cytoplasmic or nuclear membrane by compound **11d** was not detected, indicating that compound **11d** does not react non-specifically with biological membranes. Furthermore, the nucleus remained free of compound **11d** (dark nuclear area) indicating that this class of compounds would not be expected to directly damage DNA.

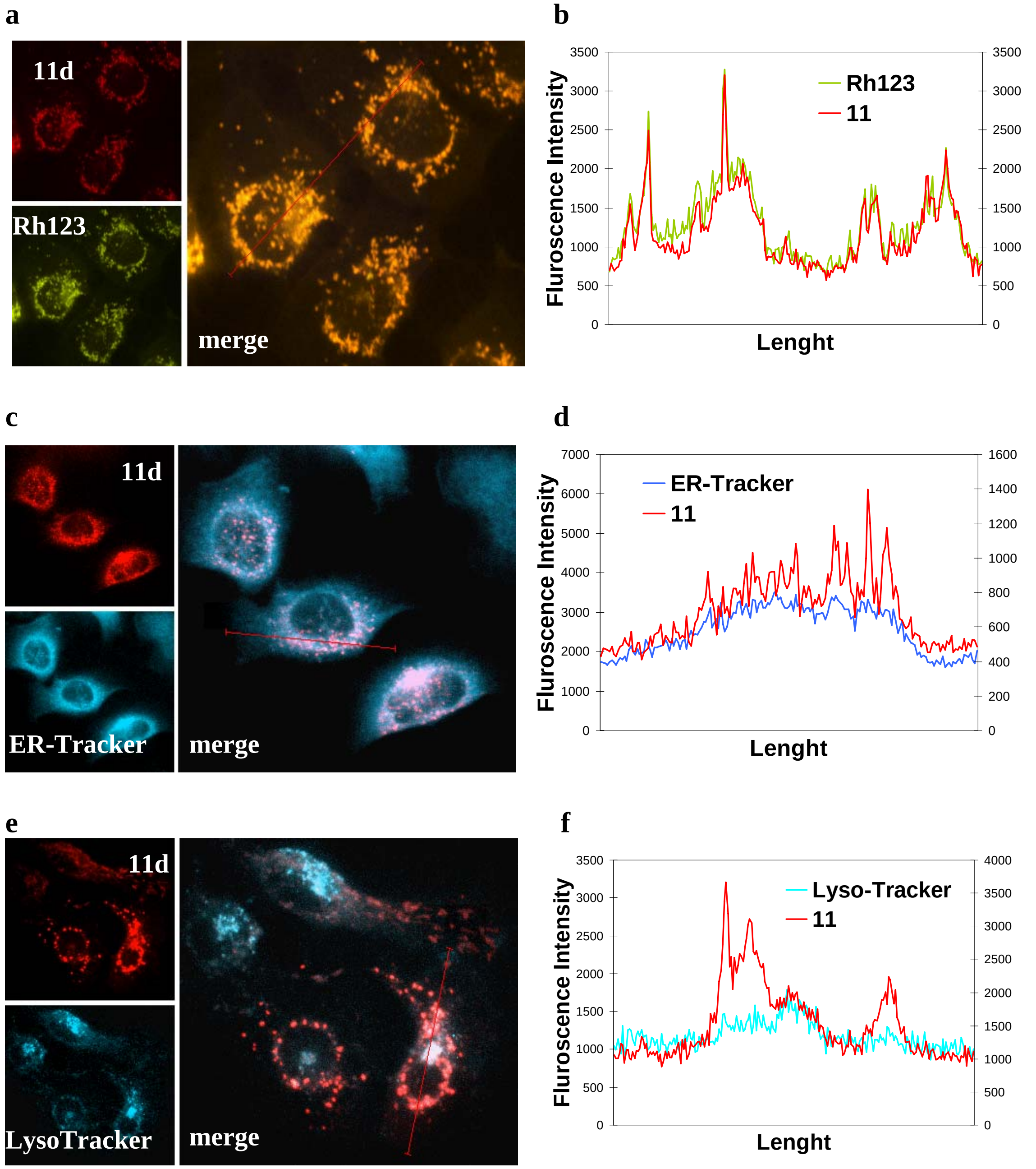

a
11d
Rh123
merge
b
Fluroscence Intensity
3500
3000
2500
2000
1500
1000
500
0
Rh123
11
Lenght
c
11d
ER-Tracker
merge
d
Fluroscence Intensity
7000
6000
5000
4000
3000
2000
1000
0
1600
1400
1200
1000
800
600
400
200
0
ER-Tracker
11
Lenght
e
11d
LysoTracker
merge
f
Fluroscence Intensity
3500
3000
2500
2000
1500
1000
500
0
4000
3500
3000
2500
2000
1500
1000
500
0
Lyso-Tracker
11
Lenght

Figure 4.11. Intracellular localization of compound **11d** in HSC2 cells. Confocal images (**a**,**c**,**e**) and fluorescence topographic profiles (**b**,**d**,**f**) of HSC2 cells double-stained with 100 nM of compound **11d** and respective organelle probes. (**a**,**b**) Mitochondria were labeled with 100 nM of Rh123 and excited at 494 nm. (**c**,**d**) Endoplasmic reticulum were labeled with 100 nM ER-Tracker and excited at 365 nm. (**e**,**f**) Lysosomes were labeled with 500 nM of LysoTracker and excited at 365 nm. Compound **11d** was excited at 575 nm. Line segment indicates the longitudinal transcellular axis analyzed to generate the topography fluorescence profiles. Objective magnification × 63.

The induction of apoptosis was quantified in flow cytometry experiments measuring the externalization of membrane phosphatidylserine through annexin V-FITC staining, which is an event considered characteristic of cells undergoing apoptosis (Figure 4.12.). Flow cytometric analysis of HSC2 cells treated with compound **11d** at IC_{50} value (0.25 μM) showed the onset of apoptosis at 8 h of compound incubation with 16% of the cells staining positive for annexin V compared to less than 10% at 0 h or 4 h time-points. The proportion of cells undergoing apoptosis continued to increase rapidly to 58 % within 24 h.

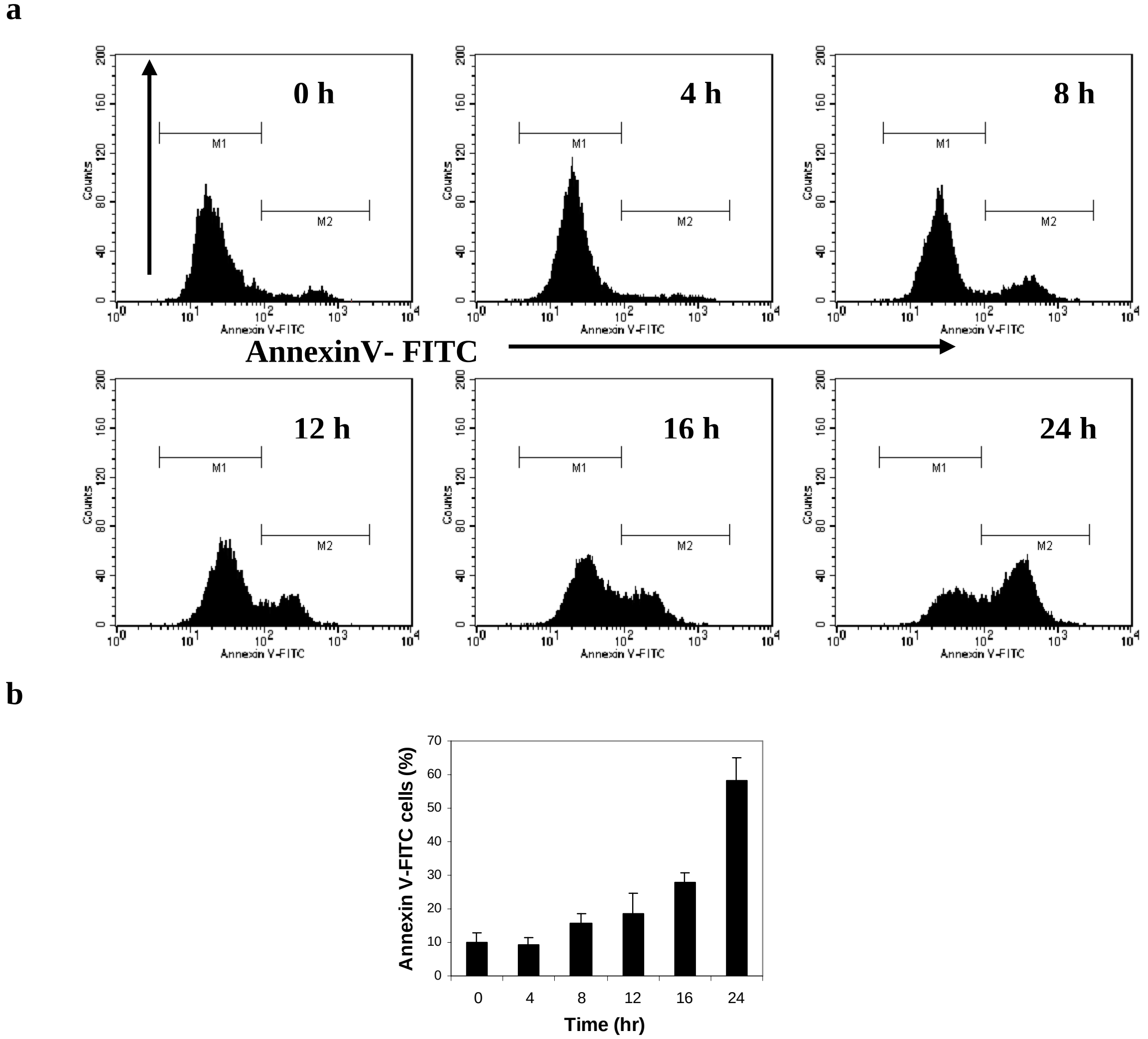


Figure 4.12. Compound **11d** induced apoptosis. (**a**) A representative histograms of event and (**b**) mean percentage of annexin V-FITC binding to PS as an indicator of apoptosis in HSC-2 cells treated with 0.5 μM compound **11d** at different duration.

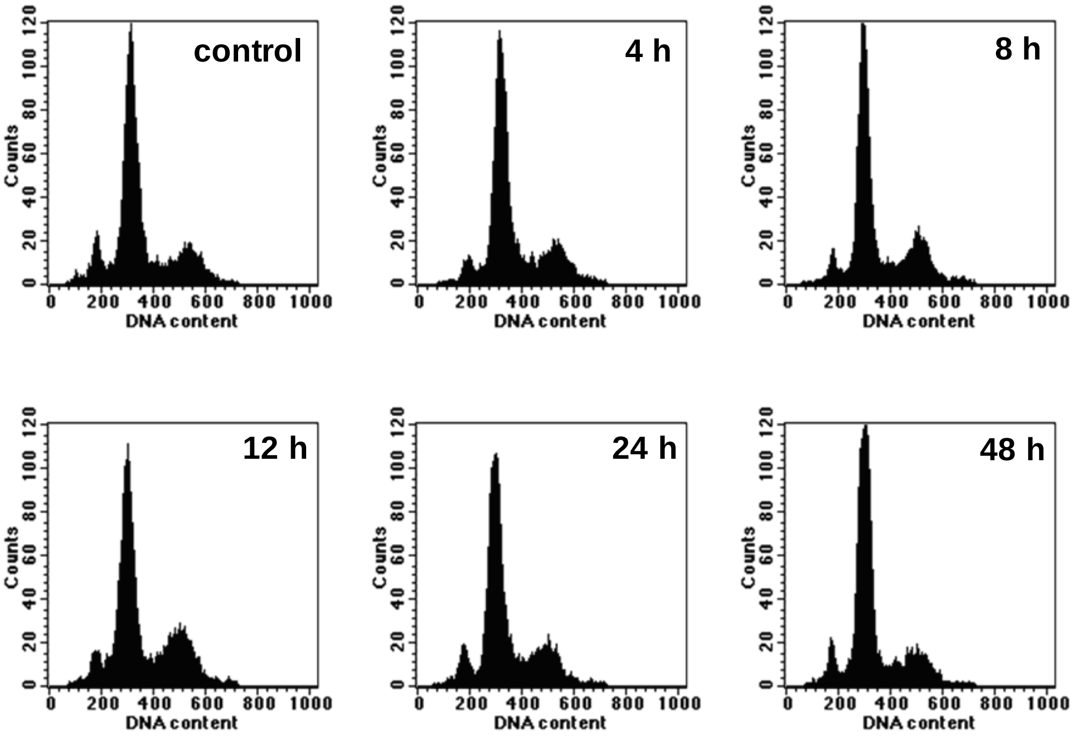


Figure 4.13. Effects of **11d** on cell cycle. HSC2 cells were treated with 0.25 μM of compound **11d** for various times and samples were harvested for flow cytometry analysis.

The cell cycle profile of HSC2 when treated with an IC$_{50}$ concentration (0.25 mM) of compound **11d** was examined in a time course experiment (Figure 4.13.). From 4 h to 48 h, the cell cycle profile remained unchanged, indicating cell death caused by compound **11d** did not occur as a result of cell cycle arrests.

4.5 Conclusions

This chapter describes efficient solution phase syntheses of rosamines with cyclic amine substituents. It is amenable to scale-up and purification of the products. Further, water-soluble derivatives can be prepared via modification of the initial products. The rosamine products are shown to be more fluorescent in dichloromethane and chloroform than in more polar solvents. However, this is not observed for some commercially available rosamines with acyclic amine substituents such as tetramethyl rosamine. This solvent dependence might be useful to probe the media polarity. Further, the data collected here illustrate that the ring size of the cyclic amine substituents has significant effects on the quantum yields; this trend has not been noted previously for rosamine dyes. Finally, the spectroscopic properties of the water soluble rosamine **25** conjugated to a protein was not significantly different to that of the free dye in aqueous media.

The anti-proliferative activity of the new rosamine derivatives against a panel of cell-lines from leukemia and solid tumors were studied. Structure activity relationship study indicated the importance of having hydrophobic substituents at the peripheral cyclic amines as well as at the *meso*-aryl groups. Structures with aryl substituents at the *meso* position, either directly attached or via a -CH_2- spacer conferred extra activity. The most active compounds **11d** and **11g** were at least 10-fold more potent than rhodamine123, a structurally similar compound whose anti-cancer properties have been extensively investigated. Furthermore, our study also showed that compounds **11d** and **11g** showed greater cytotoxicity towards oral and nasopharyngeal cancer cells compared to immortalized normal cells of the same organ type.

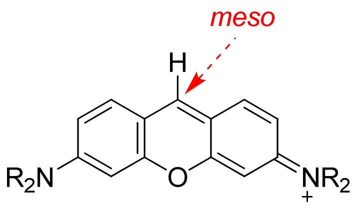
Fluorescence microscopy studies showed that compound **11d** localizes exclusively within the mitochondria. This, together with data from cell cycle analysis and onset of apoptosis studies, suggests that compound **11d** induced cell death through mitochondria-dependent apoptosis rather than through damage to nucleic materials. The intracellular localization data here also agrees with literature reports where higher mitochondrial transmembrane potential have been noted in cancer cells compared to normal epithelial cells, to result in accumulation of lipophilic cations, such as the rosamine derivatives studied here, in mitochondria. Overall, our results suggest that these compounds may offer a unique potential for the design of mitochondrial targeting agents that either directly kill or deliver cytotoxic drugs to selectively kill cancer cells.

CHAPTER V

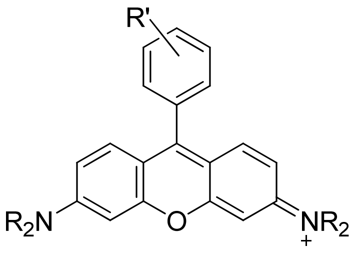
NEW PYRONIN DYES WITH MESO-HETEROATOM SUBSTITUENTS: SYNTHESIS AND SPECTROSCOPIC PROPERTIES*

5.1 Introduction

Pyronin dyes differ from rhodamines or rosamines in that the *meso*-substituent for pyronins is usually an *H*-atom instead of phenyl derivatives. Since the *meso*-substituents usually do not affect the optical properties, pyronin dyes have very similar fluorescence properties with the corresponding rhodamine derivatives.



pyronin

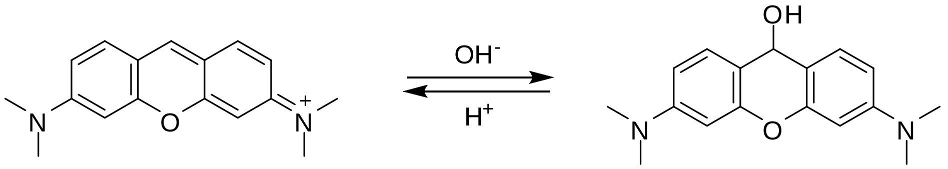


rhodamine/rosamine

A typical feature of the dye character of pyronins is their ability to be used as pH indicators since pyronin dyes hydrolyze in water especially under basic conditions which convert them into 9-ol-3,6-bisamino-9H-xanthene (xanthyldrols).[190] The color of their acidic and neutral solution is red (xanthylium 3,6-bis(dimethylamino) chloride PY: $\lambda_{max\ abs}$ = 547 nm; $\lambda_{max\ emiss}$ = 566 nm), while their alkaline solutions are colorless (PY: $\lambda_{max\ abs}$ = 390 nm; $\lambda_{max\ emiss}$ = 463 nm); a reversible change results from the variation in pH from basic to acidic. Extinction coefficients of the red solutions are much higher than these measured for the corresponding colorless solutions.

---

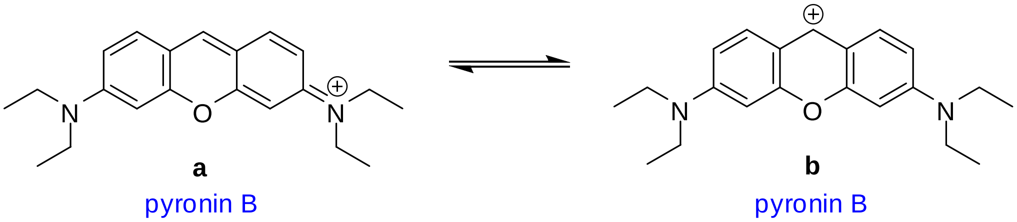
*Reprinted in part with permission from "Fluorescent Amino- and Thiopyronin Dyes", Liangxing Wu, Kevin Burgess, *Org. Lett.*, **2008**, *10*, 1779-1782. Copyright 2009 American Chemical Society.



pyronin Y (PY)
cationic: R^+
$\lambda_{max\ emiss}$: 566 nm

xanthydrols
neutral: ROH
$\lambda_{max\ emiss}$: 463 nm

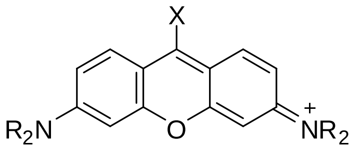
Photophysical properties of pyronin dyes are sensitive to the local environment and the photophysics of pyronins such as pyronin B (diethyl substituted derivative) has been well studied. The nonradiative $S_1 \rightarrow S_0$ internal conversion of pyronins has been studies in different solvent systems.[191,192] This process is associated with the rigidity of the xanthene-amine bond, solvent polarity and structures. Several models have been proposed to explain the effect of solvent and molecular structure on internal conversion in pyronin and rhodamine dyes: the intramolecular rotation model,[193] the twisted intramolecular charge-transfer (TICT) model[194-196] and the umbrella-like-motion (ULM) model.[197] In the intramolecular rotation model, internal conversion is mainly governed by rotation about the xanthene-amine bond. The rate of internal conversion should decrease as the solvent viscosity and size of the amino groups are increased. However, this situation has not been observed for the most rhodamine dyes as well as pyronins in solution.[191,192] In TICT model, internal conversion occurs via TICT state in dye compounds. TICT state is characterized by electron transfer from amino groups to the xanthene ring and by rotation about the xanthene-amine bond. In this model, internal conversion is controlled mainly by solvent polarity and the donor-acceptor properties of the amino group and xanthene ring.[198] In the TICT model, the energy of the TICT state mainly determines the nonradiative rate. Increasing the solvent polarity decreases the energy of the TICT state and increases the nonradiative rate. In addition, the structure of the dye molecule and solvent viscosity can also influence the formation of TICT state in this model because intramolecular rotation is involved. However, experimental studies do not confirm the TICT model for pyronin B in polar aprotic and protic solvent systems.[191,192] In the ULM model, a structural change in the amino groups from a planar structure (**a**) to a pyramidal structure (**b**) results in internal conversion in the excited state.[197,199,200] This umbrella-like motion changes the xanthene-amine bond from a double bond to a single bond and moves the positive charge from the amino group to the xanthene ring. Here, solvent polarity has to be taken into account for internal conversion process but the main factor is the specific solute-solvent interactions. This type of interactions results from hydrogen bonding formation between solute (dyes) and solvent molecules. The stability of the resonance structures of pyronins is directly related to the internal conversion because specific solute-solvent interactions are effective on either ground state or excited state structure. The ULM model is able to explain the behavior of pyronin B in different solvent systems.[191,192]



Pyronin dyes tend to form aggregates in aqueous solutions and in matrix systems as the dye concentration is increased. This behavior is extensively studied theoretically and experimentally in aqueous solutions and matrix systems such as cyclodextrans (CDs), cucurbiturils (CBs), zeolites and colloidal suspensions.[201-207]

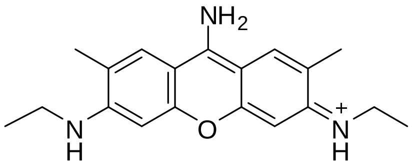
Besides the photophysics, the pyronin dyes have also been well known for its biological applications. Pyronin dyes such pyronin Y (PY) have been widely used to stain biomolecules especially RNA and DNA.[208] When used alone, pyronin Y stains both DNA and RNA and the staining capabilities of the nucleic acids vary depending upon the pyronin Y concentration at equilibrium.[209] Specificity of the staining reaction is assayed by testing sensitivity of the stainable material to RNase or DNase. The colored complexes detected by light absorption in fixed cells stained with pyronin Y are non-fluorescent and are most likely the products of condensation of single-stranded (ss) RNA or DNA by PY. The products of PY interaction with double-stranded (ds) nucleic acids are fluorescent and can be detected in cells by cytofluorometry.[210] The interactions of pyronin Y with nucleic acids are complex process.[211] PY binds by intercalation to ds nucleic acids. Its binding affinity to ds RNA is sevenfold higher than to ds DNA. Binding of PY to ss nucleic acids is initiated by PY intercalation between neighboring bases, neutralization of the polymer charge by the cationic dye and subsequent condensation and agglomeration (precipitation) of the product. In the final product, PY molecules are interspaced with bases, forming stacks of alternating dye-base composition. PY fluorescence is nearly totally quenched in these complexes. PY can denature the ds sections of nucleic acids, rendering them single-stranded and causing their condensation and precipitation. Therefore, at increasing PY concentration, the fluorescence of PY bound to ds nucleic acid is suppressed because of the progressive denaturation of the ds sections. The binding of PY to DNA can be suppressed by DNA-specific ligands such as Hoechst 33342 and methyl green.[208] In the presence of these dyes, PY can be used as a specific RNA fluorophore.[209] However, PY also stains mitochondrial membranes in live cells just like other cationic dyes, therefore, the RNA specificity obtained with PY staining needs to be further confirmed with alternative methods. Furthermore, the cytotoxicity of PY to certain cell-lines also limits its wide applications.[210]

Pyronins have also been used as laser dyes and to detect trace proteins and determine species such as As(III)[212], Cr(VI)[213] and tannins.[214] Despite all these activities, only a few pyronin derivatives are available and investigations on the structure-property relationships of pyronins are rare. There is almost no literature on any fluorescent dyes of structure **A** where X = NR_2 or SR. In fact, to the best of our knowledge, the only reported compound in this category is "Pyronin 12", *ie* structure **B**. Dyes that have been christened with a trivial name are often well known and frequently used. However, Pyronin 12 was mentioned peripherally in a paper on dimerization of dyes in aqueous solution,[215,216] but its synthesis and spectroscopic properties were not reported. Moreover, there do not appear to be any reports of others using this dye, or any like it, since it was first described. This chapter outlines syntheses and the spectroscopic properties of a range of dyes **A** where the X substituent is an amine or a thiol.



A

X = aryl or alkyl: well known
X = *NR_2 or SR: unexplored*



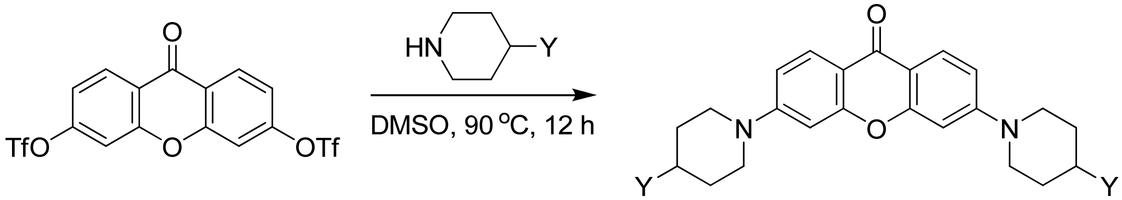
B

Pyronin 12

5.2 Syntheses of new pyronin derivatives

The approach used in this work was based on reactions of a ditriflyl xanthone **16** with piperidine derivatives as shown in Scheme 5.1. The starting material for this transformation is a known compound[179] that we prepared via a slight modification of the literature procedure (see chapter IV and Appendix information); this modification was scalable to at least 13 g since, unlike the original method, no chromatography was required. Amines related to **17** and **29** have been prepared via condensation procedures that involve several steps;[158,166,172,176-178] the reaction outlined below is more direct once the common ditriflyl xanthone starting material has been made. The product precipitates from the reaction mixture after addition of water, and is then recrystallized. There are, however, restrictions on the amines that can be used; less nucleophilic ones (*eg* $HNBu_2$, $HNEt_2$) tended to give mostly monosubstitution under the conditions studied (DMSO, 90 °C).

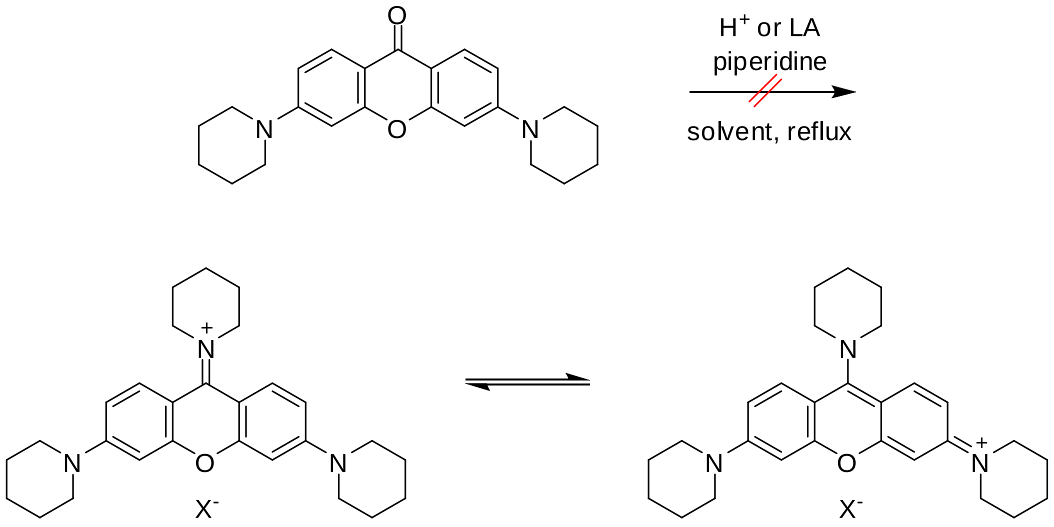
Scheme 5.1. Synthesis of 3,6 diamino-xanthones **17** and **29**.



16

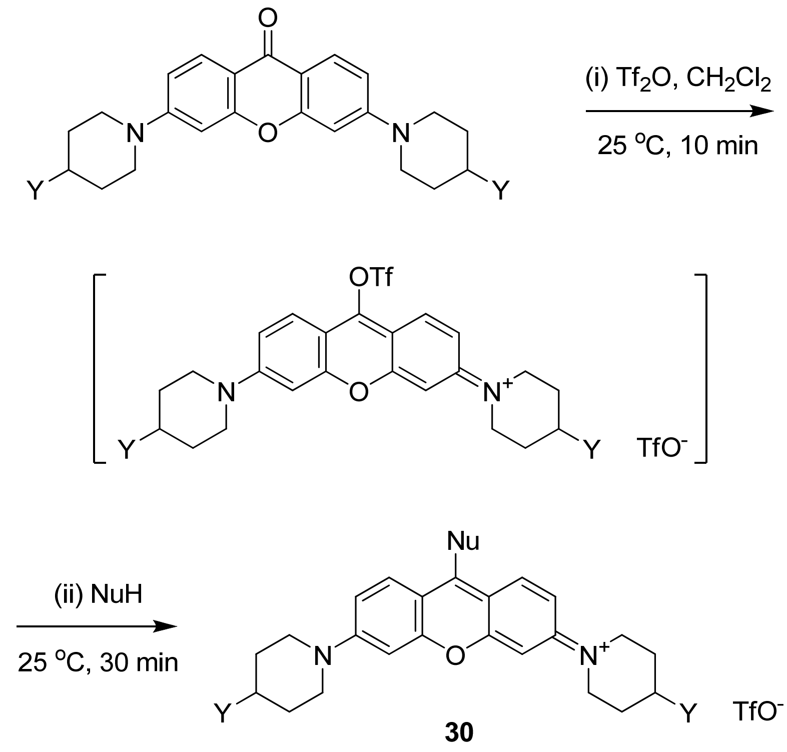
17, Y = H, 89%
29, Y = CO_2Me, 51%

Scheme 5.2. Failed attempt to synthesize new pyronin dyes.

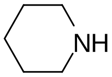
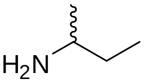
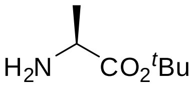
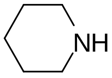


The ketone functionality of compounds **17** and **29** is less reactive because of the arylamine substituents; it may be considered as a vinylogous urea. Thus no reaction was observed when piperidine was heated with the ketones under mildly acidic conditions (Scheme 5.2.). Conversely, the nucleophilicity of the oxygen atom is increased for the same reason hence it can be triflated. These triflates were red fluorescent in CH_2Cl_2 but unfortunately they not stable. Thus they were not isolated and used to react with amines or benzyl thiol to give the fluorescent compounds **30a** – **f** (Table 5.1.). Reaction of the triflate with alcohols such as MeOH and EtOH also occurred smoothly to give yellow fluorescent solutions. However, the products formed with alcohols were not stable enough to be isolated.

Table 5.1. Synthesis of new pyronin dyes



(i) Tf_2O, CH_2Cl_2
25 °C, 10 min
OTf
TfO^-
(ii) NuH
25 °C, 30 min
Nu
Y
N
O
N^+
30

entry	Y	NuH	yield (%)
30a	H		81
30b	H	Et_2NH	61
30c	H		98
30d	H		80
30e	H	BnSH	23
30f	MeO_2C		86

A single-crystal X-ray structure of compound **30c** was obtained (Figure 5.1.). This provided two pieces of evidence that all three nitrogen atoms are involved in delocalization of the positive charge from the aromatic core. Firstly, all three nitrogens are essentially planar (deviation of the nitrogen away from the plane formed by the three substituents was observed to be 0.062 Å for the central amine and 0.175/0.223 Å for the terminal ones). Secondly, the deviation of that plane from the one which encapsulated the aromatic segment of the molecule was 9 - 22°, and the C-N distances for the bonds that connect the amine substituents to the pyronin core were 1.35 – 1.37 Å (whereas other C – N single bonds in the same molecule were 1.45 – 1.47 Å). These observations are indicative of lone pair donation from the pyronin amine-substituents into the core heterocyclic framework.

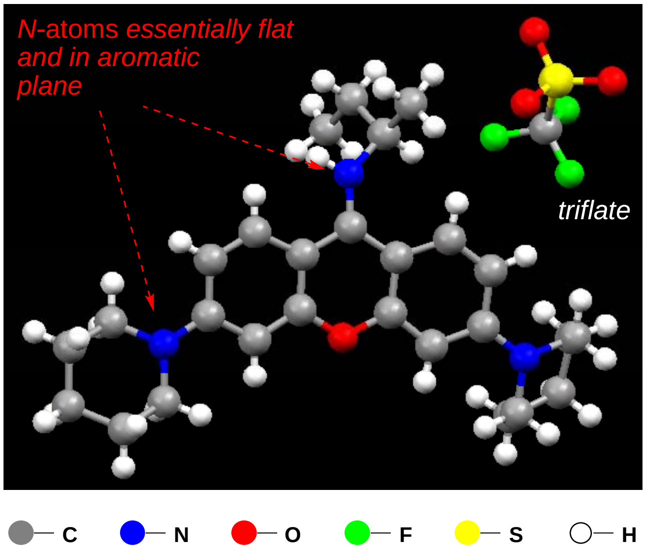


Figure 5.1. X-ray structure of novel pyronin **30c**.

5.3 Photophysical properties of new pyronin dyes

Selected spectral properties of the molecules **30** are shown in Table 5.2. The compounds formed from nitrogen nucleophiles, **30a** – **d** and **30f**, absorb in the range 456 – 499 nm with high extinction coefficients (51,000 – 62,000). They emit in the range 537 – 562 nm with quite sharp emissions (fwhm = 41 – 55 nm; *cf* the same parameter for rhodamine B in CH_2Cl_2 is 36 nm, and rhodamines are generally recognized to give sharp fluorescent emissions). Quantum yields are least in the series for the compounds from secondary amines (**30a**, **30b**, and **30f**), and these emit at longer wavelengths. Other compounds in the series that were derived from primary amines (**30c** and **30d**) had markedly higher quantum yields (0.85 and 0.90, respectively). The *S*-containing compound **30e** absorbs and emits at significantly longer wavelengths, and with a sharper output; for these reasons this is possibly the most interesting dye in the series.

Table 5.2. Spectral properties of **30** measured in CH_2Cl_2

	λ_{max} (nm)	ε_{max} ($L \cdot mol^{-1} \cdot cm^{-1}$)	λ_f (nm)	fwhm (nm) [a]	Φ_f [b]
30a	491	60400	562	55	0.14 ± 0.01
30b	499	51000	558	41	0.11 ± 0.02
30c	456	52400	537	52	0.85 ± 0.01
30d	462	62000	544	52	0.90 ± 0.02
30e	598	82600	619	38	0.79 ± 0.02 [c]
30f	482	54900	552	47	0.02 ± 0.01

[a] Full width at half-maximum height: a measurement of the sharpness of the fluorescence spectra; [b] Average of three measurements; Fluorescein in 0.1 M NaOH ($\Phi = 0.92$)[102] was used as a standard; [c] Cresyl violet in MeOH ($\Phi = 0.54$)[217] was used as a standard.

Normalized spectra that correspond to the data in Table 5.2. are shown in Figure 5.2. Compounds **30c** and **30d** absorb UV and fluoresce at shorter wavelengths than the others. Conversely, the absorption and fluorescence of compound **30e** is further to the red of the others in the series.

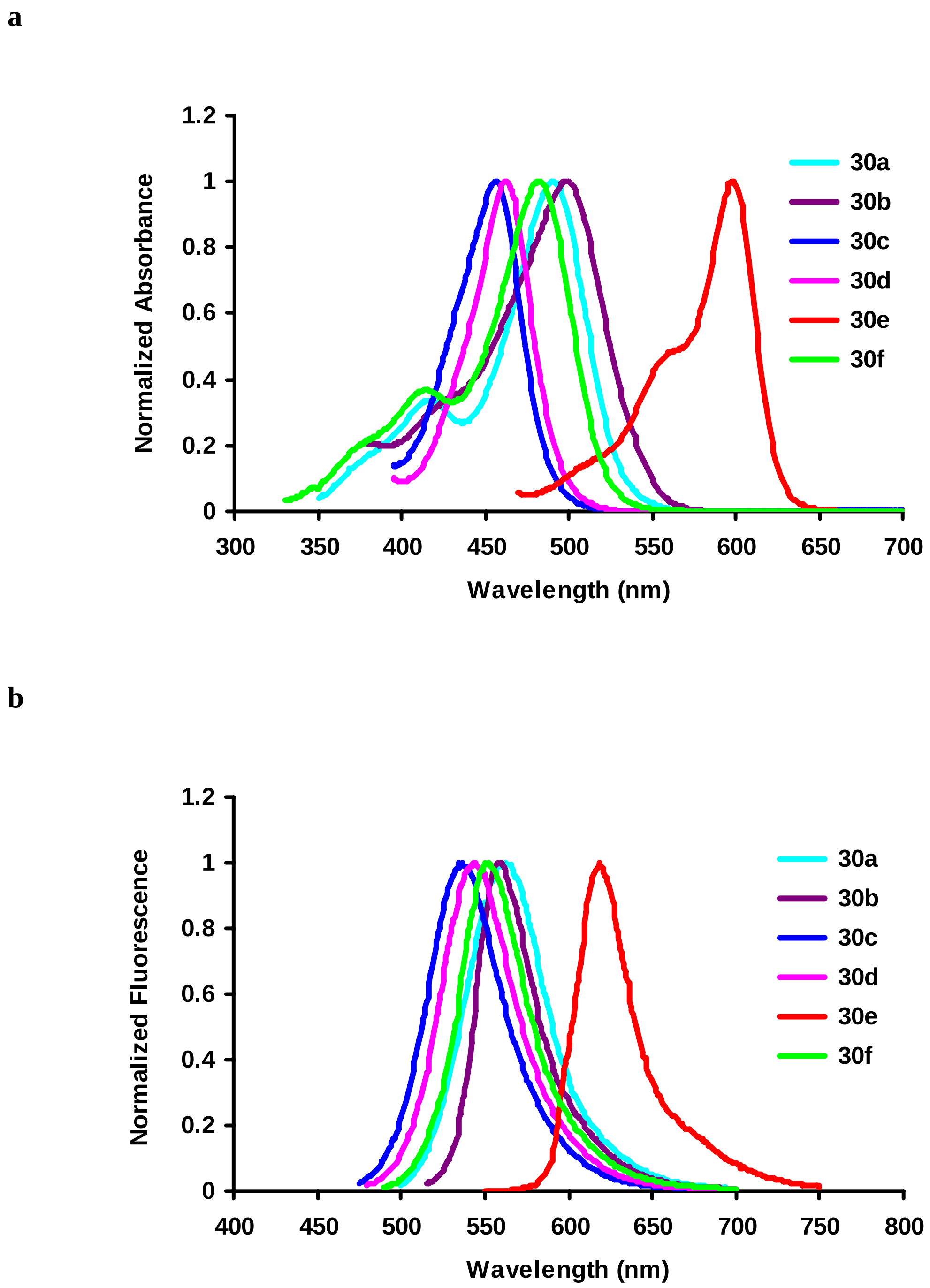


Figure 5.2. Normalized (**a**) UV absorbance and (**b**) fluorescence (excited at $\lambda_{\text{max abs}}$) of novel pyronins **30** in CH_2Cl_2 (*ca* 10^{-6} M for absorbance; 10^{-7} M for fluorescence).

Compounds **30a**, **30c**, and **30e** were selected for pH dependence studies; data for **30c** are shown here. The UV absorption of this compound was uneffected by addition of TFA, but was decreased by addition of NEt_3 (Figure 5.3.a and b). The fluorescent emission characteristics of this compound followed the same trend (Figure 5.3.c and d). Figure 5.3.e shows that the impact of NEt_3 on the fluorescence of this compound is essentially reversible. Dye **30a**, that has a tertiary amine *meso*-substituent, showed a different pH dependence (see Appendix Information) possibly because it does not have a relatively acidic proton. Thiol derivative **2e** showed no UV or fluorescence change on addition of TFA, but addition of NEt_3 caused decomposition (see Appendix Information).

a

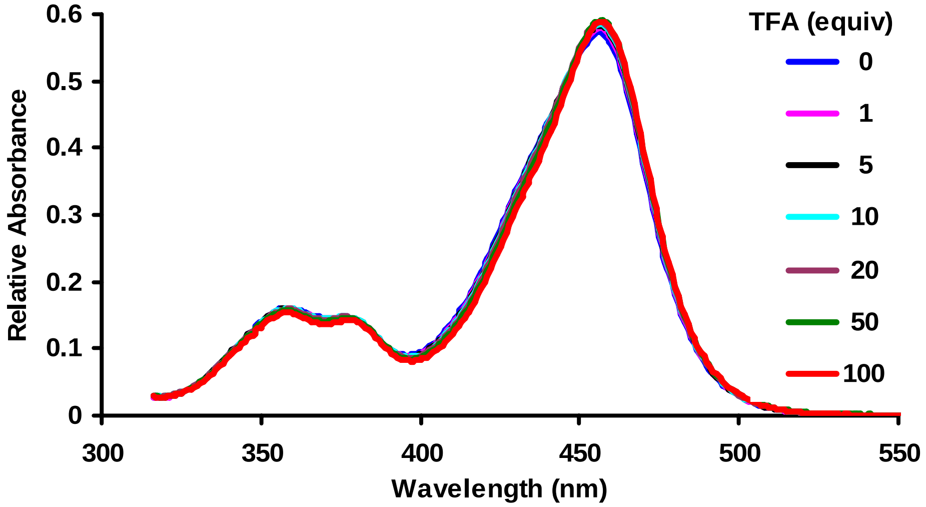
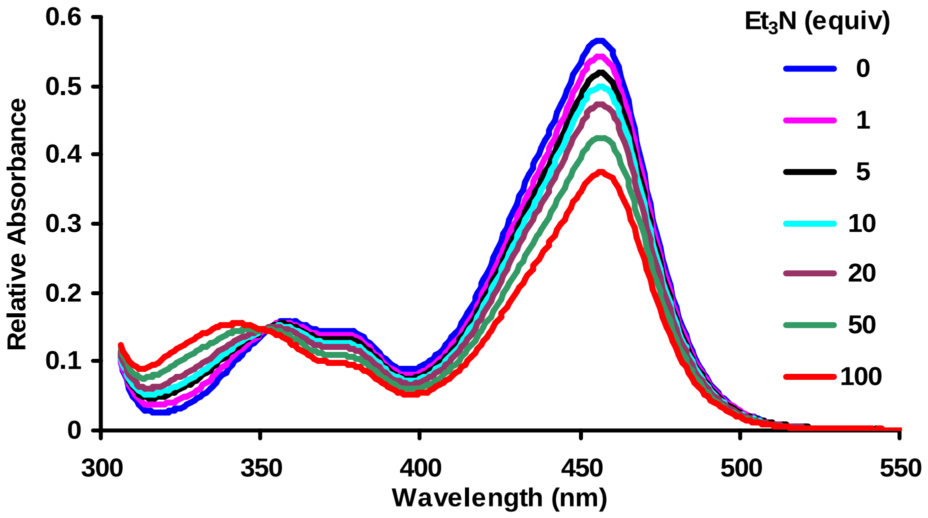


Figure 5.3. Dependence of electronic spectra of **30c** on pH. UV absorbance for **30c** under: **a** acidic; and, **b** basic conditions (CH_2Cl_2, 1.1 x 10^{-5} M). Fluorescence pH dependence of **30c** under: **c** acidic; and, **d** basic conditions (CH_2Cl_2, 2.5 x 10^{-7} M, excited at 456 nm). **e** Fluorescence intensity of **30c** under alternating basic (Et_3N) and acidic (TFA) conditions (CH_2Cl_2, 2.5 x 10^{-7} M, excited at 456 nm).

b



c

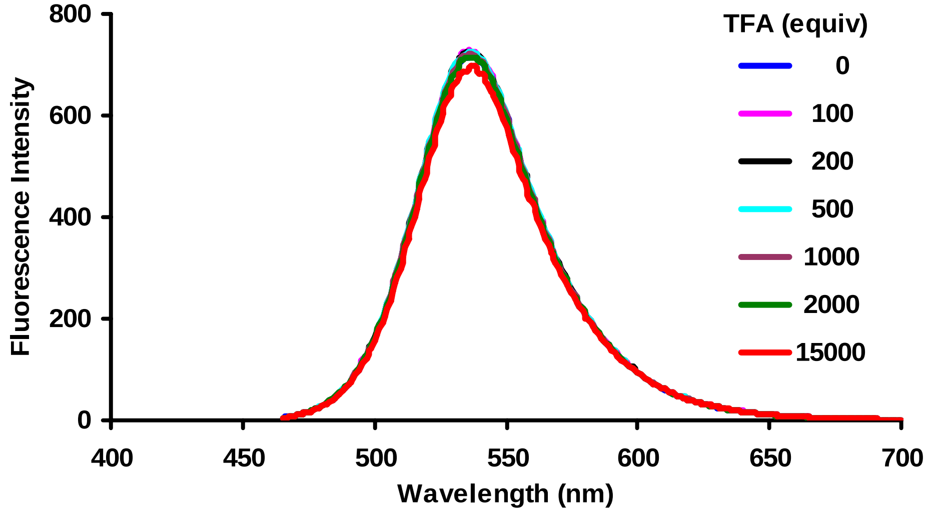
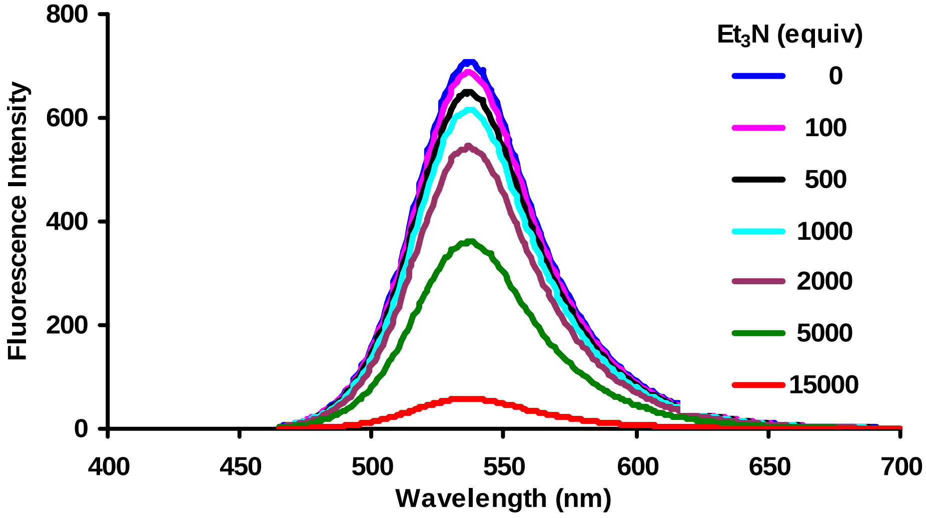


Figure 5.3. Continued.

d



e

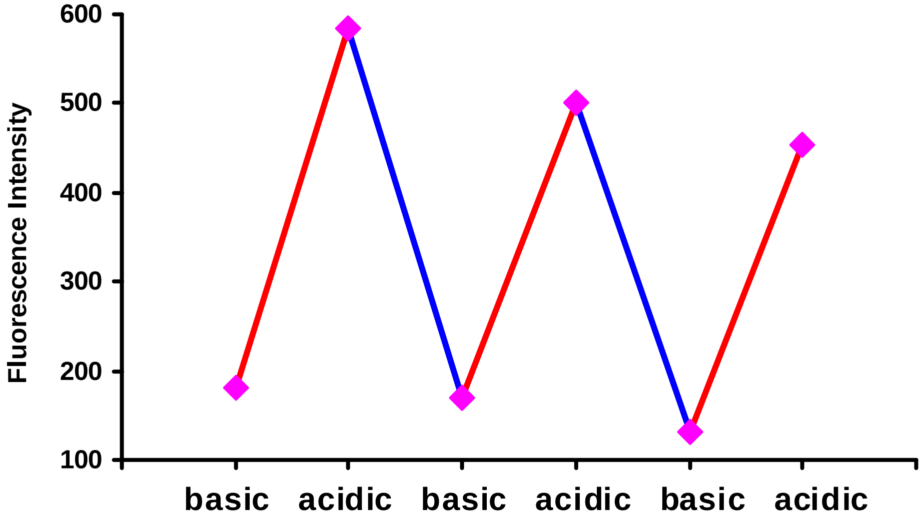
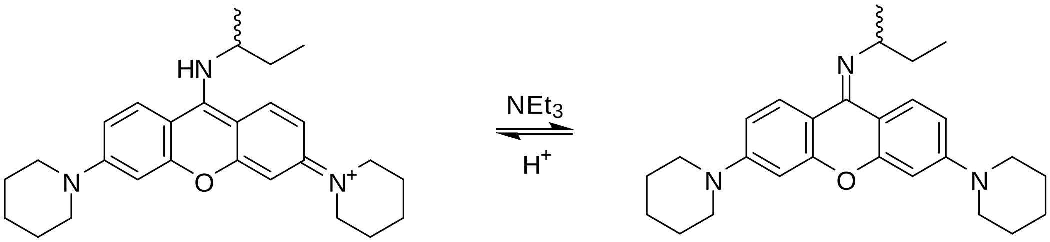


Figure 5.3. Continued.

The spectral data summarized above are consistent with reversible deprotonation of the central amine under mildly basic conditions. That proton loss gives a less conjugated and non-fluorescent imine derivative (Figure 5.4.).



30c
fluorescent form

30c
deprotonated
non-fluorescent form

Figure 5.4. Likely acid/base equilibrium for **30c**.

5.4 Conclusion

In summary, a direct route to novel heteroatom analogs of rhodamine and rosamine systems is described here. Spectral properties of the compounds are highly dependent on the nature of the nucleophile that is used to add to the central carbon atom. In that regard, the *S*-containing compound **30e** is of particular interest to the community seeking fluorophores that emit above the 600 nm mark. However, just like normal pyronin dyes, the stability of the new pyronin derivatives might be a concern for wide applications. Further improvement on the stability of such dyes is still needed.

CHAPTER VI

WATER-SOLUBLE THROUGH-BOND ENERGY TRANSFER CASSETTES: SYNTHESES, SPECTROSCOPIC PROPERTIES AND APPLICATIONS

6.1 Introduction

Multiplexing of fluorescent labels is a common requirement in biotechnology. A convenient approach to multiplex is to tag each component with a fluorescent label, and then excite the labels at a single wavelength using a commercial laser source *eg* an argon laser (488nm). The best-known examples of multiplexing are FRET[7] based cassettes in which the donor and acceptor units are connected via non-conjugated linkers (Figure 6.1.a). In these systems, a donor fragment harvests energy from an excitation source and transfers it through space, via dipole-dipole coupling, to an acceptor that then fluoresces. Well spaced fluorescence emissions over a range of wavelengths, *ie* resolution, could be achieved by choosing appropriate acceptors. Unfortunately, the sensitivity of emission with multiplexed, FRET-based cassettes is still constrained by the requirement that the fluorescence of the donor must overlap with the absorption of the acceptor. The dyes that emit at longer wavelengths usually absorb at the excitation wavelength less effectively, so their fluorescence intensities are diminished.

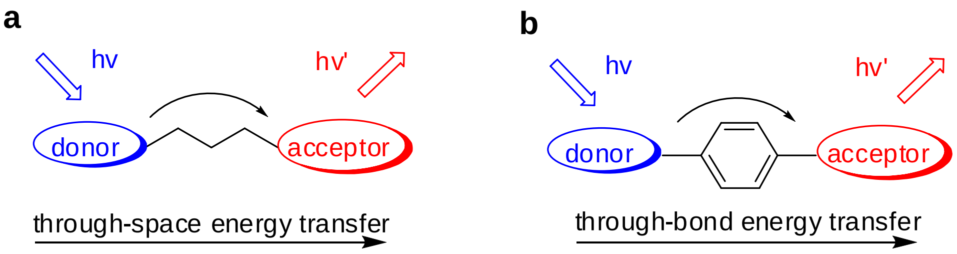


Figure 6.1. (**a**) Through-space and (**b**) through-bond energy transfer cassettes.

Our research focuses on energy transfer cassettes in which the donor is connected to acceptors via electronically conjugated linkers (Figure 6.1.b)[8,9,218-220] and energy transfer may occur via through-bond pathways such as Dexter energy transfer[12] and others. Through-bond energy transfer is not constrained by the same parameters such as spectra overlap as the corresponding through-space mechanisms. It may, therefore, be possible to separate the donor absorption from the acceptor fluorescence by many wavenumbers without compromising the intensity of the light emitted. Thus multiplexing with good resolution and high sensitivity may be achieved with sets of appropriate through-bond energy transfer cassettes.

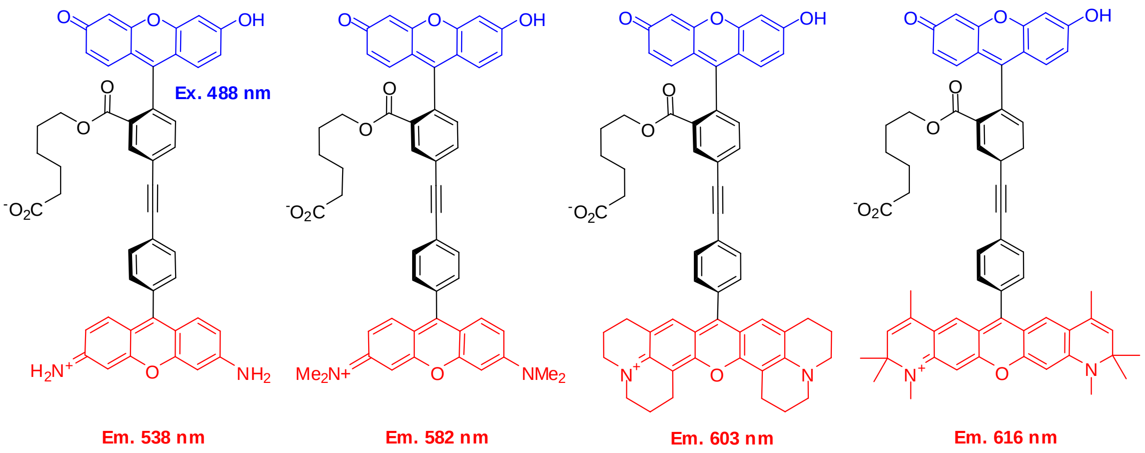
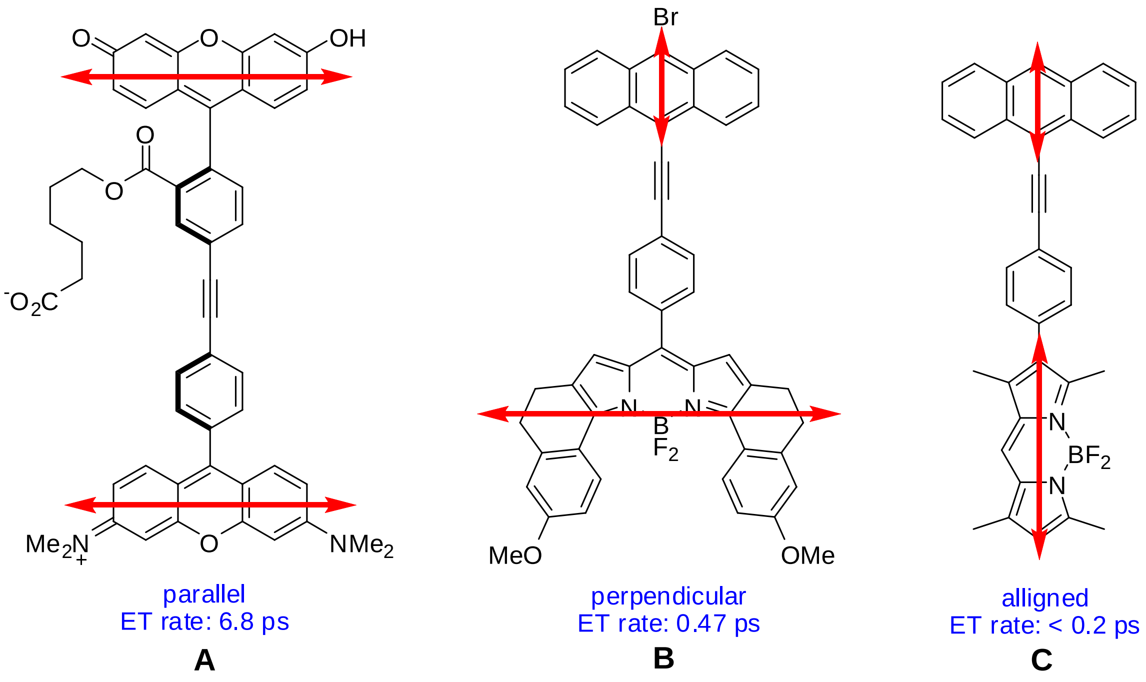


Figure 6.2. Through-bond energy transfer cassettes based on fluorescein and rosamines.

As a proof of concept, sets of through-bond energy transfer cassettes based on fluorescein donor and rosamine acceptors (Figure 6.2.) have been prepared and reported by previous group members.[11] These cassettes absorb strongly at 488 nm corresponding to the argon laser source. Excitation of the cassettes at 488 nm produces fluorescence characteristic of only the acceptor component indicating complete energy transfer in such systems. These cassettes fluoresce at 538, 582, 603 and 616 nm, respectively; hence their emissions are dispersed over a 78 nm wavelength range, and their fluorescence signals are well separated and easily differentiated. Moreover, these cassettes all fluoresce strongly when excited at the donor. Thus, these through-bond energy transfer cassettes have good resolution and sensitivity which are required for multiplexing. Further studies show that the energy transfer from donor to acceptor in these cassettes is extremely fast (6-7 ps) and is about four times faster than the corresponding FRET cassettes.[221] Due to the fast energy transfer, these cassettes are far more photostable than the fluorescein donor when excited at 488 nm. This can be explained as follows. Fluorescein decomposes via intersystem crossing to a relatively long-lived triplet state then reactions with surrounding molecules. The fluorescein donor in these cassettes, however, rapidly transfers energy to the more photostable rosamine acceptors before such intersystem crossing can occur.



parallel
ET rate: 6.8 ps
A

perpendicular
ET rate: 0.47 ps
B

alligned
ET rate: < 0.2 ps
C

Figure 6.3. Indication of the directions of the $S_1 \leftrightarrow S_0$ transition moments in different through-bond energy transfer cassettes.

In order to understand the through-bond energy transfer process and facilitate design of better through-bond energy transfer cassettes, studies on the correlation of energy-transfer rates with structures for a wide range of through-bond energy transfer cassettes were carried out by our group and collaborators.[10] Our observations indicate that energy transfer is fastest (< 0.2 ps) when the transition moments of the donor and acceptor are aligned with the linker axis as in C (Figure 6.3.). The study shows that the energy transfer in through-bond energy transfer cassettes such as A is much faster than the corresponding FRET cassettes (6.8 and 35 ps, respectively). Several findings also suggest that Förster energy transfer is not the prevalent mechanism in such cassettes. First, the transition dipole moments of the donor and acceptor in cassettes like B are perpendicular to each other. According to Förster theory, FRET should not occur. However, complete and ultrafast (0.47 ps) energy transfer was observed in cassette B and similar derivatives. Second, the studies showed no systematic correlation of the energy transfer rate with the nature of the acceptor, despite the difference in donor-acceptor energy gap. This also argues against the FRET mechanism, which depends critically on the overlap integral of the donor emission and acceptor absorption spectra. Third, experiments with cassettes having reduced donor-acceptor distances showed no significant increase in transfer rates, as would be predicted from Förster theory.

Energy transfer cassettes designed to transfer energy through bonds will always transfer a certain amount of the energy through space as well. Unfortunately, it is usually not possible to ascertain how much energy transfer occurs via through-bond mechanisms relative to the through-space pathways. This is very likely to be the case when the donor and acceptor fragments are relatively close together. Nevertheless, the overall rates of energy transfer can be measured; they in turn will be influenced by the structure of the donor, acceptor, linker fragments and the orientation of the donor-linker-acceptor connectivity.

All of the above mentioned through-bond energy transfer cassettes have poor water solubilities and this is likely to make their conjugation to biomolecules difficult, thus limits their applications in biological studies. The first water-soluble through-bond energy transfer cassette (Figure 6.4.) was reported by our group in 2006.[222] This cassette was successfully attached to a protein (ACBP) and the conjugate was transferred back into cell by a cell-penetrating peptide Chariot. The conjugate was successfully detected in COS-7 cells by fluorescence microscope. Even though it is water-soluble and has been successfully used for intracellular imaging, this cassette still has several disadvantages. First, the energy transfer in the conjugate is only about 60 %. Secondly, the emission of the acceptor is less than 600 nm and is still complicated by the cell autofluorescence. Third, there are multiple carboxylic acid handles in the cassettes which may react with proteins and cause cross coupling between different protein molecules thus might affect the protein structure and functions. Last, the synthesis of this cassette is extremely difficult, giving very low yield and requiring tedious reverse phase HPLC purification. So further effort is still needed to design and prepare better water-soluble through-bond energy transfer cassettes.

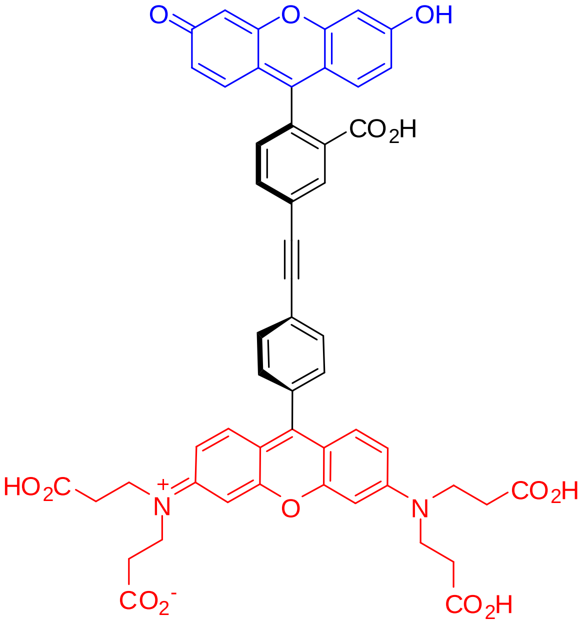
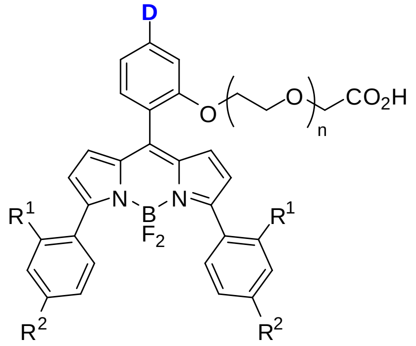


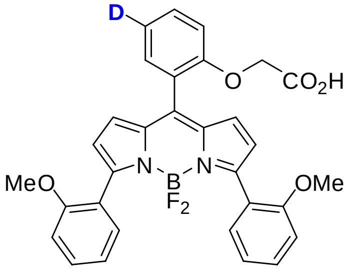
Figure 6.4. A water-soluble through-bond energy transfer cassette.

This chapter describes the design and syntheses of a series of new water-soluble through-bond energy transfer cassettes **31-33** (Figure 6.5.). Correlation between optical properties and cassette structures is investigated. Applications of these new water-soluble cassettes are also demonstrated. Thus, the specific aims of this project are to:

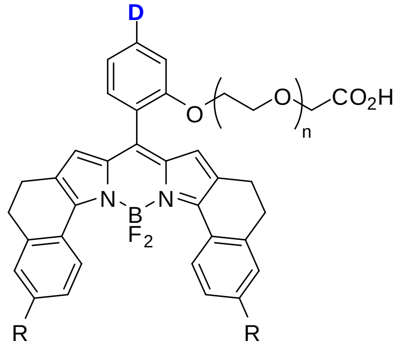
(i) design and synthesize water-soluble cassettes that emit above 600 nm
(ii) prepare sets of water-soluble cassettes for multiplexing
(iii) study the correlation between optical properties and cassette structures
(iv) demonstrate applications in biomolecular labeling and intracellular imaging
(v) monitor protein-protein interactions with these new cassettes


D
O
O
n
CO_2H
N
B
N
F_2
R^1
R^1
R^2
R^2

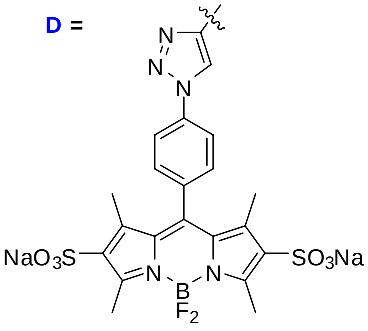
31a R^1 = OMe, R^2 = H, n = 0;
31b R^1 = $O(CH_2CH_2O)_3Me$, R^2 = H, n = 6;
31c R^1 = H, R^2 = $O(CH_2CH_2O)_3Me$, n=6

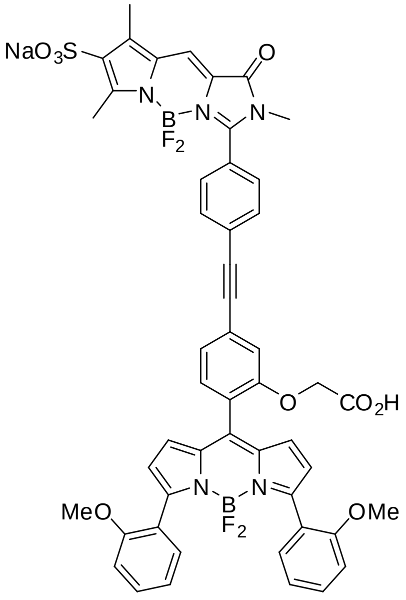

D
O
CO_2H
MeO
OMe
N
B
N
F_2

31d

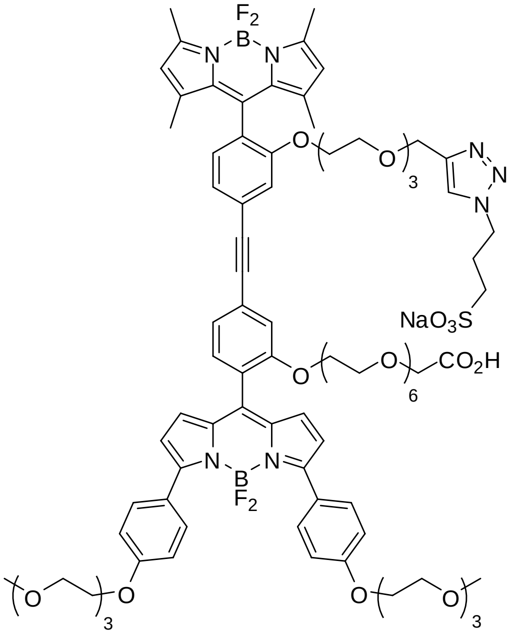

D
O
O
n
CO_2H
N
B
N
F_2
R
R

31e R = H, n = 0;
31f R = OMe, n = 0;
31g R = $O(CH_2CH_2O)_6Me$, n = 6


D =
N
N
N
NaO_3S
SO_3Na
N
B
N
F_2


NaO_3S
O
N
B
N
N
F_2
O
CO_2H
MeO
OMe
N
B
N
F_2

32


F_2
B
N
N
O
O
3
N
N
N
NaO_3S
O
O
6
CO_2H
N
B
N
F_2
O
O
3
O
O
3

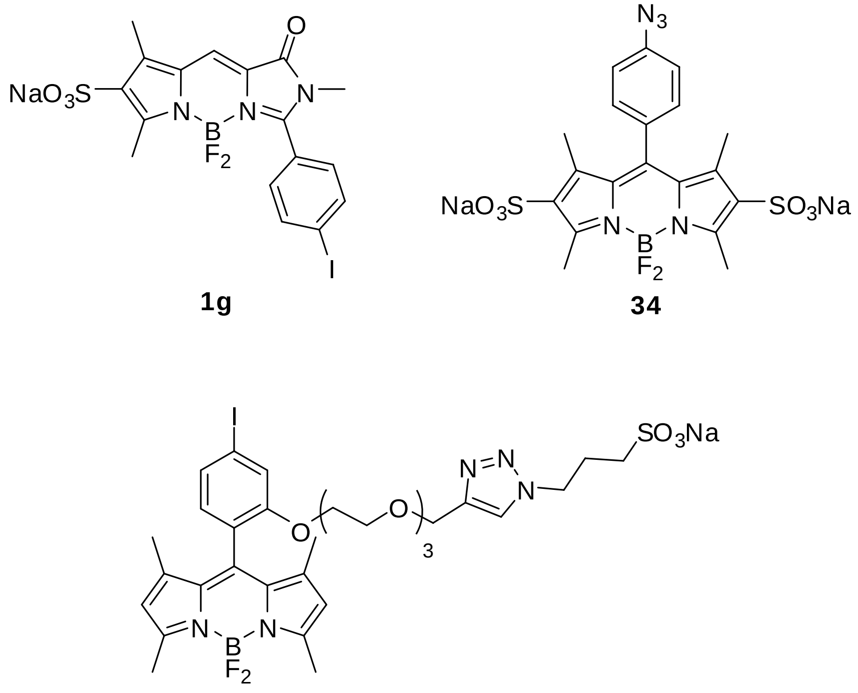
33

Figure 6.5. Structures of through-bond energy transfer cassettes **31-33**.

6.2 Syntheses of water-soluble through-bond energy transfer cassettes

6.2.1 Syntheses of donor fragments

The structures of the donor fragments that are used to prepare cassette **31-33** are showed in Figure 6.6. They are soluble in water and absorb strongly at *ca* 500 nm which makes them ideal donors for the water-soluble cassettes.

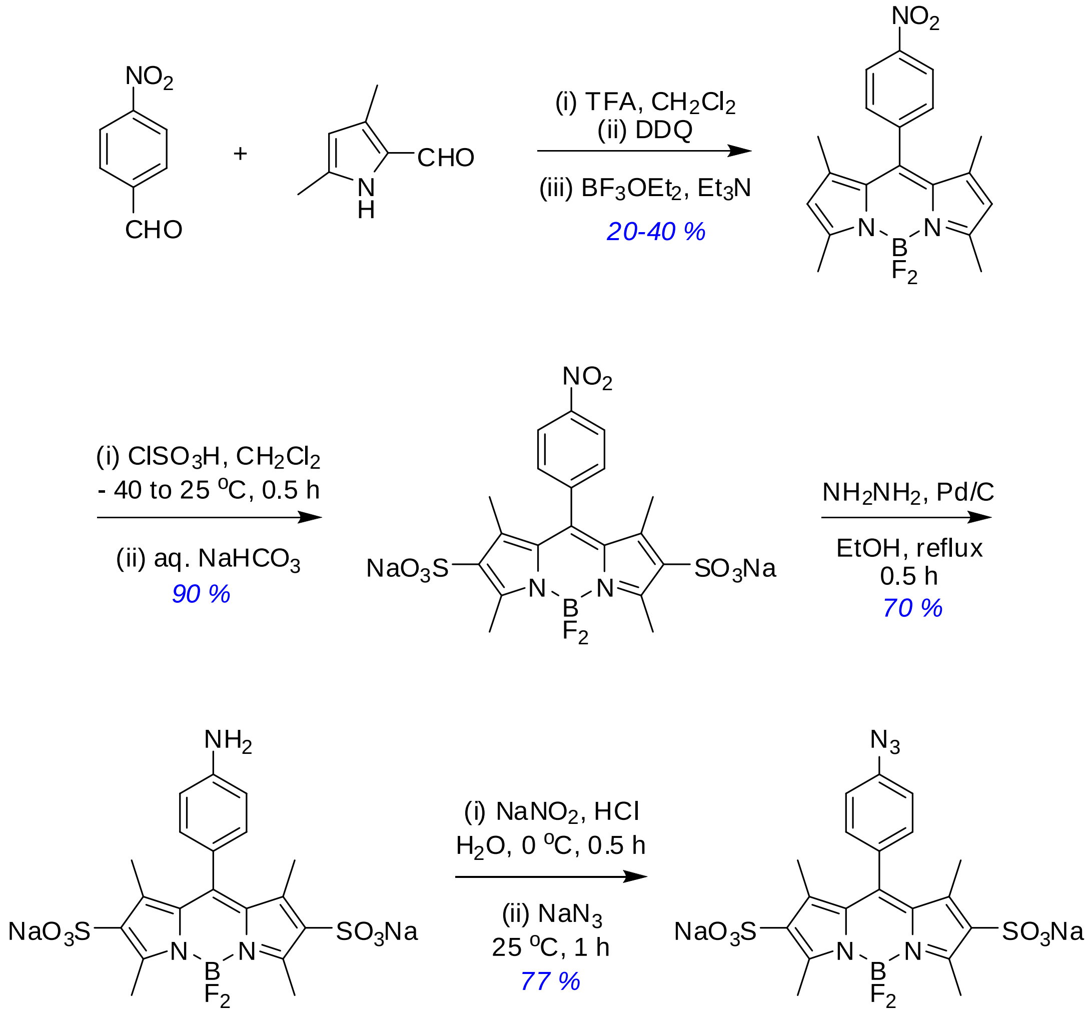


1g **34** **35**

Figure 6.6. Donor fragments for water-soluble cassettes.

The synthesis and properties of compound **1g** were described in Chapter II.[223] This compound is soluble in water and highly fluorescent. BODIPY **34** was prepared and studied by a previous group member.[224] The synthesis of this compound is summarized in Scheme 6.1.

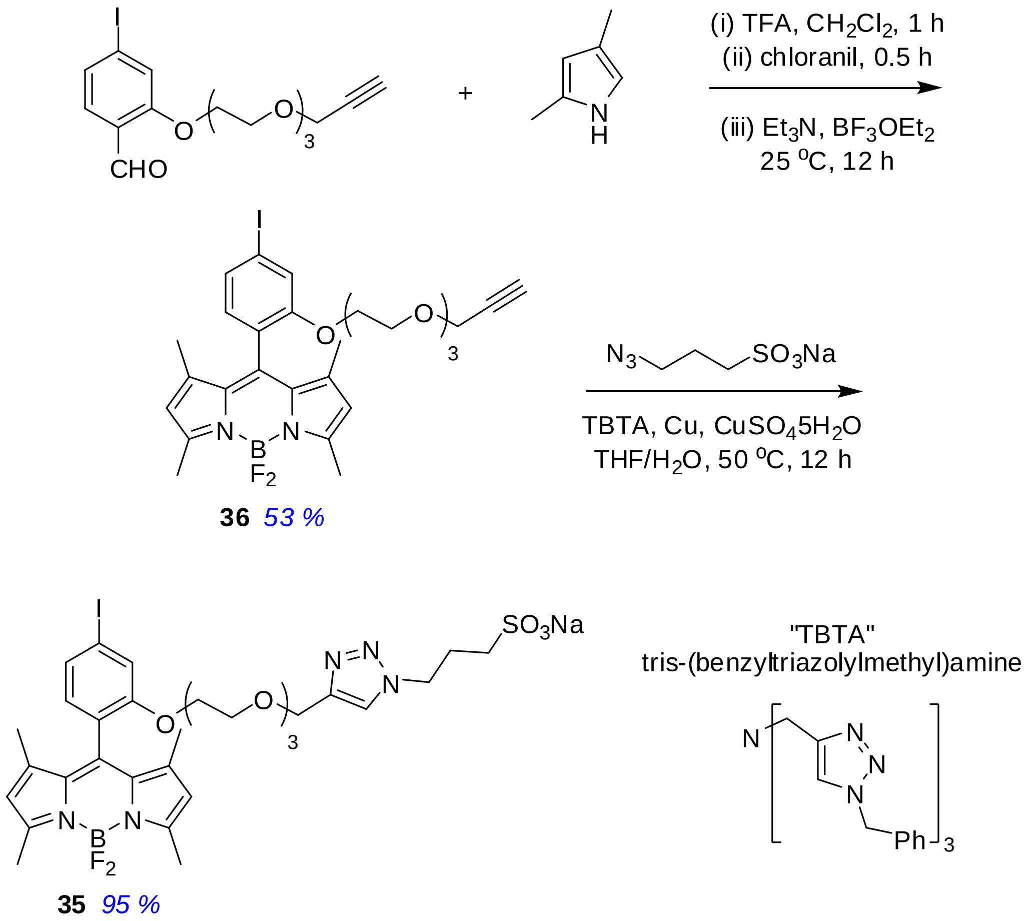
Scheme 6.1. Synthesis of donor fragment **34**.


NO2
CHO
+
CHO
N
H
(i) TFA, CH2Cl2
(ii) DDQ
(iii) BF3OEt2, Et3N
20-40 %
NO2
N
B
N
F2
(i) ClSO3H, CH2Cl2
- 40 to 25 °C, 0.5 h
(ii) aq. NaHCO3
90 %
NO2
NaO3S
SO3Na
N
B
N
F2
NH2NH2, Pd/C
EtOH, reflux
0.5 h
70 %
NH2
NaO3S
SO3Na
N
B
N
F2
(i) NaNO2, HCl
H2O, 0 °C, 0.5 h
(ii) NaN3
25 °C, 1 h
77 %
N3
NaO3S
SO3Na
N
B
N
F2

34

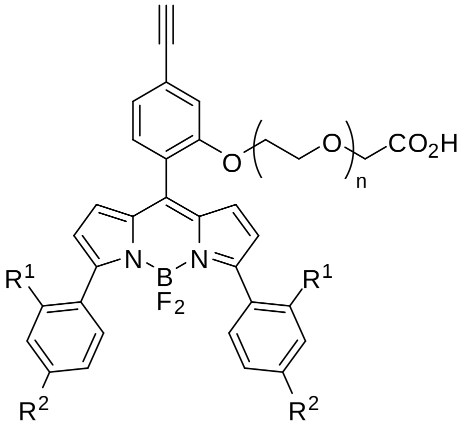
Scheme 6.2. describes the synthesis of a new compound **35**. Condensation of 2,4-dimethyl-pyrrole with a benzaldehyde derivative gave BODIPY **36**. Functionalization of **36** via "Click" reaction[225-227] leads to the water soluble derivative **35**. Interestingly, the click reaction did not proceed at room temperature even in the presence of tris-(benzyltriazolylmethyl)amine (TBTA) ligand. However, the desired product **35** was obtained in excellent yield after raising the temperature to 50 °C.

Scheme 6.2. Synthesis of donor **35**.

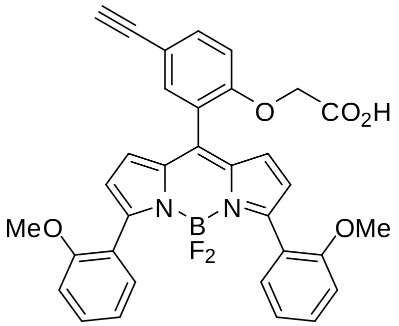

I
CHO
O
O
3
+
N
H
(i) TFA, CH$_2$Cl$_2$, 1 h
(ii) chloranil, 0.5 h
(iii) Et$_3$N, BF$_3$OEt$_2$
25 ^{o}C, 12 h
N$_3$
SO$_3$Na
TBTA, Cu, CuSO$_4$5H$_2$O
THF/H$_2$O, 50 ^{o}C, 12 h
N
B
F$_2$
36 53 %
N=N
SO$_3$Na
35 95 %
"TBTA"
tris-(benzyltriazolylmethyl)amine
Ph
3

6.2.2 Syntheses of acceptor parts

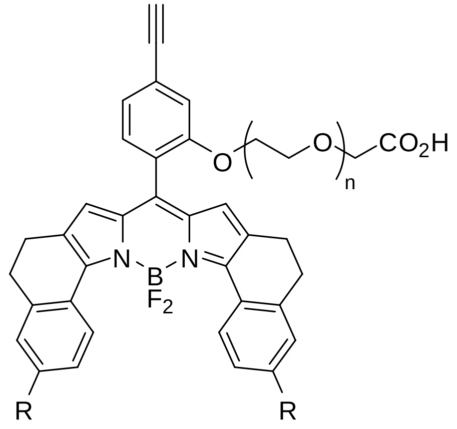
A series of BODIPYs **37** that emit above 600 nm have been prepared and exploited as acceptor fragments for water-soluble through-bond energy transfer cassettes. The structures of these BODIPY acceptors are summarized in Figure 6.7. Similar BODIPYs have been reported in literature.[228-230] However, the ones shown here all have a carboxylic acid handle which can be used as an attach point for coupling to biomolecules. They also have an alkyne functional group which allows further derivatization via cross coupling reactions such as Sonogashira and “Click” reactions to prepare through-bond energy transfer cassettes. Moreover, some of the BODIPY acceptors are water-soluble thus can be used for protein labeling and intracellular imaging.



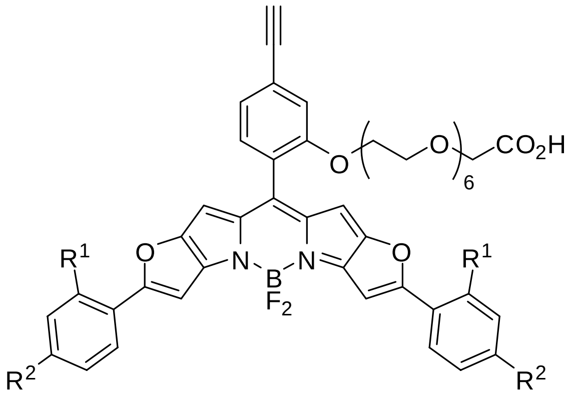
37a R^1 = OMe, R^2 = H, n = 0;
37b R^1 = $O(CH_2CH_2O)_3Me$, R^2 = H, n = 6;
37c R^1 = H, R^2 = $O(CH_2CH_2O)_3Me$, n=6



37d



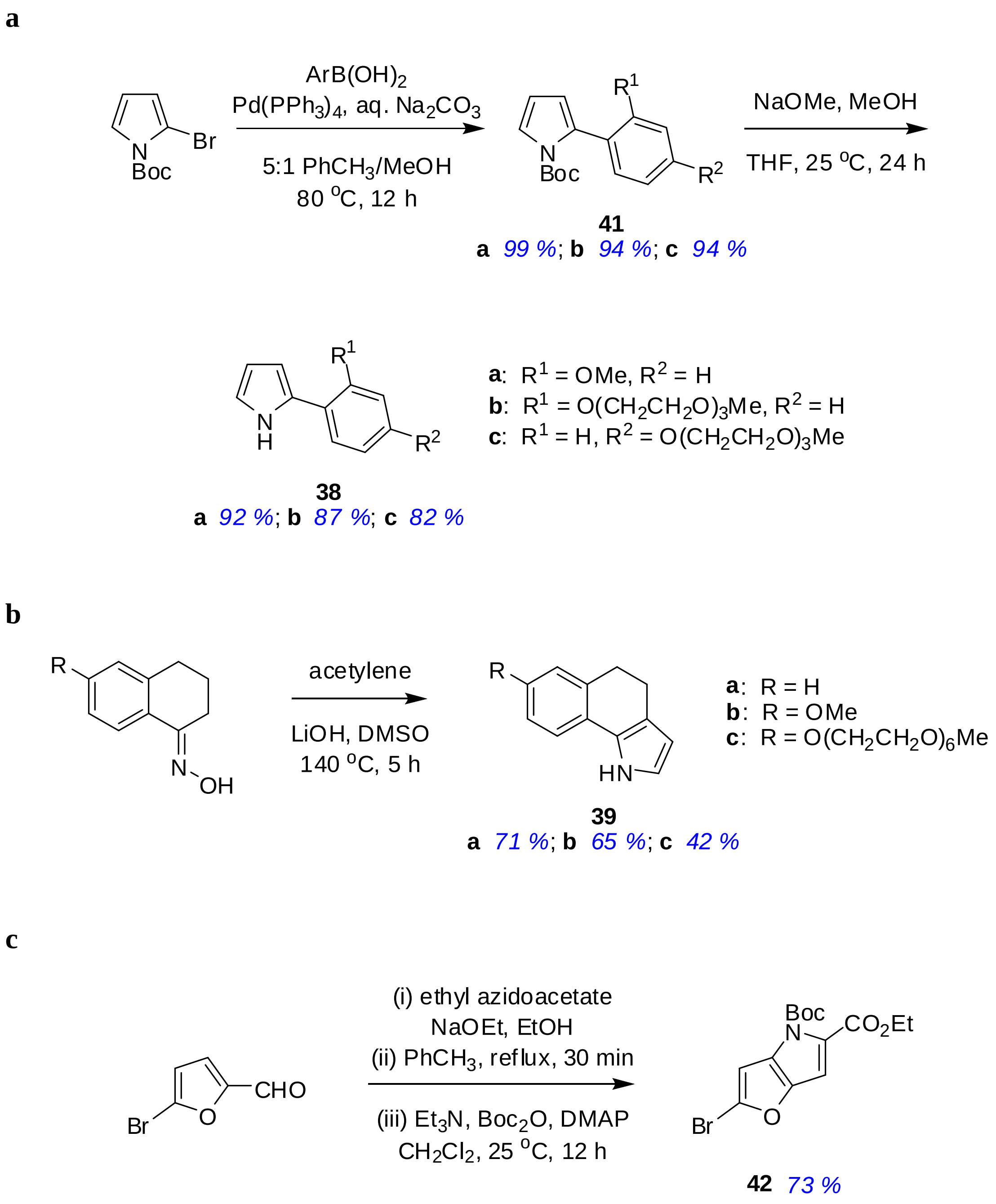
37e R = H, n = 0;
37f R = OMe, n = 0;
37g R = $O(CH_2CH_2O)_6Me$, n = 6



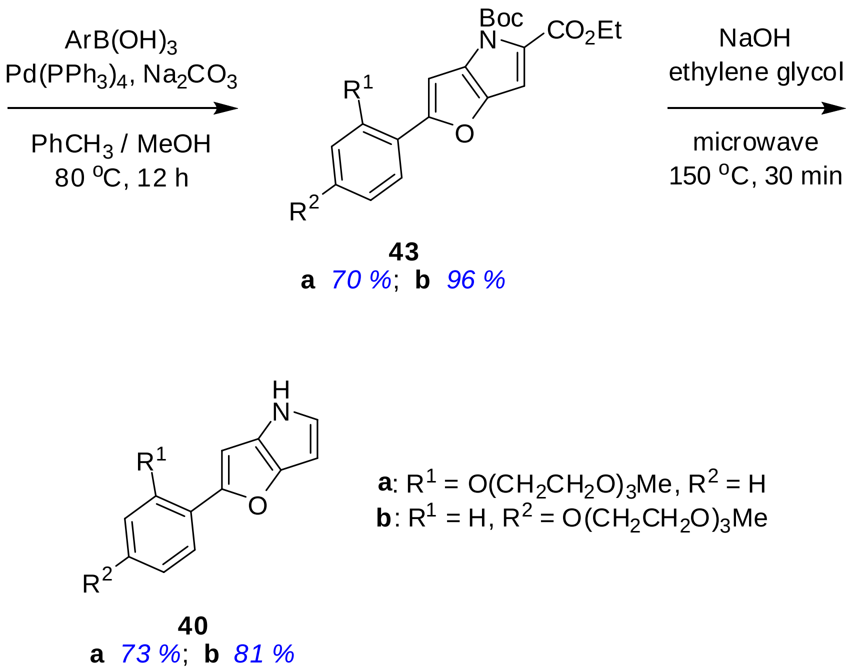
37h R^1 = $O(CH_2CH_2O)_3Me$, R^2 = H;
37i R^1 = H, R^2 = $O(CH_2CH_2O)_3Me$

Figure 6.7. BODIPY acceptor fragments for cassettes.

Scheme 6.3. Syntheses of pyrrole fragments **38-40**.

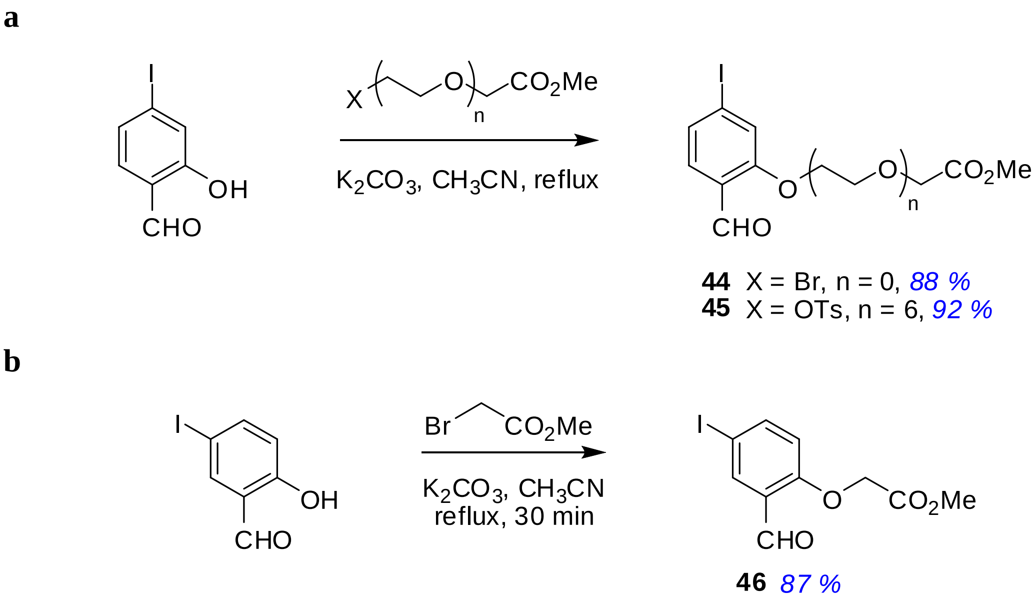

a
Br
N
Boc
ArB(OH)2
Pd(PPh3)4, aq. Na2CO3
5:1 PhCH3/MeOH
80 °C, 12 h
R1
R2
41
a 99 %; b 94 %; c 94 %
NaOMe, MeOH
THF, 25 °C, 24 h
N
H
38
a 92 %; b 87 %; c 82 %
a: R1 = OMe, R2 = H
b: R1 = O(CH2CH2O)3Me, R2 = H
c: R1 = H, R2 = O(CH2CH2O)3Me
b
R
N
OH
acetylene
LiOH, DMSO
140 °C, 5 h
HN
39
a 71 %; b 65 %; c 42 %
a: R = H
b: R = OMe
c: R = O(CH2CH2O)6Me
c
Br
O
CHO
(i) ethyl azidoacetate
NaOEt, EtOH
(ii) PhCH3, reflux, 30 min
(iii) Et3N, Boc2O, DMAP
CH2Cl2, 25 °C, 12 h
Boc
CO2Et
42 73 %

Scheme 6.3. Continued.


ArB(OH)3
Pd(PPh3)4, Na2CO3
PhCH3 / MeOH
80 °C, 12 h
Boc
N
CO2Et
R1
R2
O
43
a 70 %; b 96 %
NaOH
ethylene glycol
microwave
150 °C, 30 min
H
N
R1
R2
O
40
a 73 %; b 81 %
a: R1 = O(CH2CH2O)3Me, R2 = H
b: R1 = H, R2 = O(CH2CH2O)3Me

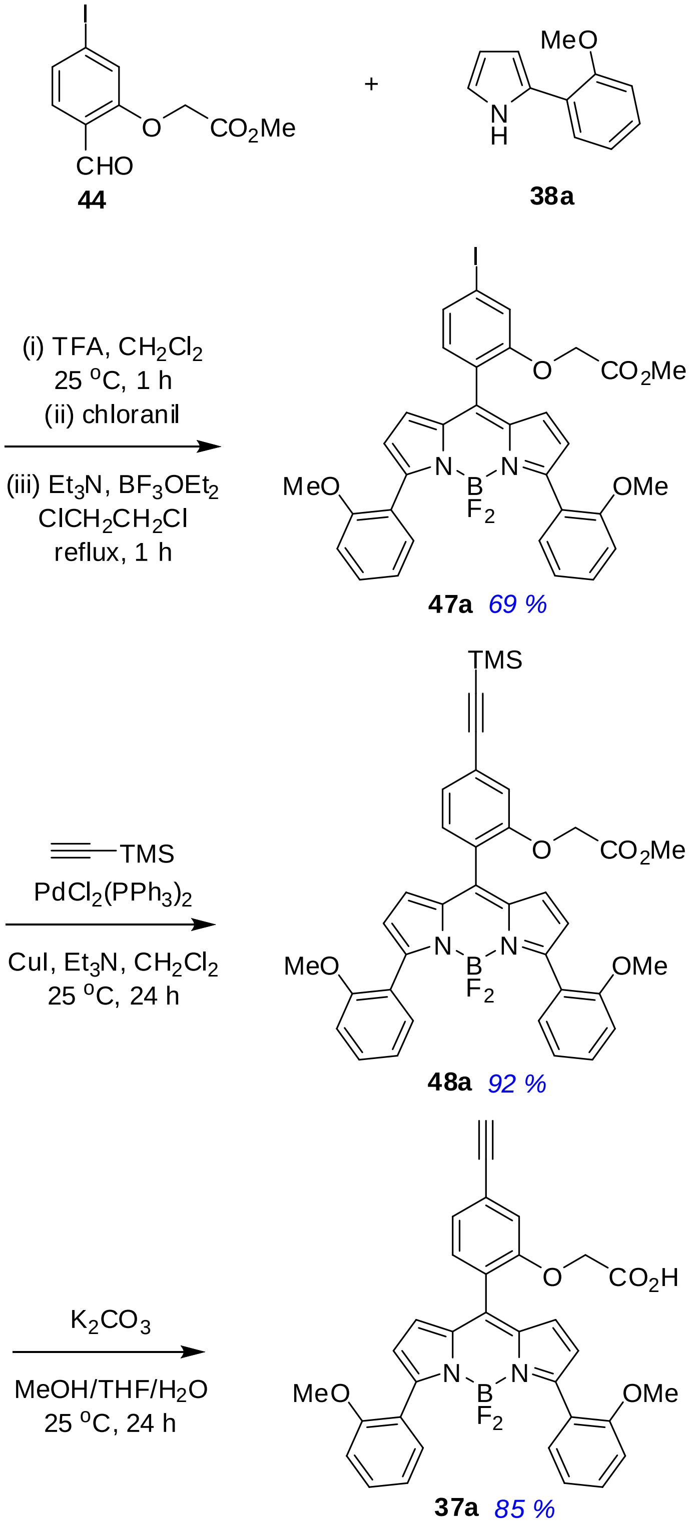
The first steps necessary for syntheses of BODIPY acceptors **37** were preparation of the starting pyrroles **38-40** and iodobenzaldehyde derivatives **44-46**. Scheme 6.3. describes the syntheses of the pyrrole derivatives. Pyrrole **38a** is a known compound and its synthesis has been reported.[229] Compounds **38b-c** were prepared using similar conditions reported for **38a** (Scheme 6.3.a). Pyrrole **39c** was synthesized according to a known procedure[228] developed for **39a-b** (Scheme 6.3.b). Analogs of pyrrole **40a-b** have also been reported.[230] Attempt to synthesize **40a-b** using the literature procedures failed to give the final products. Thus a new approach was devised (Scheme 6.3.c). In the new synthesis, a common intermediate furapyrrole was synthesized from 5-bromo-2-furaldehyde[231] and then protected by tert-butyloxylcarbonyl (Boc) group to afford compound **42** which was stable and easy to manipulate. Suzuki coupling of **42** with arylboronic acid followed by decarboxylation and removal of Boc under microwave irradiation gave pyrroles **40a-b**. Compared with the reported conditions, the new approach described here is more convergent and easier to work with because the intermediates involved are stable and easy to handle. Scheme 6.4. shows the syntheses of iodobenzaldehyde derivatives **44-46**. They were obtained in high yields by alkylation of the phenol precursors.

Scheme 6.4. Syntheses of iodo-benzaldehyde fragments **44-46**.


a
I
OH
CHO
X
O
CO2Me
n
K2CO3, CH3CN, reflux
I
O
O
n
CO2Me
CHO
44 X = Br, n = 0, 88 %
45 X = OTs, n = 6, 92 %
b
I
OH
CHO
Br
CO2Me
K2CO3, CH3CN
reflux, 30 min
I
O
CO2Me
CHO
46 87 %

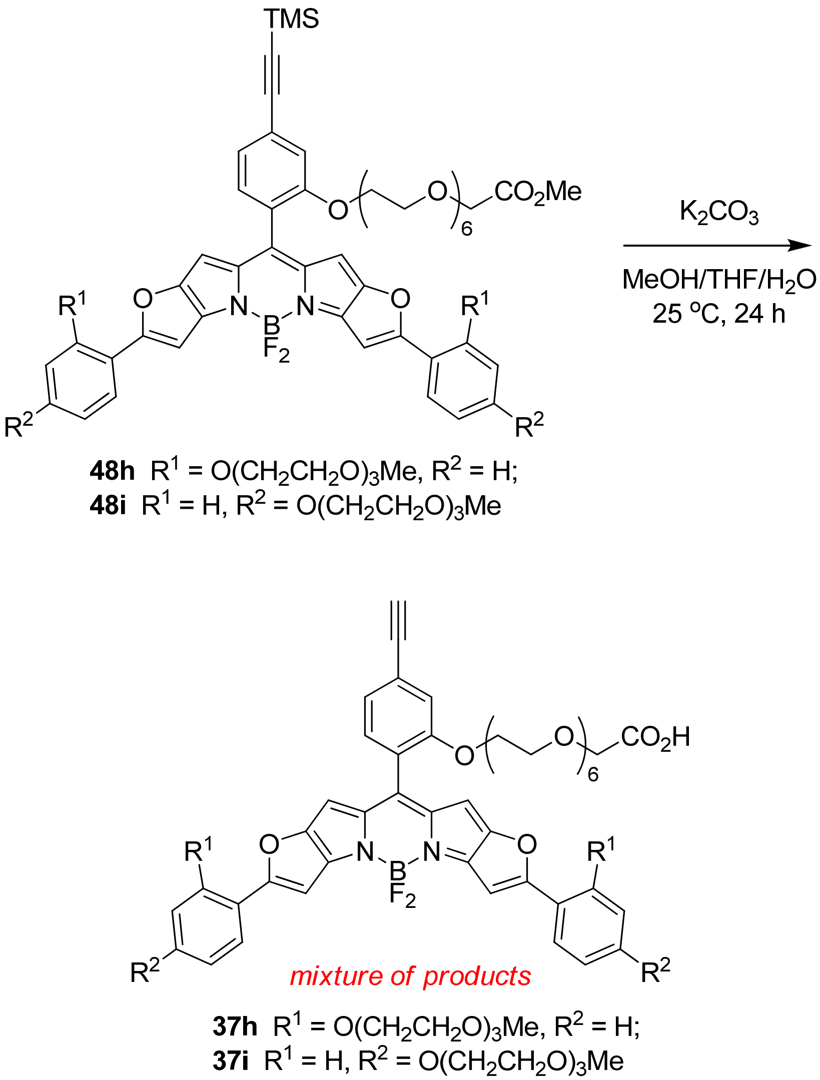
With the pyrroles and iodobenzaldehydes in hand, BODIPY acceptors can be prepared from corresponding precursors by standard BODIPY synthesis. Illustrative synthesis of a BODIPY acceptor **37a** is showed in Scheme 6.5. Condensation of pyrrole **38a** with aryl aldehyde **44** followed by oxidation with chloranil or DDQ afforded the dipyrromethene (DIPY) intermediate, which was passed through a short pad of alumina to remove most of the by-products. Subsequent chelation of DIPY with boron gave the BODIPY **47a**. Palladium catalyzed Sonogashira reaction of iodo-BODIPY **47a** with trimethylsilyl acetylene occurred smoothly, giving compound **48a** in excellent yield. Removal of trimethylsilyl protecting group and hydrolysis of methyl ester in BODIPY **48a** were achieved in a single step by stirring a solution of this compound in a mixture of solvents (MeOH/THF/H_2O) in the presence of potassium carbonate. BODIPY acceptors **37b-g** were successfully prepared using similar conditions.

Scheme 6.5. Illustrative synthesis of BODIPY acceptor fragment **37a**.


I
O
CO2Me
CHO
44
+
MeO
N
H
38a
(i) TFA, CH2Cl2
25 °C, 1 h
(ii) chloranil
(iii) Et3N, BF3OEt2
ClCH2CH2Cl
reflux, 1 h
47a 69 %
TMS
PdCl2(PPh3)2
CuI, Et3N, CH2Cl2
25 °C, 24 h
48a 92 %
K2CO3
MeOH/THF/H2O
25 °C, 24 h
CO2H
N B N
F2
MeO
OMe
37a 85 %

There was a problem for the synthesis of BODIPY acceptors **37h-i**. Compounds **48h-i** were efficiently prepared using the above described procedures. However, the deprotection step did not work well for **48h-i**. A mixture of products were formed when compound **48h-i** were stirred with K_2CO_3 in MeOH/THF/H_2O (Scheme 6.6.). Attempt to isolate products **37h-i** from the reaction mixture was failed due to close polarity of the product with unknown by-products. Further optimization of the reaction was not carried out at this stage.

Scheme 6.6. Syntheses of acceptors **37h-i**.


TMS
O
O
6
CO_2Me
O
N
B
F_2
N
O
R^1
R^1
R^2
R^2
K_2CO_3
MeOH/THF/H_2O
25 °C, 24 h
O
O
6
CO_2H
mixture of products

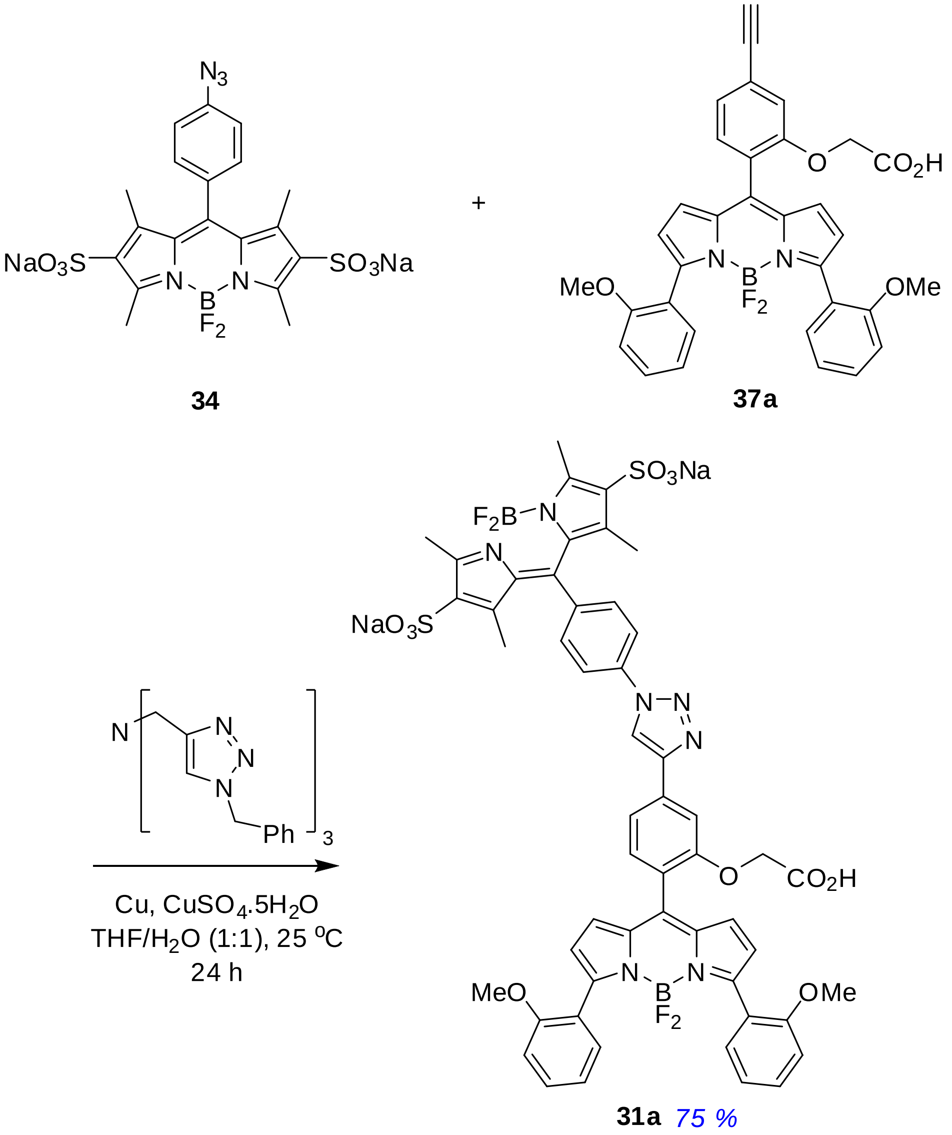
48h $R^1 = O(CH_2CH_2O)_3Me$, $R^2 = H$;
48i $R^1 = H$, $R^2 = O(CH_2CH_2O)_3Me$

37h $R^1 = O(CH_2CH_2O)_3Me$, $R^2 = H$;
37i $R^1 = H$, $R^2 = O(CH_2CH_2O)_3Me$

6.2.3 Syntheses of cassettes

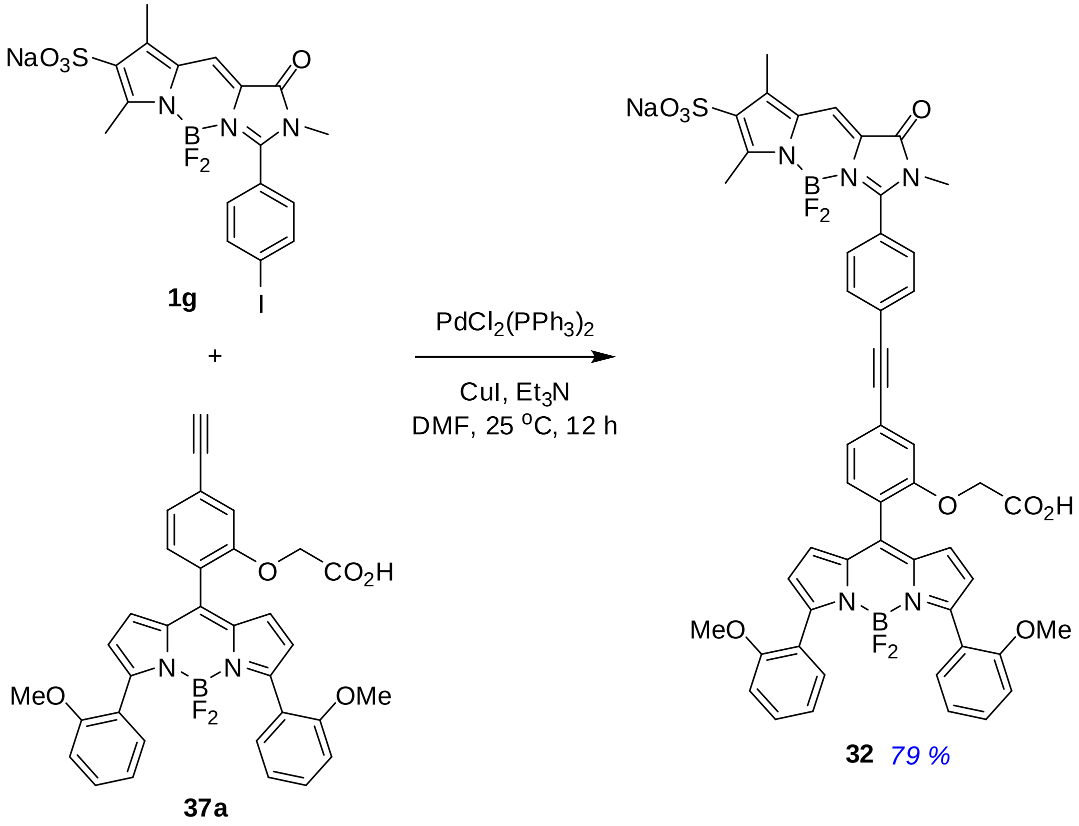
Cassettes **31** were prepared via copper-mediated azide/alkyne cycloadditions, which is also called "Click" reaction, from the corresponding azido-BODIPY donor **34** and alkynyl-BODIPY acceptors **37**. The click reaction occurred smoothly in the presence of a tris-(benzyltriazolylmethyl)amine "TBTA" ligand[232] to give a 1,2,3-triazole based linker. One illustrative systhesis of such cassettes was described for **31a** in Scheme 6.7. Cassettes **31b-g** were prepared using similar conditions.

Scheme 6.7. Illustrative synthesis of BODIPY cassettes.


N_3
NaO_3S
SO_3Na
N
B
N
F_2
34
+
O
CO_2H
MeO
OMe
N
B
N
F_2
37a
N
N
N
N
Ph
3
Cu, $CuSO_4.5H_2O$
THF/H_2O (1:1), 25 °C
24 h
SO_3Na
F_2B
N
N
NaO_3S
N
N
N
O
CO_2H
MeO
OMe
N
B
N
F_2
31a 75 %

Cassettes **32** and **33** were prepared to test the effect of the linkers and donor parts on optical properties of the cassettes. These two cassettes were prepared from the corresponding donor and acceptor fragments via Palladium catalyzed Sonogashira reactions. An aryl-alkyne linker was formed to join the donor and acceptor parts. The synthesis for cassette **32** was demonstrated in Scheme 6.8. Cassette **33** was made similarly from donor **35** and acceptor **37c** in 45 % yield.

Scheme 6.8. Synthesis of cassette **32**.



NaO_3S
O
N
B
F_2
N
N
I
1g
+
$PdCl_2(PPh_3)_2$
CuI, Et_3N
DMF, 25 ^oC, 12 h
O
CO_2H
MeO
OMe
37a
32 79 %

Table 6.1. summarizes the yields for the syntheses of cassettes **31**, corresponding BODIPY acceptors **37** and the intermediates **47-48**.

Table 6.1. Summary of the yields for preparation of **47**, **48**, **37** and cassettes **31**

yield (%)	**a**	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**
47	69	15	64	50	88	80	47	75	41
48	92	86	100	83	88	87	80	100	94
37	85	76	67	95	80	100	80	N/A	N/A
31	75	78	80	48	60	35	25	N/A	N/A

6.3 Spectroscopic properties of BODIPY acceptors and cassettes

The absorption and fluorescence emission properties of BODIPY acceptors are summarized in Table 6.2. BODIPY **37a** and its water-soluble derivative **37b** have similar absorption and emission maxima. They also have almost the same quantum yield. The absorption and emission wavelengths for compound **37d** are almost the same as its regioisomer **37a**. However, BODIPY **37d** has much higher quantum yield than **37a** (0.81 and 0.48 in MeOH, respectively). The oxygen atom at *para*-position of the aryl substituent shifts the emission maxima *ca* 25 nm to the red (compare **37c** and **37b**, **37f** and **37e**). Molecular constraints in compound **37e** extend the conjugation and cause red shift of the absorption and emission wavelengths compared with uncyclized derivatives. BODIPYs **37b**, **37c** and **37g** are water-soluble. Their properties in aqueous buffer are also summarized in Table 6.2. The absorption and emission maxima of these compounds in buffer are very similar to that in MeOH. However, the quantum yield and extinction coefficient for these compounds are slightly lower than that in MeOH. Illustrative absorbance and fluorescence spectra of the BODIPY acceptors and the donor fragment **34** are showed in Figure 6.8.

Table 6.2. Photophysical properties of BODIPY acceptors

	$\lambda_{\mathbf{max\ abs}}$ **(nm)**	log ($\varepsilon_{\mathbf{max}}$)	$\lambda_{\mathbf{max\ emiss}}$ **(nm)**	**fwhm (nm)**	**Φ**
		MeOH			
37a	543	4.61	598	49	0.48 ± 0.01[a]
37b	553	4.71	605	47	0.50 ± 0.01[a]
37d	542	4.70	599	47	0.81 ± 0.02[a]
37c	585	4.71	627	44	0.76 ± 0.02[a]
37e	626	5.12	642	32	0.75 ± 0.02[b]
37f	650	4.98	669	37	0.57 ± 0.01[b]
37g	653	5.10	672	36	0.69 ± 0.02[b]
		0.1 M PBS buffer (pH 7.4)			
37b	544	4.65	604	56	0.44 ± 0.02[a]
37c	582	4.52	629	50	0.66 ± 0.01[a]
37g	660	4.84	678	40	0.48 ± 0.02[b]

[a] Cresyl Violet (Φ=0.54 in MeOH); [b] Nile Blue (Φ = 0.27 in EtOH) were used as standards.

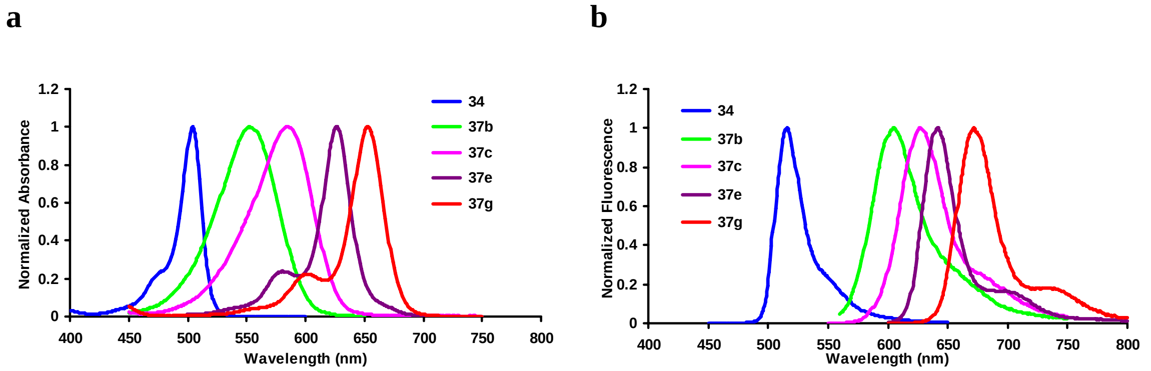


Figure 6.8. Normalized (**a**) absorbance; and (**b**) fluorescence spectra of BODIPY donor **34** and illustrative acceptors **37** in MeOH (10^{-6} M for absorbance; 10^{-7} M for fluorescence).

The absorption spectra of all the cassettes display two distinct peaks. One stong absorption band around 500 nm corresponds to the λ_{max} of the donor component. The other peak is due to the absorption of the acceptor fragment. The absorption spectra of the cassettes resemble those that would be obtained by superimposing the donor and acceptor spectra, indicative of two independent dye fragments. However, when irradiated at the absorption maxima of the donor, the donor fluorescence is almost completely quenched and only acceptor fluorescence is observed in MeOH. Thus energy is absorbed at wavelengths corresponding to the donor part and emitted at wavelengths governed by the acceptor fragment, hence gives large pseudo-Stokes' shift. Complete energy transfer from donor to acceptor is observed for all the cassettes when MeOH is used as solvent. However, some energy leak from the donor fragment is detected for cassettes **31f** and **31g** when they are dissolved in aqueous media. Illustrative absorption and emission spectra of these cassettes are shown in Figure 6.9.

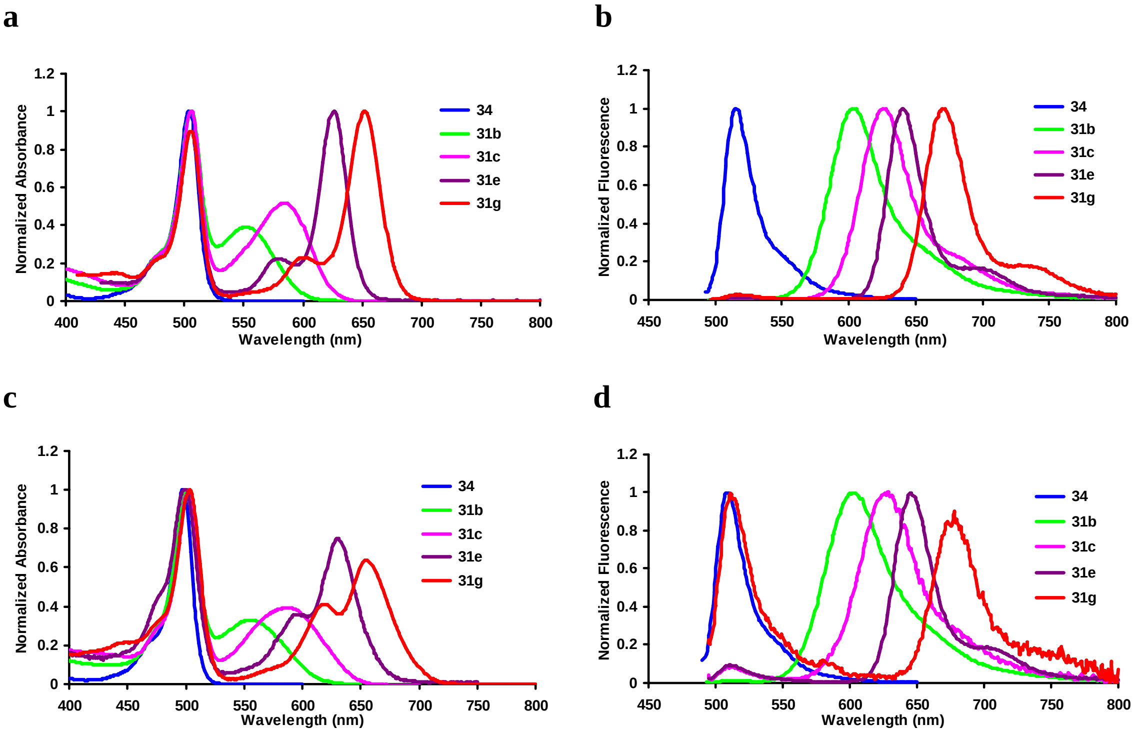


Figure 6.9. Normalized (**a**) absorbance, (**b**) fluorescence in MeOH; (**c**) absorbance, (**d**) fluorescence in 0.1 M phosphate buffer (pH 7.4) of BODIPY donor **34** and cassettes **31**. Throughout, 10^{-6} M for absorbance, 10^{-7} M for fluorescence, ex. at 488nm.

Salient photophysical properties of the BODIPY-BODIPY cassettes are shown in Table 6.3. Through-bond energy transfer cassettes are usually designed to absorb light at the donor excitation wavelength, relay it to the acceptor part, then emit fluorescence from there. The term "energy transfer efficiency" (ETE %) quantifies this and it is defined as follows:

$$\text{ETE \%} = \frac{\text{quantum yield of the acceptor fragment in the cassette excited at the donor}}{\text{quantum yield of the acceptor fragment in the cassette excited at the acceptor}} \times 100$$

ETE % is a measure of the energy transfer of the cassette when irradiated at the donor. It reflects the extent of energy transfer which is emitted as photons by the acceptors. This term eliminates the negative effects of non-radiative loss in the transfer process.

The fluorescence quantum yields of the acceptor fragments in the cassettes were determined when excited at the donor and acceptor parts. Values of the ETE % for the cassettes were calculated using these quantum yields. In MeOH, the quantum yields were almost the same whether the donor or the acceptor parts were excited, indicating nearly complete energy transfer from the donor to the acceptor. The ETE % values for all the cassettes are summarized in Table 6.3. The results in Table 6.2. and Table 6.3. also show that the quantum yields of the acceptor fragments were almost the same in the presence and absence of the donor parts. The absorption and emission maxima of the acceptor fragments in the cassettes were also very similar to that of the BODIPY acceptors alone. However, in aqueous buffer, the quantum yields were significantly reduced. Cassettes **31c**, **31d**, **31f**, **31g** were almost non-fluorescent in aqueous media. The ETE % for all the cassettttes were also lower than that in MeOH. Several cassettes have obvious energy leak from the donor fragment in aqueous media (Table 6.3.). The absorption and emission wavelengths were similar whether MeOH or aqueous buffer was used as solvent. Another observation was that the extinction coefficients were also less in aqueous media. Cassettes **32** and **33** also have complete energy transfer and high quantum yields in MeOH (see Appendix Information). However, they were not soluble in aqueous media, thus their photophysical properties were not further studied. Interestingly, cassettes **31a**, **31b** and **31e** still have moderate quantum yields (0.24-0.30) and almost complete energy transfer in aqueous media. Cassettes **31a** and **31b** emit in the same range (*ca* 600 nm) while cassette **31e** fluoresce at longer wavelength (*ca* 650 nm). Thus cassettes **31b** and **31e** were chosen to monitor protein-protein interactions (see next section).

Table 6.3. Photophysical properties of BODIPY cassettes

	$\lambda_{\mathbf{abs\,D}}$ (nm)	$\lambda_{\mathbf{abs\,A}}$ (nm)	$\lambda_{\mathbf{emiss\,A}}$ (nm)	log($\varepsilon_{\mathbf{max}}$)		Φ of acceptor		**E.T.E** (%)[d]
				D	A	Ex. D[a]	Ex. A	
				MeOH				
31a	506	542	595	4.96	4.52	0.50	0.50[a]	100
31b	506	552	601	5.15	4.74	0.46	0.47[b]	98
31c	506	584	623	5.02	4.73	0.69	0.73[b]	95
31d	506	542	594	5.02	4.58	0.77	0.79[a]	97
31e	506	626	640	4.67	4.72	0.69	0.71[c]	97
31f	506	649	666	4.81	4.86	0.47	0.51[c]	92
31g	506	652	668	4.90	4.95	0.55	0.58[c]	95
				0.1M PBS buffer, pH 7.4				
31a	500	535	594	4.94	4.41	0.25	0.27[a]	93
31b	501	556	599	5.06	4.58	0.30	0.32[b]	94
31c	503	589	628	4.94	4.53	0.030	0.035[b]	86
31d	500	535	592	4.95	4.42	0.069	0.077[a]	90
31e	499	631	645	4.52	4.39	0.24	0.26[c]	92
31f	500	655	674	4.65	4.48	0.032	0.037[c]	86
31g	503	654	673	4.82	4.63	0.008	0.012[c]	67

[a] Rhodamine 6G (Φ=0.94 in EtOH) ; [b] Cresyl Violet (Φ=0.54 in MeOH); [c] Nile Blue (Φ = 0.27 in EtOH) were used as standards; [d] energy transfer efficiency (E.T.E) was calculated by dividing the quantum yield of acceptor when excited at donor by the quantum yield of acceptor when excited at acceptor.

6.4 Protein-protein interaction studies

FRET donor Atto425 and through-bond energy transfer cassettes **1** (cassette **31b**) and **2** (cassette **31e**) were used to study protein-protein interactions and test our hypothesis that combination of FRET and TBET will enable detection of multiple protein-protein interacctions simutaneously (See Chapter I, Figure 1.11.). The structures and spectral properties of Atto425, cassettes **1** and **2** are shown in Figure 6.10. Several key spectral data make Atto425 a good FRET donor for cassettes **1** and **2**. First, the fluorescence emission of Atto425 overlaps perfectly with absorbance of the donor part in cassettes **1** and **2**. Second, cassettes **1** and **2** do not absorb at Atto425's absorption maxima. Thus cassettes **1** and **2** will not be directly excited during the FRET studies.

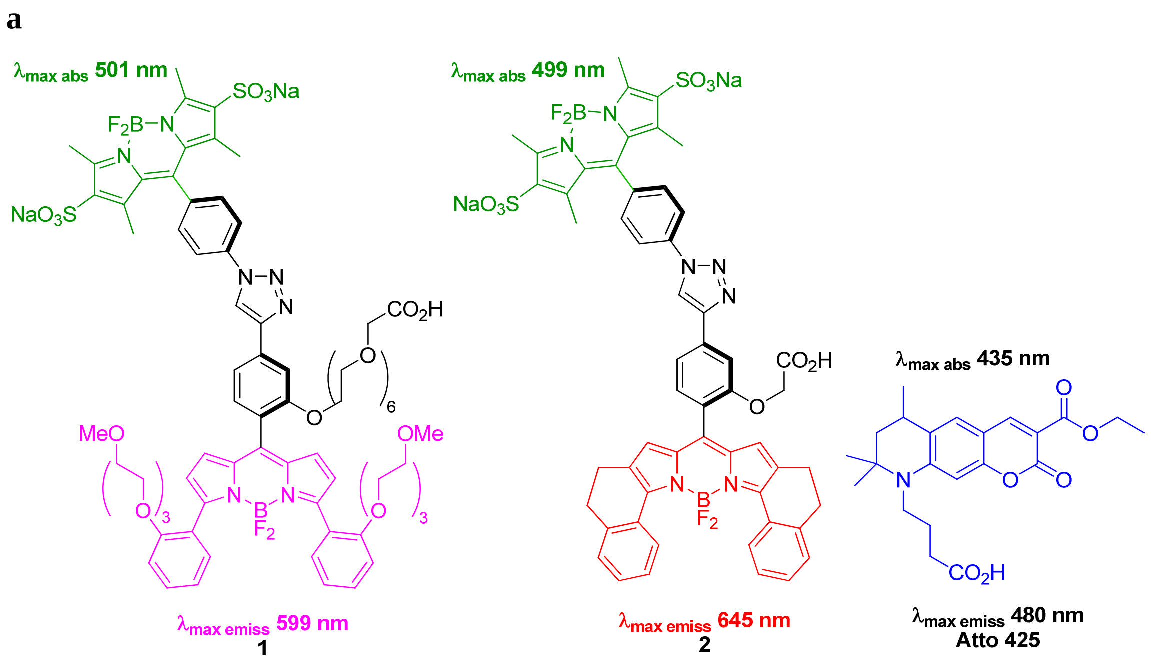

a
λmax abs 501 nm
λmax emiss 599 nm
1
λmax abs 499 nm
λmax emiss 645 nm
2
λmax abs 435 nm
λmax emiss 480 nm
Atto 425

Figure 6.10. FRET donor **Atto425**, cassette **1** and cassette **2**: (**a**) structures; (**b**) absorbance and fluorescence spectra in 0.1 M phosphate buffer (pH 7.4).

b

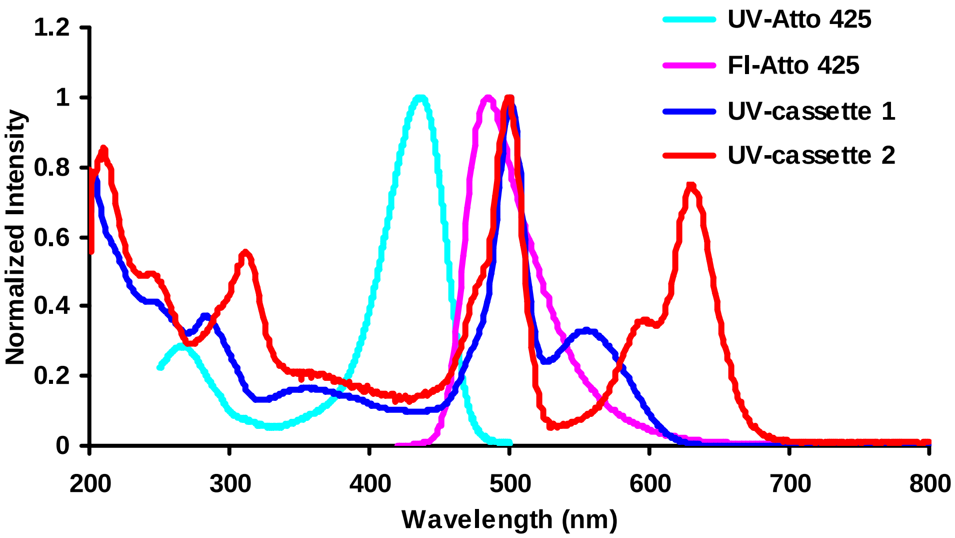


Figure 6.10. Continued.

A well-behaved system for studying protein-protein interactions was needed to test cassettes **1** and **2**. This was generated via the interaction of biotin-functionalized bovine serum albumin (BSA) with avidin or streptavidin.[233-235] The vitually irreversible nature of the interaction was ideal for this proof of concept study. Thus, BSA was reacted with *N*-hydroxysuccinimide (NHS) activated forms of: (i) the FRET donor Atto-425 (a coumarin derivative; Sigma); then, (ii) biotin. The dye:biotin:BSA ratio was measured[101,236] as 5.2:6.6:1.0. Cassette **1** was activated (di-*iso*propylcarbodiimide, NHS) then coupled to streptavidin (dye:protein 2.5:1.0). Similarly, cassette **2** was coupled to avidin (dye:protein 1.6:1.0). The strategy followed was to mix these proteins, excite the Atto425 donor, and observe emissions from cassettes **1** and **2**. Strong enhancements mean that binary complexes like **A** and **B**, three-component tertiary complexes like **C**, and/or higher order assemblies were formed (Figure 6.11.). This experiment would *not* be possible using Atto425 as a FRET donor and single dyes based on the acceptor components of **1** and **2** because the fluorescence emission from Atto425 (centered on 480 nm) has minimal overlap with the absorbance spectrum of the TBET-acceptor from **1** and almost none with that in **2** ($\lambda_{\text{max abs}}$ 556 and 631 nm, respectively, in pH 7.4 phosphate buffer).

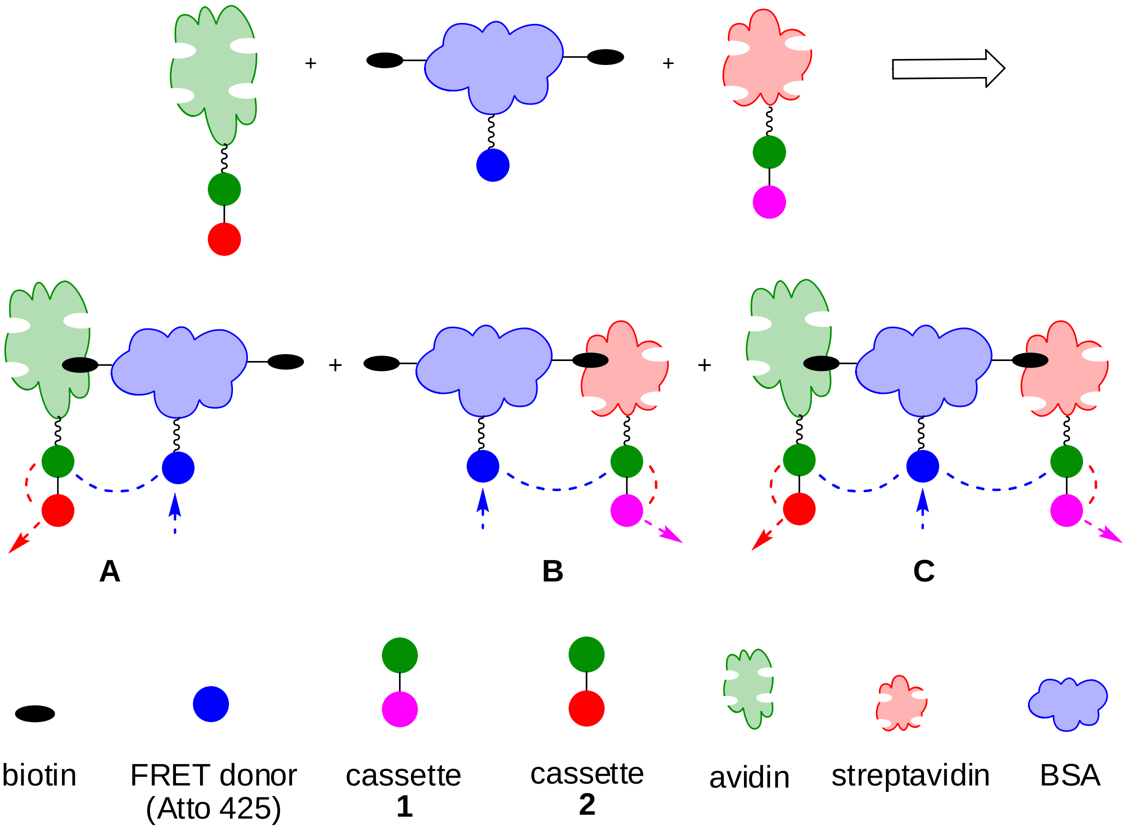


Figure 6.11. FRET from one protein can be used to excite the donors of TBET-cassettes on another protein to reveal interactions.

Interactions between two proteins were studied in a fluorescence cuvette to test our hypothesis and the artificial protein interaction system. Several precautions were necessary before the key energy transfer experiment could be performed and interpreted. First, **1**-streptavidin and **2**-avidin were pretreated with BSA to suppress non-specific interactions. Second, steps were taken to correct for variations of dye fluorescence properties in different protein environments. This was done by observing each probe for each relevant protein-protein interaction, one dye at a time to give the baseline "before FRET" values. Cassettes **1** and **2** were slightly influenced by the protein environment, but Atto425 showed more marked changes. The "after FRET" values were direct fluorometer reading for the protein mixtures. The difference between "before FRET" and "after FRET" values were the changes due to FRET during protein interactions.

The fluorescence profiles for the two protein interactions were summarized in Figure 6.12. Experiment (i) tested the interaction between Atto425-BSA-biotin and **1**-streptavidin. Mixing of these two proteins gave significant fluorescence enhancement at around 600 nm corresponding to cassette-**1** and the fluorescence of Atto425 (*ca* 480 nm) decreased accordingly. The biotin/streptavidin interactions brought the proteins together such that the FRET donor was able to relay energy to the TBET cassette donor parts. Thus, the fluorescence output of this experiment clearly showed that Atto425-BSA-biotin and 1-steptavidin were interacting. Experiment (ii) represented a control study for experiment (i). In this experiment, Atto425-BSA and **1**-streptavidin were mixed in a fluorescence cuvette. Since the BSA was not functionalized with biotin hence no protein-protein interactions were expected. This was confirmed by the fluorescence output. No obvious changes between "before FRET" and "after FRET" were observed in this experiment. The interactions between Atto425-BSA-(biotin) and **2**-avidin were studied similarly as shown in Figure 6.12. **c** and **d** (experiment iii and iv, respectively). The results showed that the interactions between these two proteins were also successfully detected using our through-bond energy transfer cassettes and a FRET donor Atto425. Our method was further validated by testing the interactions between Atto425-BSA-(biotin) with **1**-BSA (experiment v and vi, Figure 6.12.**e** and **f**). Whether Atto425-BSA was functionalized with biotin or not, no interactions were expected between these two proteins thus no significant changes in the fluorescence intensities should be detected. This was confirmed by the experimental data as shown in Figure 6.12.**e** and **f**. All these studies proved our hypothesis and showed that combination of FRET and TBET can be used to monitor protein-protein interactions.

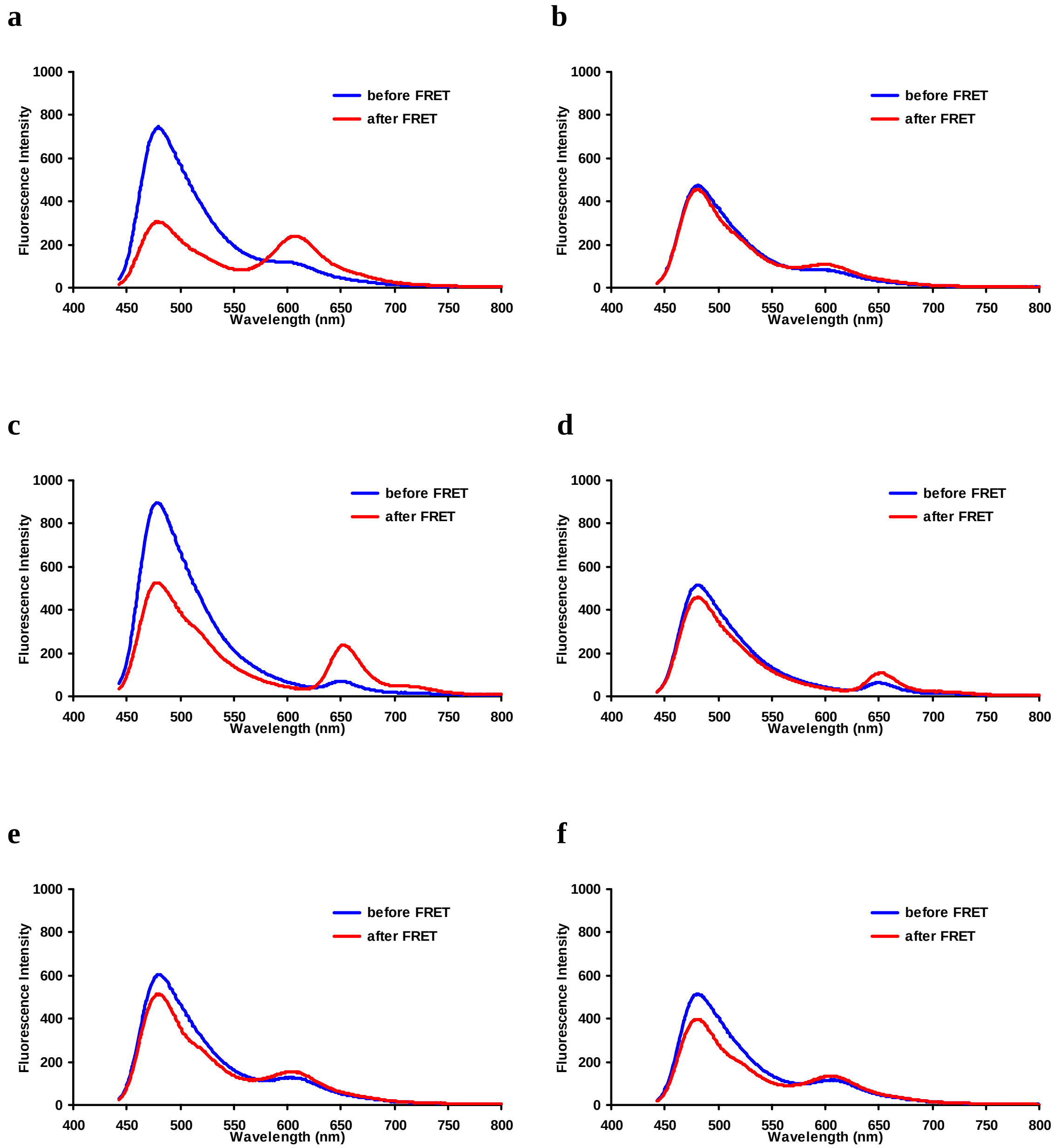


Figure 6.12. *In vitro* two protein interaction studies. (i) Atto425-BSA-biotin and **1**-streptavidin; (ii) Atto425-BSA and **1**-streptavidin; (iii) Atto425-BSA-biotin and **2**-avidin; (iv) Atto425-BSA and **2**-avidin; (v) Atto425-BSA-biotin and **1**-BSA; and, (vi) Atto425-BSA and **1**-BSA. **a** spectra for experiment (i); **b** for experiment (ii); **c** for experiment (iii); **d** for experiment (iv); **e** for experiment (v); **f** for experiment (vi). Throughout, the mixtures were excited at 430 nm. The concentrations of the proteins used were (x 10^{-7} M, in pH 7.4 phosphate buffer): Atto-425-BSA, 2; Atto-425-BSA-biotin, 2; **1**-streptavidin, 3; **2**-avidin, 9; **1**-BSA, 3.

The next step is to test our method by studying more complex protein-protein interactions such as interactions between three proteins. Figure 6.13. describes four key experiments of observing protein-protein interactions between three proteins. Experiment (i) represents a critical control wherein Atto-425-BSA and **1**-streptavidin and **2**-avidin were mixed. In this experiment the BSA was not functionalized with biotin hence no protein-protein interactions were expected. The slight changes in fluorescence intensities were due to errors from quantification calibration and/or residual non-specific interactions (even though all samples were pretreated with BSA). Experiment (ii) was the most important one. Here the combination of Atto-425-BSA-biotin, **1**-streptavidin and **2**-avidin gave significant fluorescence enhancements for cassettes **1** and **2** relative to experiment (i), and the fluorescence of the Atto-425 FRET donor decreased accordingly. These observations indicate the biotin-to-avidin/streptavidin interactions brought the proteins together such that the FRET donor was able to relay energy to the TBET cassette donor parts. Exactly the same combination was used in experiment (iii) except that **1**-streptavidin and **2**-avidin were pretreated with biotin. This was a negative control in which the BSA/avidin (or streptavidin) interaction was blocked by pretreating with biotin. This almost completely suppressed the fluorescence enhancements that were observed in experiment (ii). In experiment (iv), cassette **1**, bound this time to BSA, did not show significant energy transfer enhancement, while **2**-avidin did, proving that the origin was the biotin/avidin (or streptavidin) interaction.

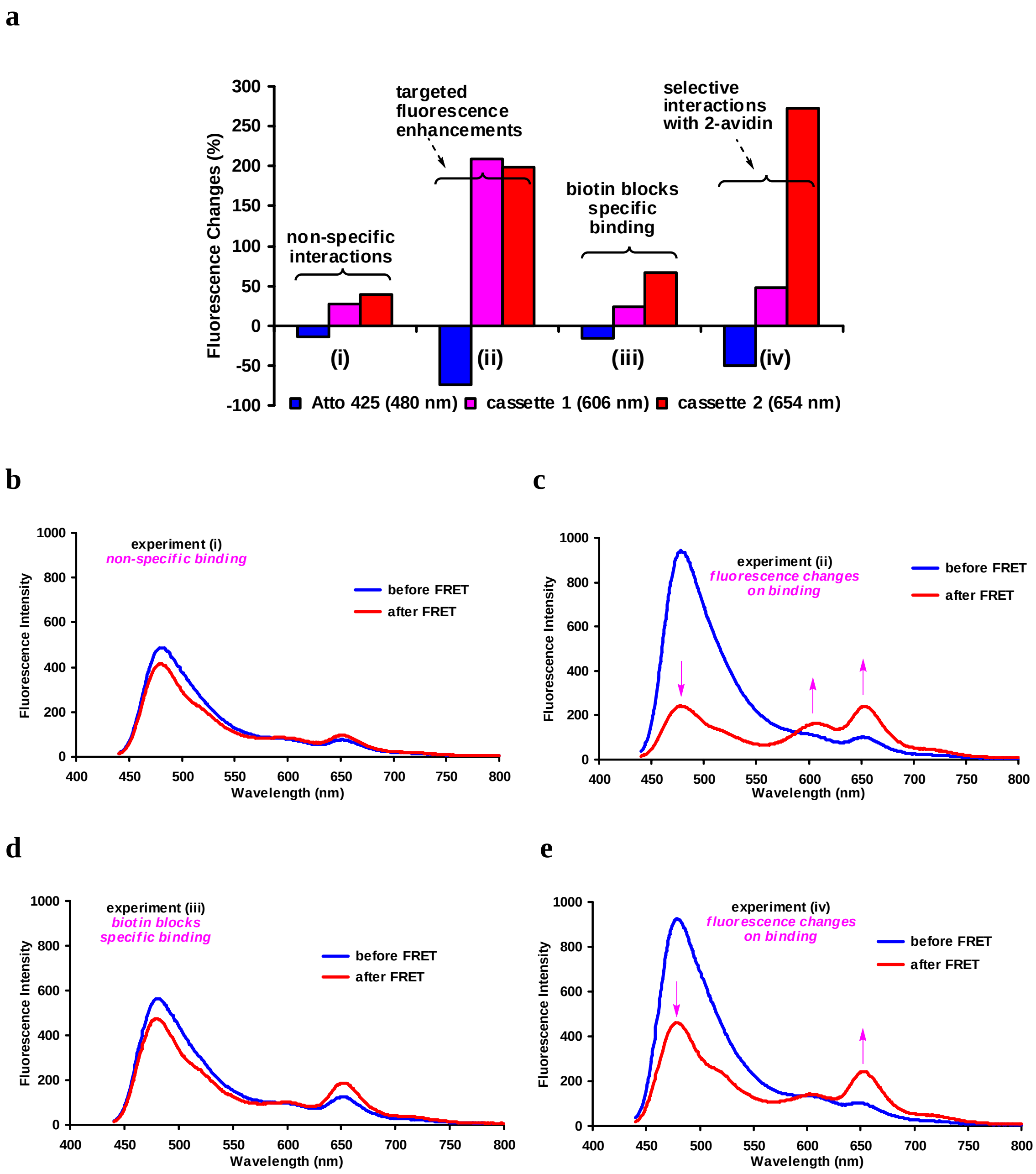


Figure 6.13. *In vitro* three protein interactions. **a** Deconvoluted fluorescence changes for combination of (i) Atto-425-BSA, **1**-streptavidin and **2**-avidin; (ii) Atto-425-BSA-biotin, **1**-streptavidin and **2**-avidin; (iii) **1**-streptavidin and **2**-avidin were pretreated with excess biotin then Atto-425-BSA-biotin was added; and, (iv) **1**-BSA, **2**-avidin and Atto-425-BSA-biotin. **b** spectra for experiment (i); **c** for experiment (ii); **d** for experiment (iii); and, **e** for experiment (iv), (before deconvolution). Throughout, the mixtures were excited at 430 nm. The concentrations of the proteins used were (x 10^{-7} M, in pH 7.4 phosphate buffer): Atto-425-BSA, 2; Atto-425-BSA-biotin, 2; **1**-streptavidin, 3; **2**-avidin, 9; **1**-BSA, 3.

The protein-protein interactions were then investigated in live COS-7 cells using Pep-1[52,53,57,237] to import the labeled proteins inside the cells at 37 °C (see Appendix for detailed information). No significant fluorescence upon excitation at 458 nm, where FRET donor Atto425 absorbs, was observed in the acceptors channels (565-615 nm and 640-704 nm), when Atto425-BSA-biotin, **1**-streptavidin and **2**-avidin were imported into cells individually (Figure 6.14.**a-c**). However, significant signal from the acceptor proteins (**1**-streptavidin or **2**-avidin) was observed upon excitation at 458 nm when the donor protein (Atto425-BSA-biotin) and the acceptor protein were present in the same cell (Figure 6.14.**d-e**). When all three components were present, interactions between the FRET donor and the two cassettes labeled proteins were evident (Figure 6.14.**f**).

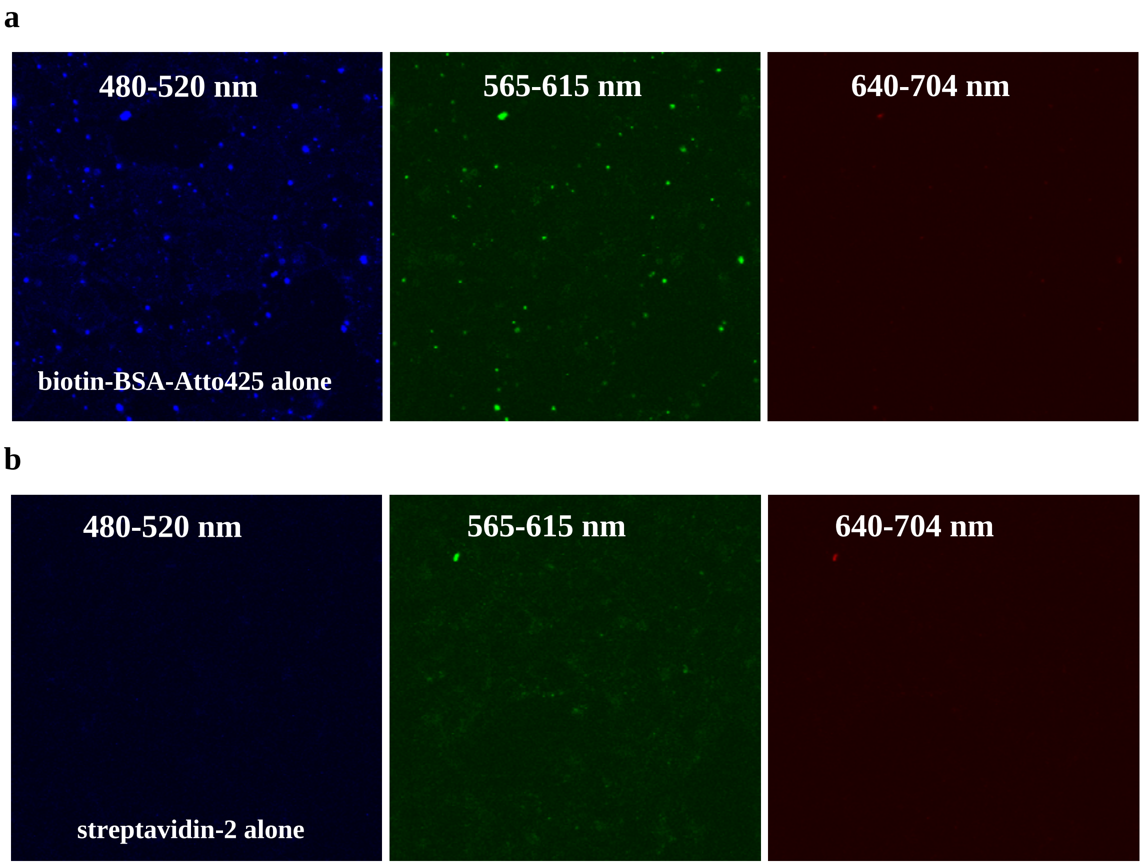


Figure 6.14. Studies of protein interactions in COS-7 cells. (**a**) biotin-BSA-Atto425; (**b**) streptavidin-cassette**1**; (**c**) avidin-cassette**2**; (**d**) biotin-BSA-Atto425 + streptavidin-cassette**1**; (**e**) biotin-BSA-Atto425 + avidin-cassette**2**; (**f**) biotin-BSA-Atto425 + streptavidin-cassette**1** + avidin-cassette**2**. Throughout, excited at 458 nm.

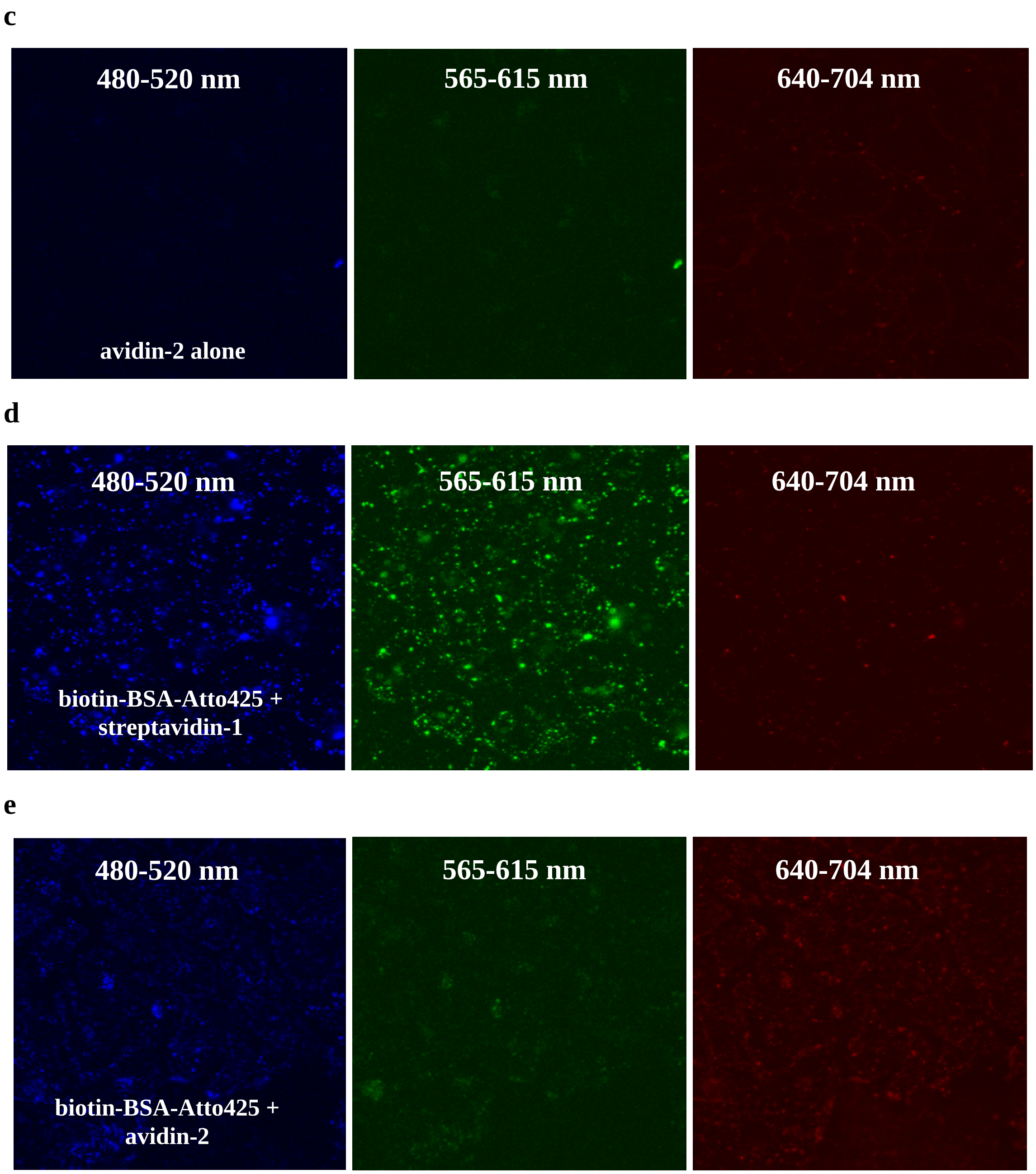

c
480-520 nm
565-615 nm
640-704 nm
avidin-2 alone
d
480-520 nm
565-615 nm
640-704 nm
biotin-BSA-Atto425 +
streptavidin-1
e
480-520 nm
565-615 nm
640-704 nm
biotin-BSA-Atto425 +
avidin-2

Figure 6.14. Continued.

f

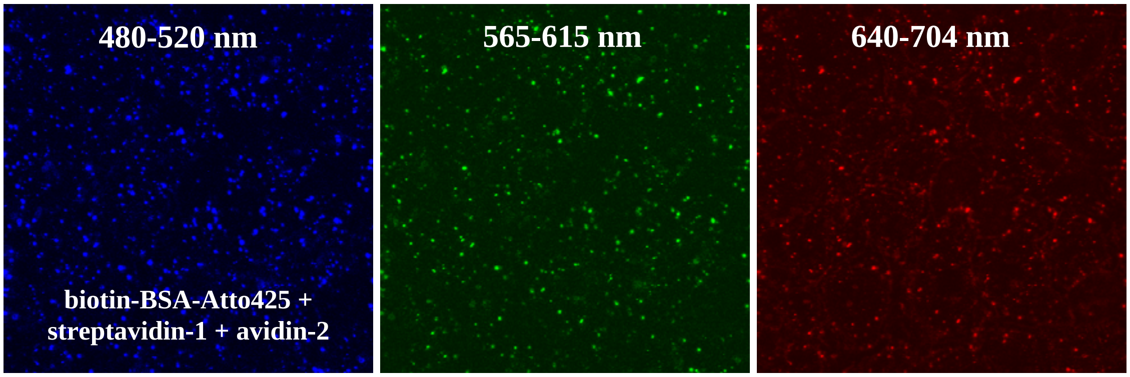


Figure 6.14. Continued.

6.5 Conclusions

Several through bond energy transfer cassettes **31-33** based on BODIPY fragments were prepared. In organic solvents such as MeOH, these cassettes transferred energy efficiently from the donor to the acceptors when irradiated in the donor absorption region. The overall quantum yields for the cassettes were almost the same as the quantum yield of the acceptor fragments alone. All of the cassettes except **32** and **33** were also soluble in water. Unfortunately, the quantum yields for the cassettes reduced dramatically in water. Several cassettes (**31c**, **31d**, **31f**, **31g**) were almost non-fluorescent in aqueous buffer. The energy transfer efficiencies were also lower than that in MeOH. No systematic correlation between the behavior in water and cassette structures were observed.

Interestingly, three cassettes **31a**, **31b** and **31e** still had very good photophysical properties in aqueous buffer. Two of them (**31b** and **31e**) were used to monitor protein-protein interactions. An artificial protein interaction system was built utilizing biotin/(strept)avidin interactions to test the water-soluble cassettes. The protein interactions with the model system were successfully detected *in vitro* and in living cells using the combination of a FRET donor and two cassettes.

CHAPTER VII

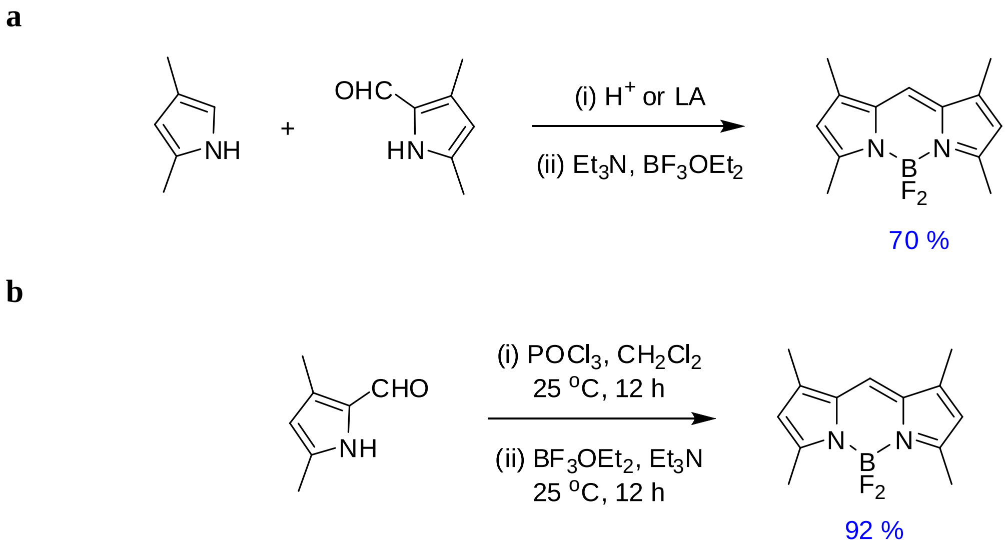
CONCLUSIONS AND OUTLOOK

7.1 Conclusions

7.1.1 New methodologies for the syntheses of BODIPY and rosamine

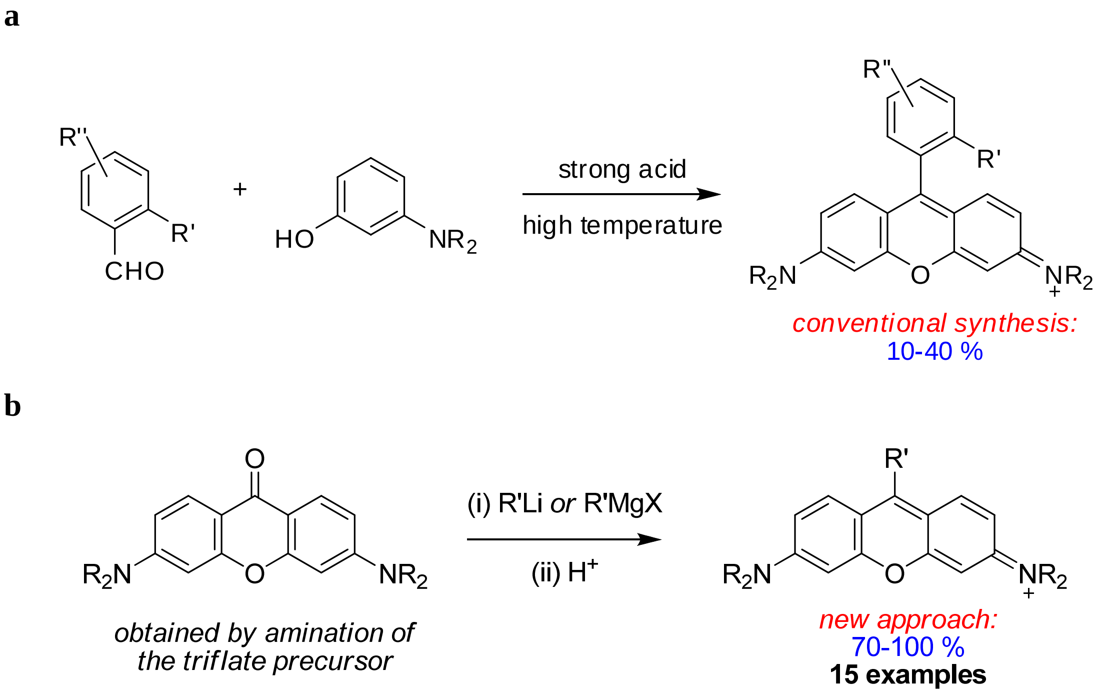
A new synthesis of symmetric BODIPY dyes without *meso*-substituents was discovered (Scheme 7.1.). The kinetics and mechanisms for the new reaction were systematically investigated. Possible mechanisms were proposed for the new transformation. Series of BODIPYs were easily prepared using the new approach. Some of the substrates gave superior yields of desired product compared to conventional synthesis.

Scheme 7.1. Illustrative syntheses of symmetric BODIPYs without *meso*-substituents. (**a**) Conventional synthesis; (**b**) a new approach.



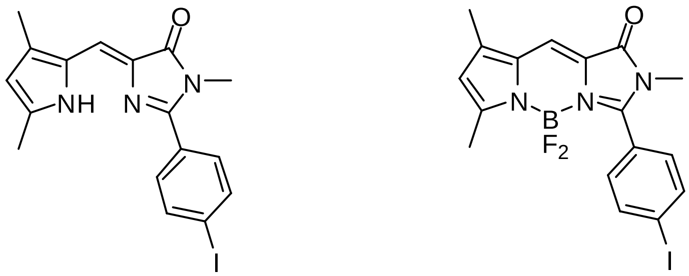
New methods were developed for the efficient syntheses of rosamine derivatives (Scheme 7.2.). Compared with traditional synthesis, the new approach described here uses milder conditions but leads to much higher yields with easy separation of the desired product. Water-soluble derivatives were also prepared by functionalization of the initial products. The fluorescence properties of these new rosamines showed interesting dependence on the amine substituents and the solvent media. Several new rosamines were found to be highly cytotoxic toward cancer cells. Some of these new rosamines exhibited at least 10-fold lower IC_{50} values compared to Rh123 which is a well studied anti-tumor compound. Further studies revealed that the new compounds showed greater cytotoxicity towards oral and nasopharyngeal cancer cells compared to immortalized normal cells of the same organ type cancer cells which offers potential for the design of new cytotoxic drugs for selective tumor cell killing.

Scheme 7.2. Syntheses of rosamine derivatives. (**a**) Conventional synthesis; (**b**) a new approach.


a
R''
R'
CHO
+
HO
NR2
strong acid
high temperature
R''
R'
R2N
O
NR2
+
conventional synthesis:
10-40 %
b
O
R2N
O
NR2
obtained by amination of
the triflate precursor
(i) R'Li or R'MgX
(ii) H+
R'
R2N
O
NR2
+
new approach:
70-100 %
15 examples

7.1.2 Syntheses of novel fluorescent compounds

Highly fluorescent GFP-chromophore analogs have been successfully designed and prepared (Figure 7.1.). The correlation between optical properties and structures of these new compounds was investigated. This work provided the first experimental proof that the rigidity of the chromophore inside the protein was responsible for its strong fluorescence. Moreover, functional groups were incorporated, which not only provided the fluorophores with water-solubility but also were amenable to anchor them onto biomolecules.



4c	**1c**
almost non-fluorescent	**highly fluorescent**
458 515 40200 0.0005	**492 532 57300 0.86**

absorption (nm), emission (nm), extinction coefficient ($M^{-1}cm^{-1}$), quantum yield
in MeOH

Figure 7.1. Highly fluorescent GFP-chromophore analog **1c** and its precursor **4c**.

Novel pyronin derivatives with meso-heteroatom substituents (Figure 7.2.) have been efficiently prepared. This represents the first synthesis of such dyes. The photophysical properties of these compounds are highly dependent on the nature of the heteroatom that is attached to the central carbon atom. Series of dyes with different emission colors can be easily accessed simply by varying the nucleophiles.

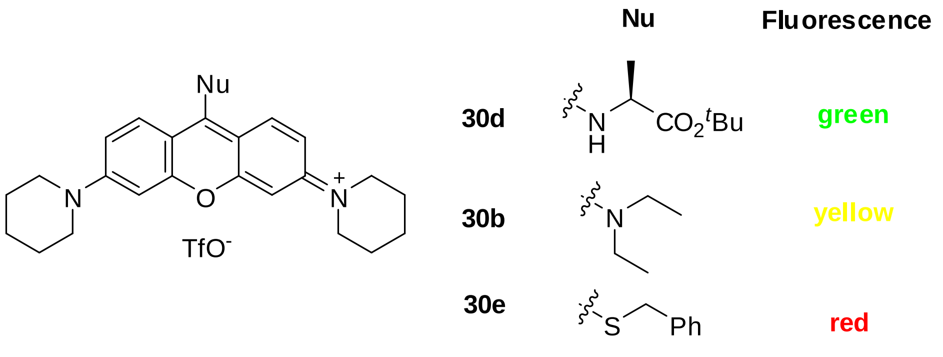
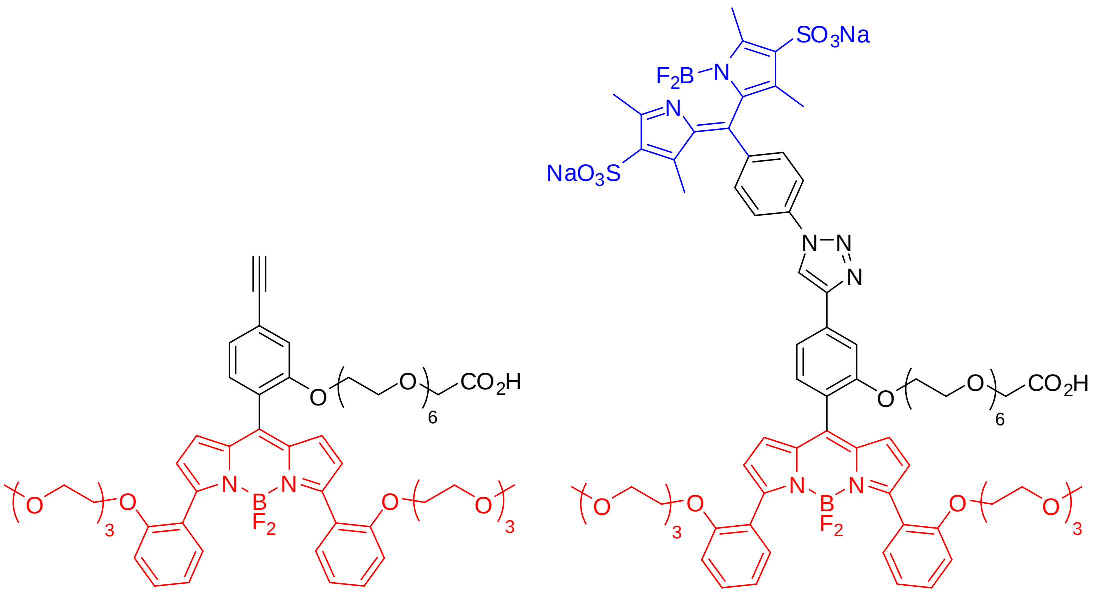


Figure 7.2. Novel pyronin derivatives.

Series of water-soluble BODIPY dyes with emission above 600 nm were prepared. A carboxylic acid was also incorporated into these BODIPYs so that they can be used for biomolecular labeling and intracellular imaging. Moreover, these new BODIPYs have also been used as acceptors to prepare various water-soluble through bond energy transfer cassettes (Figure 7.3.). The cassettes all have complete energy transfer and high quantum yields in polar organic solvents (*eg* MeOH). Several cassettes still have good energy transfer and quantum yields in aqueous media and even on proteins. Applications of such cassettes in biological studies were also demonstrated.



37b

Φ 0.50; MeOH
$\lambda_{\text{max abs}}$: 553 nm
$\lambda_{\text{max fl}}$: 605 nm

Φ 0.44; pH 7.4 buffer
$\lambda_{\text{max abs}}$: 544 nm
$\lambda_{\text{max fl}}$: 604 nm

31b

Φ 0.46; MeOH
$\lambda_{\text{max abs}}$: 506, 552 nm
$\lambda_{\text{max fl}}$: 601 nm

Φ 0.30; pH 7.4 buffer
$\lambda_{\text{max abs}}$: 501, 556 nm
$\lambda_{\text{max fl}}$: 599 nm

Figure 7.3. Illustrative examples of water-soluble BODIPY acceptor and through-bond energy transfer cassette.

7.1.3 Monitoring of protein-protein interactions *in vitro* and in living cells

One of the major applications of fluorescent dyes is to label biomolecules and then study their intracellular localization and interactions with surrounding molecules. For instance, the water-soluble through-bond energy transfer cassette **31b** was used to label several proteins such as BSA, avidin, streptavidin and β-galactosidase. The quantum yield of the cassette was reduced upon attached to proteins. The absorption and emission wavelength of these conjugates were very similar to these of the cassette (Table 7.1.). The fluorescent conjugates were then introduced back into cells such as COS-7 cells by cell penetrating peptide (*eg* pep-1, Arg_8). These conjugates were easily detected by fluorescence microscope. The intracellular localization was also studied.

Table 7.1. Optical properties of cassette and its protein conjugates

	Dye/ protein ratio	λ_{abs} D (nm)	λ_{abs} A (nm)	λ_{emiss} A (nm)	Φ of acceptor ex. D[a]	Φ of acceptor ex. A[b]	**E.T.E** (%)[c]
Cassette **31b**	n/a	501	556	599	0.30	0.32	94
BSA-cassette	4.0	505	558	604	0.094	0.102	92
β-Galactosidase-cassette	4.2	505	556	601	0.11	0.12	92
Avidin-cassette	1.9	503	554	605	0.09	0.10	90
Streptavidin-cassette	2.5	502	558	604	0.075	0.080	94

[a] Rhodamine 6G (Φ=0.94 in EtOH) ; [b] Cresyl Violet (Φ=0.54 in MeOH) were used as standards; [c] energy transfer efficiency (E.T.E) was calculated by dividing the quantum yield of acceptor when excited at donor by the quantum yield of acceptor when excited at acceptor.

The water-soluble through-bond energy transfer cassettes were also used to detect multiple protein-protein interactions simultaneously (Figure 7.4.). As a proof of concept study, an artificial three protein interacting system was built by utilizing the biotin/(strept)avidin interactions. Thus, FRET donor (Atto425) labeled BSA was functionalized with biotin. Avidin and streptavidin were labeled with two different cassettes. The interactions between the Atto425-BSA-biotin and (strept)avidin were successfully detected *in vitro* and in COS-7 cells.

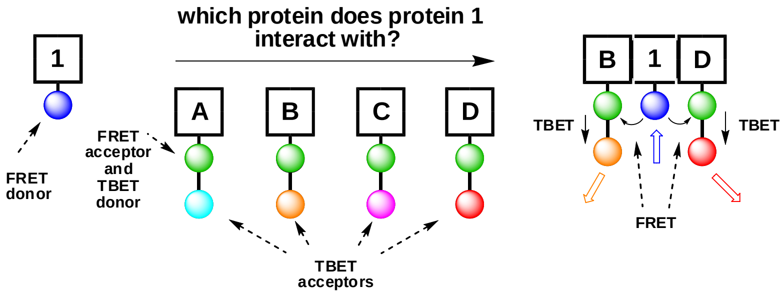

which protein does protein 1 interact with?
1
A
B
C
D
FRET donor
FRET acceptor and TBET donor
TBET acceptors
B
1
D
TBET
TBET
FRET

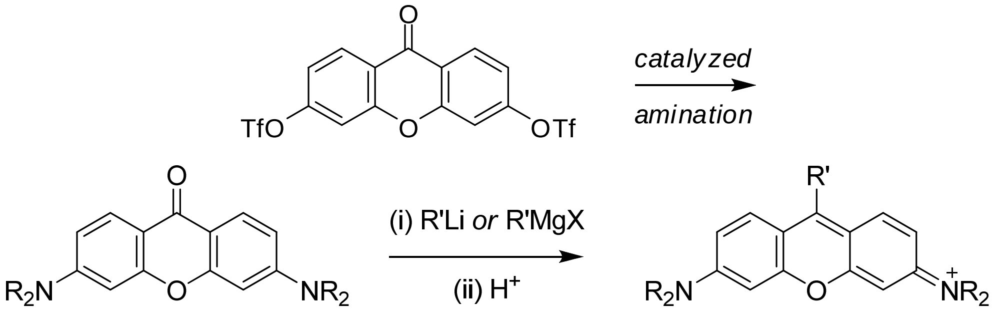
Figure 7.4. Scheme for detecting protein-protein interactions.

7.2 Outlook

7.2.1 Rosamine derivatives

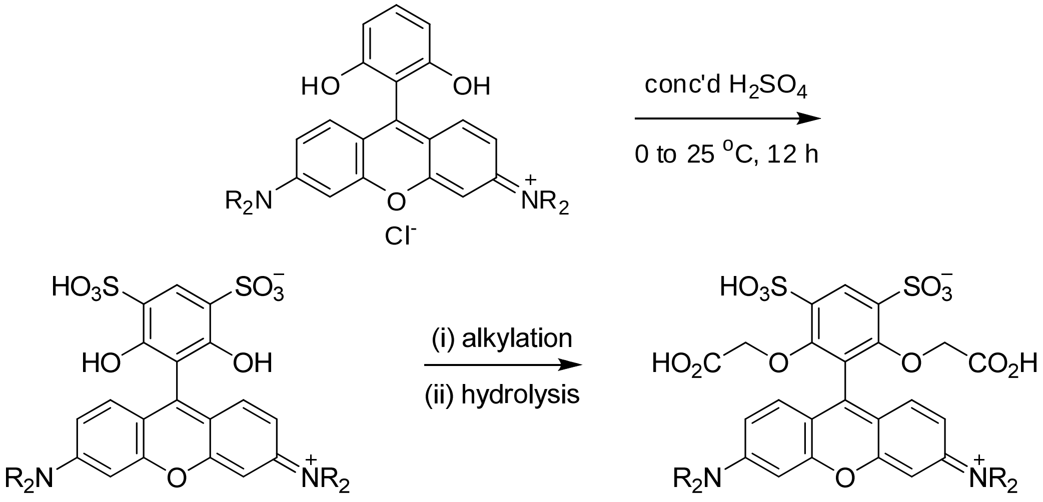
There is still a lot work to be done in this field. More rosamine derivatives with different amine substituents are needed to figure out the origin of the dependence of the optical properties on the amine substituents. This can be achieved be employing transition metal catalyzed amination, such as Buchwald-Hartwig reaction, of the 3,6-di-OTf-xanthone (Scheme 7.3.).

Scheme 7.3. New synthesis of rosamine derivatives.


O
TfO
O
OTf
catalyzed
amination
O
R2N
O
NR2
(i) R'Li or R'MgX
(ii) H+
R'
R2N
O
+
NR2

The sulfonation reaction as showed in Chapter IV (reaction 4.3) can be further exploited to prepare useful water-soluble rosamines. One illustrative example is demonstrated in scheme 7.4.

Scheme 7.4. Synthesis of water-soluble rosamines.


HO
OH
R2N
O
+
NR2
Cl-
conc'd H2SO4
0 to 25 °C, 12 h
HO3S
SO3-
HO
OH
R2N
O
+
NR2
(i) alkylation
(ii) hydrolysis
HO3S
SO3-
HO2C
O
O
CO2H
R2N
O
+
NR2

New rosamine derivatives are needed for further structure-activity relationship (SAR) studies (Figure 7.5.). Better anti-tumor candidates could be prepared by derivatizing the active hits. For example, compound **11d** was found to be highly toxic against cancer cells. The bromo/chloro derivatives of **11d** can be prepared and used to study the halogen effect on the activity. Rosamine **11g** was another active compound. Furan analog should be used to test the heterocycle effect. Iodination of **11g** should lead to more active derivatives due to heave atom effect. The unsymmetrical rosamine **22i**, which had a combination of piperidine and morpholine substituents interestingly had the lowest IC_{50} value among compounds with 2-methyl phenyl at the *meso* position, alluding to the possible importance of an amphilic structure with contrasting hydrophobic and hydrophilic halves. New analogs such as 4-IPh or thiophene derivatives should be test and compared with **11d** and **11g** to double check the previous conclusion.

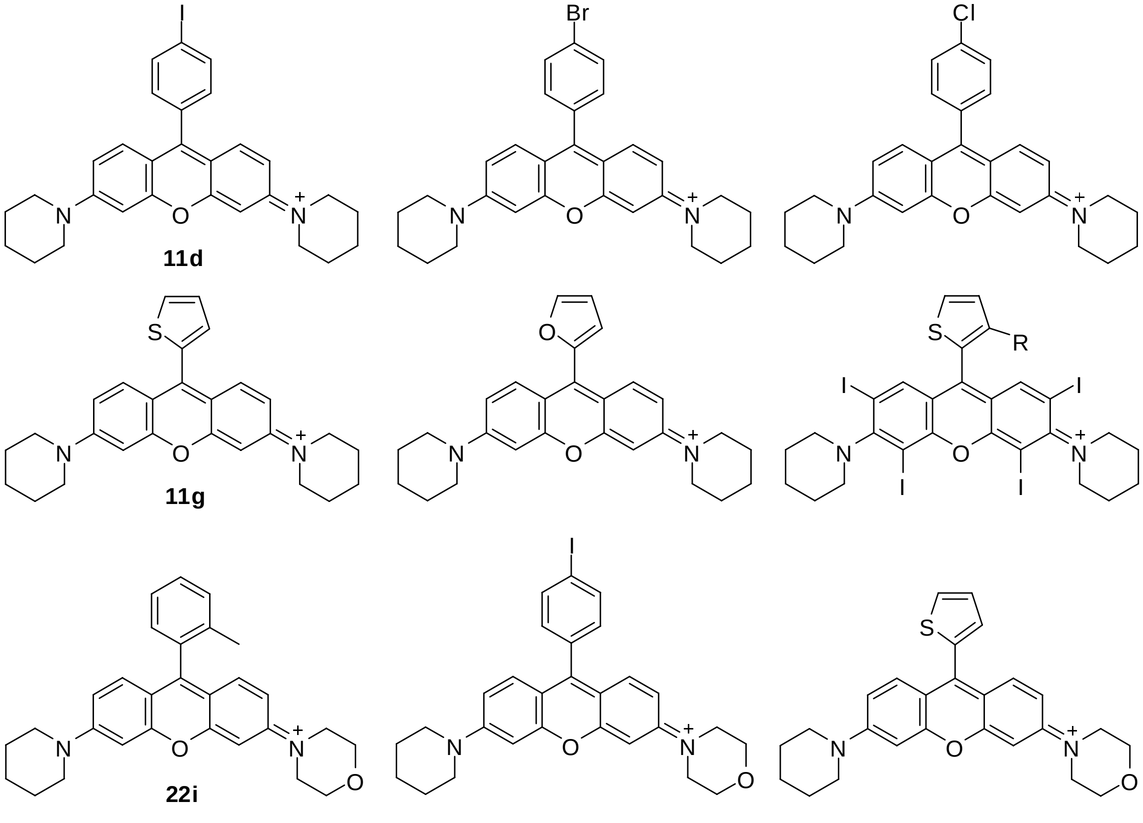
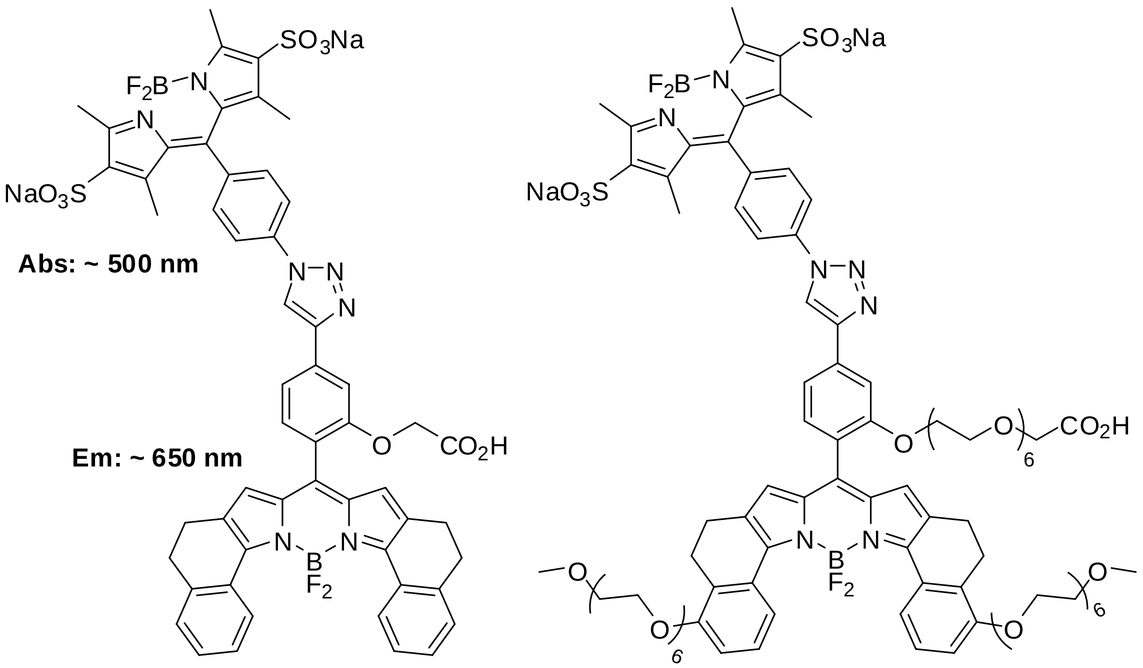


Figure 7.5. Rosamine derivatives for anti-tumor activity study.

7.2.2 Water-soluble through-bond energy transfer cassettes

There are still several problems with the cassettes described in Chapter VI. First, the quantum yields of these cassettes are reduced when they are attached to proteins; hence they are not optimal for biomolecular labeling. Second, some cassettes have very low quantum yields and the energy transfer is not complete in aqueous media. This limits their application in biological studies. Third, the emission of the cassettes is below 700 nm. Thus, these cassettes are still not ideal for *in vivo* imaging. Further improvement is needed to solve these problems.

The problem of low quantum yield in water or on protein might be solved by improving the water solubility of the cassettes. This can be achieved by incorporating more water solubilizing groups into the cassettes. For example, as showed in Figure 7.6., a new cassette **49** should have similar absorption and emission wavelength with **31e**. But it is more soluble in water than **31e** thus should give better quantum yield in aqueous media and on proteins. Scheme 7.5. describes the synthesis of the new cassette **49**.

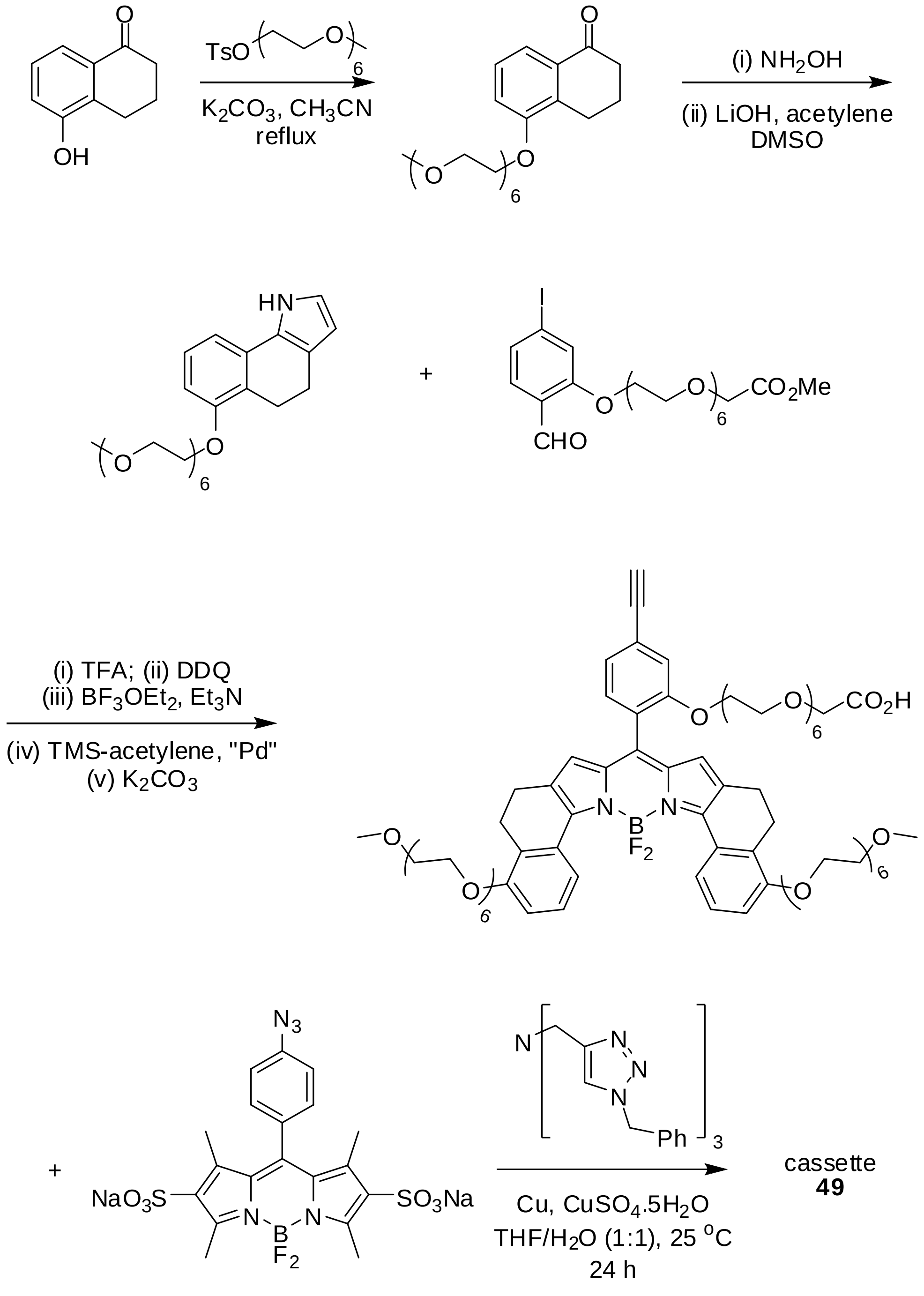


31e
MeOH: Φ: 0.69 , ET: 97%
pH 7.4 PBS Buffer: Φ: 0.24 , ET: 92%
avidin: Φ: 0.03 , ET: 60%

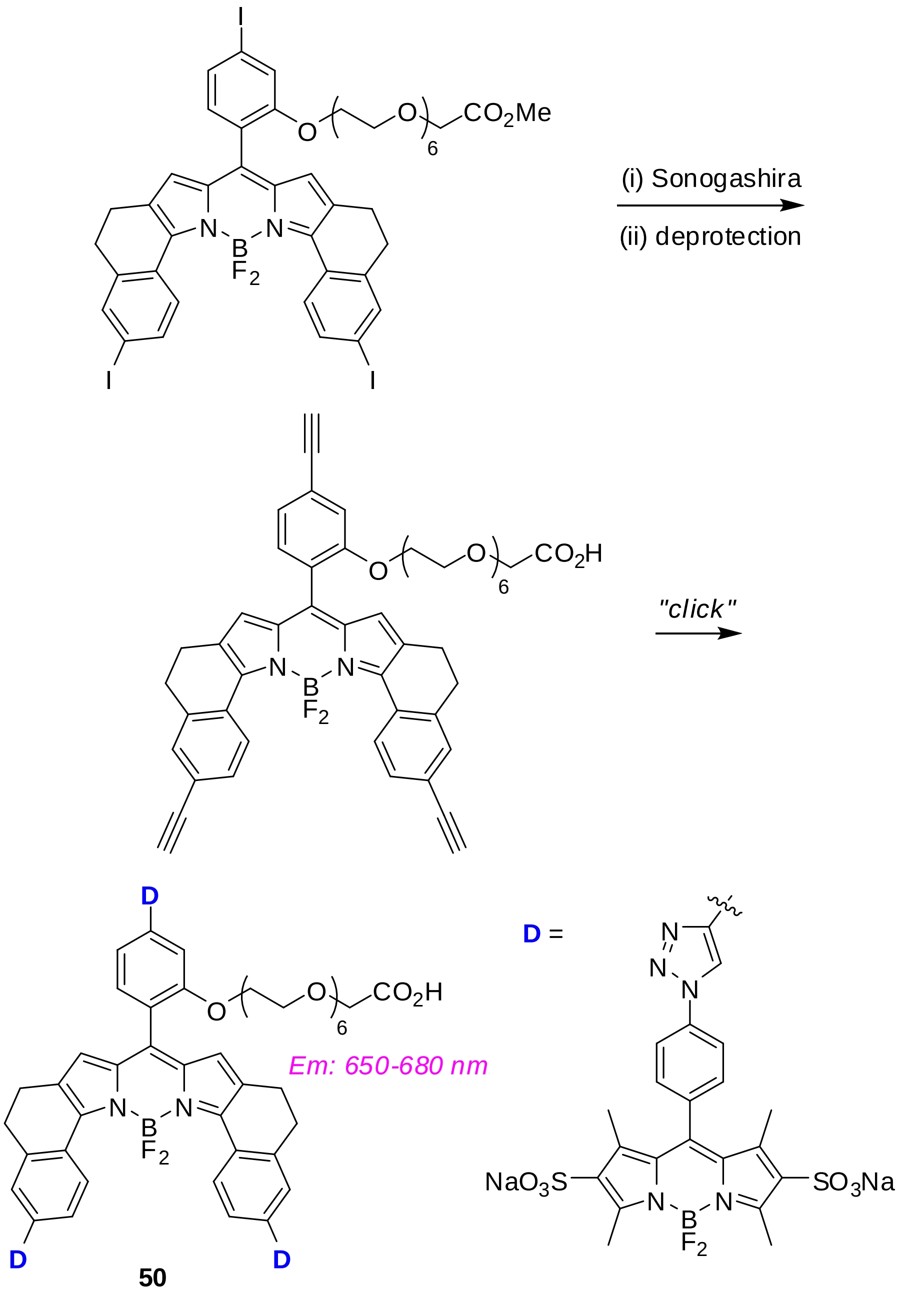
49

Figure 7.6. Water-soluble through-bond energy transfer cassettes.

Scheme 7.5. Synthesis of new water-soluble cassette **49**.


OH
TsO, O, 6
K$_2$CO$_3$, CH$_3$CN
reflux
(i) NH$_2$OH
(ii) LiOH, acetylene
DMSO
HN
+
I
CHO
CO$_2$Me
(i) TFA; (ii) DDQ
(iii) BF$_3$OEt$_2$, Et$_3$N
(iv) TMS-acetylene, "Pd"
(v) K$_2$CO$_3$
CO$_2$H
N, B, N, F$_2$
+
N$_3$
NaO$_3$S
SO$_3$Na
Ph
3
Cu, CuSO$_4$.5H$_2$O
THF/H$_2$O (1:1), 25 °C
24 h
cassette
49

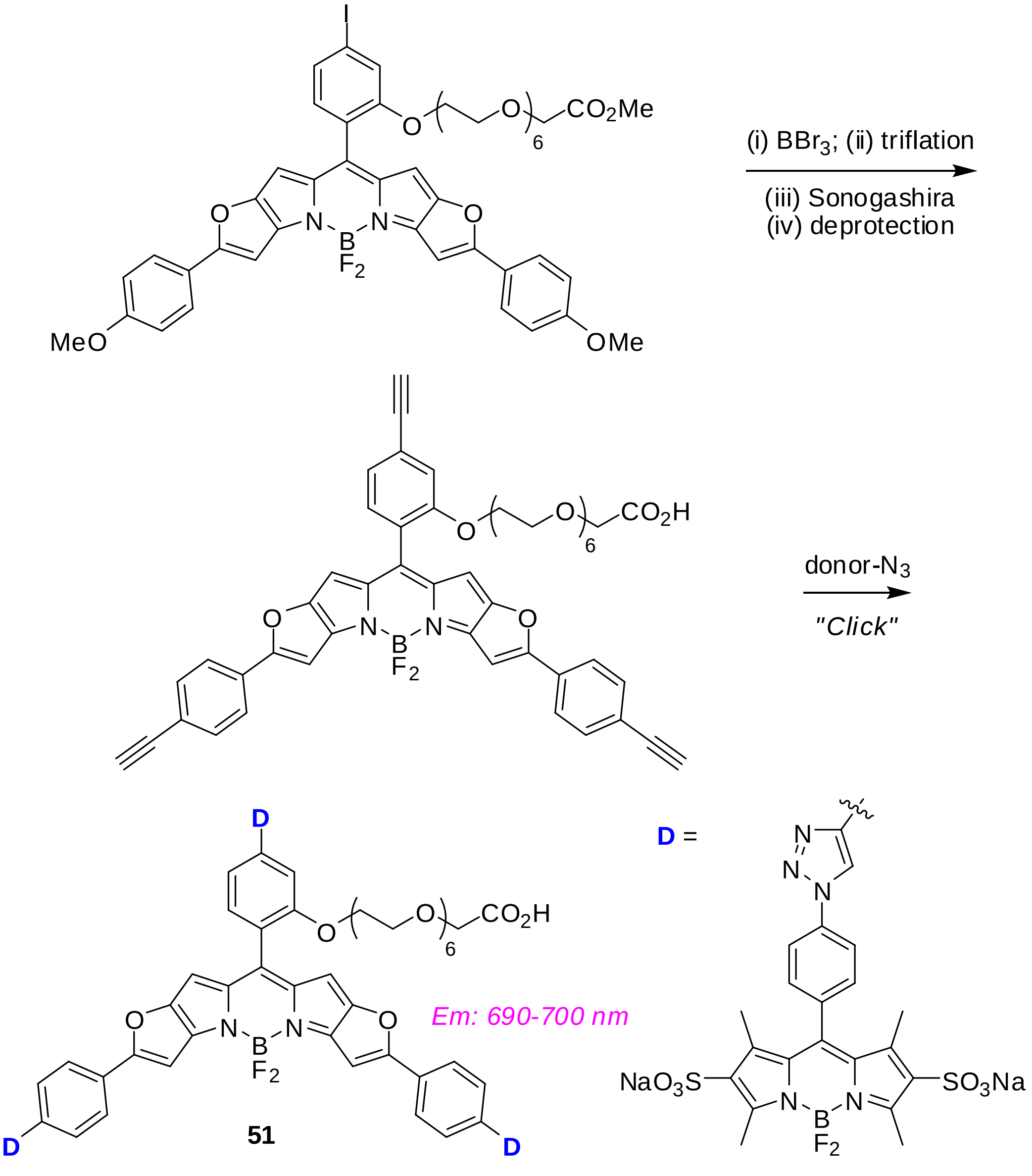
Scheme 7.6. Synthesis of cassette **50**.


I
O
O
6
CO_2Me
N
B
N
F_2
I
I
(i) Sonogashira
(ii) deprotection
O
O
6
CO_2H
N
B
N
F_2
"click"
D
O
O
6
CO_2H
Em: 650-680 nm
N
B
N
F_2
D
D
50
D =
N
N
N
NaO_3S
SO_3Na
N
B
N
F_2

Another possible water-soluble cassette **50** is showed in Scheme 7.6. In this cassette, there are three donor fragments attached to one acceptor thus it can absorb more photons and increase the intensity of emission. Since there will be six sulfonic acids and one hexa-ethylene glycol chain in the cassette, it is going to be highly soluble in aqueous media and the quantum yield should also be high.

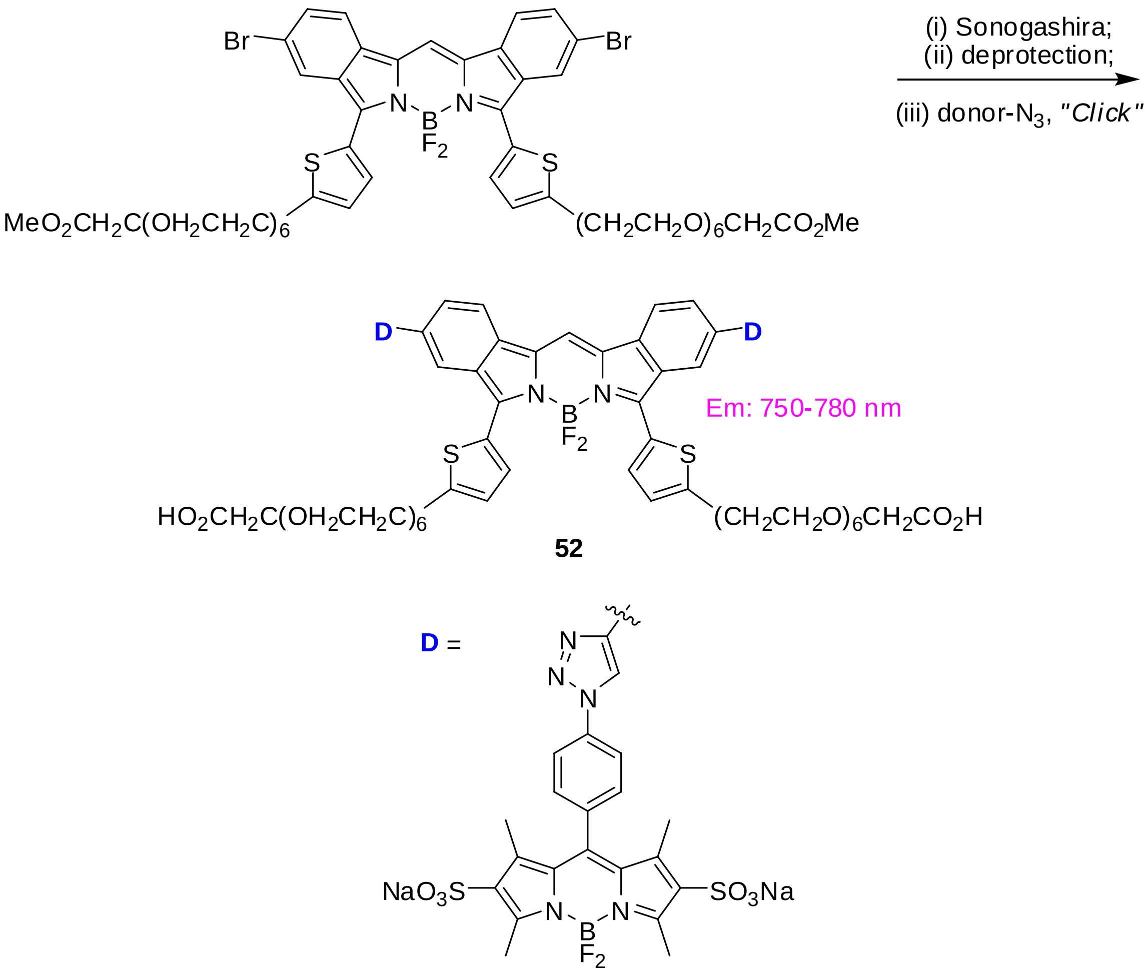
Scheme 7.7. Syntheses of NIR water-soluble BODIPY-BODIPY cassettes **51-52**.

a


I
CO2Me
MeO
OMe
(i) BBr3; (ii) triflation
(iii) Sonogashira
(iv) deprotection
CO2H
donor-N3
"Click"
D
D =
Em: 690-700 nm
51
NaO3S
SO3Na
N
B
F2
O
6

Scheme 7.7. Continued.

b


Br
Br
MeO2CH2C(OH2CH2C)6
(CH2CH2O)6CH2CO2Me
(i) Sonogashira;
(ii) deprotection;
(iii) donor-N3, "Click"
D
D
Em: 750-780 nm
HO2CH2C(OH2CH2C)6
(CH2CH2O)6CH2CO2H
D =
NaO3S
SO3Na

52

Near IR emitting cassettes are highly desirable for *in vivo* imaging because their emission can penetrate tissue deeper than common dyes thus facilitates detection of the labeled targets. Two illustrative examples of such cassettes **51-52** are described in Scheme 7.7. BODIPY dyes that are similar to the proposed acceptor fragments have been reported.[230,238] Therefore, the acceptor parts can be prepared accordingly. Incorporation of alkyne functionalities allows subsequent coupling with donor moiety to afford the through-bond energy transfer cassettes. There are multiple donor fragments in both cassettes thus the brightness of these cassettes will be high.

REFERENCES

(1) Lackowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer: New York, **2006**.

(2) Goppert-Mayer, M. *Annalen der Physik (Berlin, Germany)*, **1931**, *9*, 273-294.

(3) Masters Barry, R.; So Peter, T. C. *Microscopy Research and Technique*, **2004**, *63*, 3-11.

(4) Berti, L.; Xie, J.; Medintz, I. L.; Glazer, A. N.; Mathies, R. A. *Anal. Biochem.*, **2001**, *292*, 188-197.

(5) Ju, J.; Glazer, A. N.; Mathies, R. A. *Nature Med.*, **1996**, *2*, 246-249.

(6) Forster, T. *Discuss. Faraday Soc.*, **1959**, *No. 27*, 7-17.

(7) Forster, T. *Naturwissenschaften*, **1946**, *6*, 166-175.

(8) Burghart, A.; Thoresen, L. H.; Chen, J.; Burgess, K.; Bergstrom, F.; Johansson, L. B. A. N. G. *Chem. Commun.*, **2000**, 2203-2204.

(9) Wan, C. W.; Burghart, A.; Chen, J.; Bergstrom, F.; Johansson, L. B.; Wolford, M. F.; Kim, T. G.; Topp, M. R.; Hochstrasser, R. M.; Burgess, K. *Chemistry-A Euro. J.*, **2003**, *9*, 4430-4441.

(10) Kim, T. G.; Castro, J. C.; Loudet, A.; Jiao, J. G. S.; Hochstrasser, R. M.; Burgess, K.; Topp, M. R. *J. Phys. Chem. A*, **2006**, *110*, 20-27.

(11) Jiao, G.-S.; Thoresen, L. H.; Burgess, K. *J. Am. Chem. Soc.*, **2003**, *125*, 14668-14669.

(12) Dexter, D. L. *J. Chem. Phys.*, **1953**, *21*, 836-850.

(13) Ghiggino, K. P.; Yeow, E. K. L.; Haines, D. J.; Scholes, G. D.; Smith, T. A. *J. Photochem. Photobio., A*, **1996**, *102*, 81-86.

(14) Hermanson, G. T. *Bioconjugate Techniques*, 2nd ed., Academic Press: San Diego, **2008**.

(15) Antos, J. M.; Francis, M. B. *Curr. Opin. Chem. Biol.*, **2006**, *10*, 253-262.

(16) Li, X.; Zhang, L.; Hall, S. E.; Tam, J. P. *Tetrahedron Lett.*, **2000**, *41*, 4069-4073.

(17) Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.*, **1992**, *3*, 138-146.

(18) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.*, **2006**, *45*, 5307-5311.

(19) Yeo, D. S. Y.; Srinivasan, R.; Chen, G. Y. J.; Yao, S. Q. *Chemistry-A Euro. J.*, **2004**, *10*, 4664-4672.

(20) Giriat, I.; Muir, T. W. *J. Am. Chem. Soc.*, **2003**, *125*, 7180-7181.

(21) Dibowski, H.; Schmidtchen, F. P. *Angew. Chem., Int. Ed.*, **1998**, *37*, 476-478.

(22) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.*, **2006**, *35*, 225-249.

(23) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. *Nat. Chem. Biol.*, **2009**, *advanced online publication*.

(24) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science*, **1994**, *263*, 802-805.

(25) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *Trends Biochem. Sci.*, **1995**, *20*, 448-455.

(26) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.*, **2004**, *126*, 7754-7755.

(27) George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.*, **2004**, *126*, 8896-8897.

(28) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods*, **2005**, *2*, 99-104.

(29) Chen, I.; Ting, A. Y. *Curr. Opin. Biotechnol.*, **2005**, *16*, 35-40.

(30) Jager, M.; Nir, E.; Weiss, S. *Protein Science*, **2006**, *15*, 640-646.

(31) Lin, C.-W.; Ting, A. Y. *J. Am. Chem. Soc.*, **2006**, *128*, 4542-4543.

(32) Duckworth, B. P.; Zhang, Z.; Hosokawa, A.; Distefano, M. D. *ChemBioChem*, **2007**, *8*, 98-105.

(33) Dierks, T.; Schmidt, B.; Borissenko, L. V.; Peng, J.; Preusser, A.; Mariappan, M.; von Figura, K. *Cell*, **2003**, *113*, 435-444.

(34) Rush, J. S.; Bertozzi, C. R. *J. Am. Chem. Soc.*, **2008**, *130*, 12240-12241.

(35) Gautier, A.; Juillerat, A.; Heinis, C.; Correa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chemistry & Biolology*, **2008**, *15*, 128-136.

(36) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.*, **2003**, *21*, 86-89.

(37) O'Hare, H. M.; Johnsson, K.; Gautier, A. *Curr. Opin. Struct. Biol.*, **2007**, *17*, 488-494.

(38) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chem. Biol.*, **2008**, *3*, 373-382.

(39) Farinas, J.; Verkman, A. S. *J. Biol. Chem.*, **1999**, *274*, 7603-7606.

(40) Miller, L. W.; Cornish, V. W. *Curr. Opin. Chem. Biol.*, **2005**, *9*, 56-61.

(41) Clackson, T.; Yang, W.; Rozamus, L. W.; Hatada, M.; Amara, J. F.; Rollins, C. T.; Stevenson, L. F.; Magari, S. R.; Wood, S. A.; Courage, N. L.; Lu, X.; Cerasoli, F.; Gilman, M.; Holt, D. A. *Proc. Natl. Acad. Sci. U. S. A.*, **1998**, *95*, 10437-10442.

(42) Kaufman, R. J.; Bertino, J. R.; Schimke, R. T. *J. Biol. Chem.*, **1978**, *253*, 5852-5860.

(43) Miller, L. W.; Sable, J.; Goelet, P.; Sheetz, M. P.; Cornish, V. W. *Angew. Chem. Int. Ed.*, **2004**, *43*, 1672-1675.

(44) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science*, **1998**, *281*, 269-272.

(45) Adams, S. R.; Tsien, R. Y. *Nat. Protoc.*, **2008**, *3*, 1527-1534.

(46) Marks, K. M.; Nolan, G. P. *Nat. Methods*, **2006**, *3*, 591-596.

(47) Rozinov, M. N.; Nolan, G. P. *Chemistry & Biology*, **1998**, *5*, 713-728.

(48) Guignet, E. G.; Hovius, R.; Vogel, H. *Nat. Biotechnol.*, **2004**, *22*, 440-444.

(49) Ojida, A.; Honda, K.; Shinmi, D.; Kiyonaka, S.; Mori, Y.; Hamachi, I. *J. Am. Chem. Soc.*, **2006**, *128*, 10452-10459.

(50) Magzoub, M.; Graeslund, A. *Q. Rev. Biophys.*, **2004**, *37*, 147-195.

(51) Brooks, H.; Lebleu, B.; Vives, E. *Adv. Drug Delivery Rev.*, **2005**, *57*, 559-577.

(52) Morris, M. C.; Robert-Hebmann, V.; Chaloin, L.; Mery, J.; Heitz, F.; Devaux, C.; Goody, R. S.; Divita, G. *J. Biol. Chem.*, **1999**, *274*, 24941-24946.

(53) Morris, M. C.; Depollier, J.; Mery, J.; Heitz, F.; Divita, G. *Nat. Biotechnol.*, **2001**, *19*, 1173-1176.

(54) Futaki, S. *Adv. Drug Delivery Rev.*, **2005**, *57*, 547-558.

(55) Kosuge, M.; Takeuchi, T.; Nakase, I.; Jones, A. T.; Futaki, S. *Bioconjugate Chem.*, **2008**, *19*, 656-664.

(56) Didenko, V. V.; Ngo, H.; Baskin, D. S. *Anal. Biochem.*, **2005**, *344*, 168-173.

(57) Deshayes, S.; Heitz, A.; Morris, M. C.; Charnet, P.; Divita, G.; Heitz, F. *Biochemistry*, **2004**, *43*, 1449-1457.

(58) Henriques, S. T.; Castanho, M. A. R. B. *Biochemistry*, **2004**, *43*, 9716-9724.

(59) Weller, K.; Lauber, S.; Lerch, M.; Renaud, A.; Merkle, H. P.; Zerbe, O. *Biochemistry*, **2005**, *44*, 15799-15811.

(60) Mendelsohn, A. R.; Brent, R. *Science*, **1999**, *284*, 1948-1950.

(61) Phizicky, E. M.; Fields, S. *Microbiol. Rev.*, **1995**, *59*, 94-123.

(62) Fields, S.; Song, O. K. *Nature*, **1989**, *340*, 245-246.

(63) Griffiths, A. D.; Winter, G. *Chemtracts: Biochem. Mol. Biol.*, **1990**, *1*, 35-37.

(64) Rossi, F.; Charlton, C. A.; Blau, H. M. *Proc. Natl. Acad. Sci. U. S. A.*, **1997**, *94*, 8405-8410.

(65) Miyawaki, A.; Tsien, R. Y. *Methods Enzymol.*, **2000**, *327*, 472-500.

(66) Zimmer, M. *Chem. Rev.*, **2002**, *102*, 759-781.

(67) Tsien, R. Y. *Annu. Rev. Biochem.*, **1998**, *67*, 509-544.

(68) Davenport, D.; Nicol, J. A. C. *Proc. R. Soc. London, B*, **1955**, *144*, 399-411.

(69) Osamu Shimomura, F. H. J. Y. S. *J. Cell. Comp. Physiol.*, **1962**, *59*, 223-239.

(70) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. *Gene*, **1992**, *111*, 229-233.

(71) Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U. S. A.*, **1994**, *91*, 12501-12504.

(72) Helm, R.; Cubitt, A. B.; Tsien, R. Y. *Nature*, **1995**, *373*, 663-664.

(73) Gross, L. A.; Baird, G. S.; Hoffman, R. C.; Baldridge, K. K.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U. S. A.*, **2000**, *97*, 11990-11995.

(74) Shimomura, O. *FEBS Lett.*, **1979**, *104*, 220-222.

(75) Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F. G.; Ward, W. W. *Biochemistry*, **1993**, *32*, 1212-1218.

(76) Orm, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science*, **1996**, *273*, 1392-1395.

(77) Barondeau, D. P.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Biochemistry*, **2005**, *44*, 1960-1970.

(78) Kojima, S.; Ohkawa, H.; Hirano, T.; Maki, S.; Niwa, H.; Ohashi, M.; Inouye, S.; Tsuji, F. I. *Tetrahedron Lett.*, **1998**, *39*, 5239-5242.

(79) Erlenmeyer, E. *Justus Liebigs Ann. Chem.*, **1893**, *275*, 1-8.

(80) Follenius-Wund, A.; Bourotte, M.; Schmitt, M.; Iyice, F.; Lami, H.; Bourguignon, J.-J.; Haiech, J.; Pigault, C. *Biophys. J.*, **2003**, *85*, 1839-1850.

(81) Bourotte, M.; Schmitt, M.; Follenius-Wund, A.; Pigault, C.; Haiech, J.; Bourguignon, J.-J. *Tetrahedron Lett.*, **2004**, *45*, 6343-6348.

(82) Morise, H.; Shimomura, O.; Johnson, F. H.; Winant, J. *Biochemistry*, **1974**, *13*, 2656-2662.

(83) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. *Proc. Natl. Acad. Sci. U. S. A.*, **1996**, *93*, 13617-13622.

(84) Stoner-Ma, D.; Melief, E. H.; Nappa, J.; Ronayne, K. L.; Tonge, P. J.; Meech, S. R. *J. Phys. Chem. B*, **2006**, *110*, 22009-22018.

(85) He, X.; Bell, A. F.; Tonge, P. J. *Org. Lett.*, **2002**, *4*, 1523-1526.

(86) Dong, J.; Solntsev, K. M.; Poizat, O.; Tolbert, L. M. *J. Am. Chem. Soc.*, **2007**, ASAP.

(87) Chen, K.-Y.; Cheng, Y.-M.; Lai, C.-H.; Hsu, C.-C.; Ho, M.-L.; Lee, G.-H.; Chou, P.-T. *J. Am. Chem. Soc.*, **2007**, *129*, 4510-4511.

(88) Yampolsky, I. V.; Remington, S. J.; Martynov, V. I.; Potapov, V. K.; Lukyanov, S.; Lukyanov, K. A. *Biochemistry*, **2005**, *44*, 5788-5793.

(89) Herz, W. *J. Am. Chem. Soc.*, **1949**, *71*, 3982-3984.

(90) Takeuchi, H.; Hagiwara, S.; Eguchi, S. *Tetrahedron*, **1989**, *45*, 6375-6386.

(91) He, X.; Bell, A. F.; Tonge, P. J. *FEBS Lett.*, **2003**, *549*, 35-38.

(92) de Dios, A.; Luz de la Puente, M.; Rivera-Sagredo, A.; Espinosa, J. F. *Can. J. Chem.*, **2002**, *80*, 1302-1307.

(93) Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Chem. Theory Comput.*, **2006**, *2*, 292-299.

(94) Ando, R.; Mizuno, H.; Miyawaki, A. *Science*, **2004**, *306*, 1370-1373.

(95) Habuchi, S.; Dedecker, P.; Hotta, J.-i.; Flors, C.; Ando, R.; Mizuno, H.; Miyawaki, A.; Hofkens, J. *Photochem. Photobiol. Sci.*, **2006**, *5*, 567-576.

(96) Mizuno, H.; Mal, T. K.; Walchli, M.; Kikuchi, A.; Fukano, T.; Ando, R.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Ikura, M.; Miyawaki, A. *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 9227-9232.

(97) Pakhomov, A. A.; Martynov, V. I. *Chemistry & Biology*, **2008**, *15*, 755-764.

(98) Loudet, A.; Burgess, K. *Chem. Rev.*, **2007**, *107*, 4891-4932.

(99) Sonogashira, K.; Tohda, Y.; Hagihara, N. *Tetrahedron Lett.*, **1975**, *16*, 4467-4470.

(100) Albarran, B.; To, R.; Stayton, P. S. *Protein Eng.*, **2005**, *18*, 147-152.

(101) http://probes.invitrogen.com; Invitrogen Corporation: 2006.

(102) Weber, G.; Teale, F. W. J. *Trans. Faraday Soc.*, **1958**, *54*, 640-648.

(103) Treibs, A.; Kreuzer, F. H. *Justus Liebigs Ann. Chem.*, **1968**, *718*, 208-223.

(104) Ulrich, G.; Ziessel, R.; Harriman, A. *Angew. Chem. Int. Ed.*, **2008**, *47*, 1184-1201.

(105) Li, Z.; Mintzer, E.; Bittman, R. *J. Org. Chem.*, **2006**, *71*, 1718-1721.

(106) Wagner, R. W.; Lindsey, J. S. *Pure Applied Chem.*, **1996**, *68*, 1373-1380.

(107) Goud, T. V.; Tutar, A.; Biellmann, J.-F. *Tetrahedron*, **2006**, *62*, 5084-5091.

(108) Nicolaou, K. C.; Claremon, D. A.; Papahatjis, D. P. *Tetrahedron Lett.*, **1981**, *22*, 4647-4650.

(109) Tahtaoui, C.; Thomas, C.; Rohmer, F.; Klotz, P.; Duportail, G.; Mely, Y.; Bonnet, D.; Hibert, M. *J. Org. Chem.*, **2007**, *72*, 269-272.

(110) Van Koeveringe, J. A.; Lugtenburg, J. *Recl. Trav. Chim. Pays-Bas*, **1977**, *96*, 55-58.

(111) Vos de Wael, E.; Pardoen, J. A.; Van Koeveringe, J. A.; Lugtenburg, J. *Recl. Trav. Chim. Pays-Bas*, **1977**, *96*, 306-309.

(112) Wories, H. J.; Koek, J. H.; Lodder, G.; Lugtenburg, J.; Fokkens, R.; Driessen, O.; Mohn, G. R. *Recl. Trav. Chim. Pays-Bas*, **1985**, *104*, 288-291.

(113) Falk, H.; Hofer, O.; Lehner, H. *Monatsh. Chem.*, **1974**, *105*, 169-178.

(114) Wood, T. E.; Berno, B.; Beshara, C. S.; Thompson, A. *J. Org. Chem.*, **2006**, *71*, 2964-2971.

(115) Shah, M.; Thangaraj, K.; Soong, M. L.; Wolford, L.; Boyer, J. H.; Politzer, I. R.; Pavlopoulos, T. G. *Heteroatom Chem.*, **1990**, *1*, 389-399.

(116) Sens, R.; Drexhage, K. H. *J. Lumin.*, **1981**, *24-25*, 709-712.

(117) Wittmershaus, B. P.; Skibicki, J. J.; McLafferty, J. B.; Zhang, Y.-Z.; Swan, S. *J. Fluoresc.*, **2001**, *11*, 119-128.

(118) Magde, D.; Wong, R.; Seybold, P. G. *Photochem. Photobiol.*, **2002**, *75*, 327-334.

(119) Haugland, R. P. *Handbook of Fluorescent Probes and Research Chemicals*, 9th ed., Molecular Probes: Eugene, OR, **2002**.

(120) Noelting, E.; Dziewonski, K. *Ber. Dtsch. Chem. Ges.*, **1905**, *38*, 3516-3527.

(121) Liu, J.; Diwu, Z.; Leung, W.-Y.; Lu, Y.; Patch, B.; Haugland, R. P. *Tetrahedron Lett.*, **2003**, *44*, 4355-4359.

(122) Panchuk–Voloshina, N.; Haugland, R. P.; Bishop–Stewart, J.; Bhalgat, M. K.; Millard, P. J.; Mao, F.; Leung, W.-Y.; Haugland, R. P. *J. Histochem. Cytochem.*, **1999**, *47*, 1179-1188.

(123) Dujols, V.; Ford, F.; Czarnik, A. W. *J. Am. Chem. Soc.*, **1997**, *119*, 7386-7387.

(124) Kommareddi, S.; Abramowsky, C. R.; Swinehart, G. L.; Hrabak, L. *Hum. Pathol.*, **1984**, *15*, 1085-1089.

(125) Tombline, G.; Holt, J. J.; Gannon, M. K.; Donnelly, D. J.; Wetzel, B.; Sawada, G. A.; Raub, T. J.; Detty, M. R. *Biochemistry*, **2008**, *47*, 3294-3307.

(126) Tombline, G.; Donnelly, D. J.; Holt, J. J.; You, Y.; Ye, M.; Gannon, M. K.; Nygren, C. L.; Detty, M. R. *Biochemistry*, **2006**, *45*, 8034-8047.

(127) Gottesman, M. M.; Ling, V. *FEBS Lett.*, **2006**, *580*, 998-1009.

(128) Gottesman, M. M.; Fojo, T.; Bates, S. E. *Nat. Rev. Cancer*, **2002**, *2*, 48-58.

(129) Ambudkar, S. V.; Kim, I.-W.; Sauna, Z. E. *Eur. J. Pharm. Sci.*, **2006**, *27*, 392-400.

(130) Al-Shawi, M. K.; Omote, H. *J. Bioenerg. Biomembr.*, **2005**, *37*, 489-496.

(131) Gibson, S. L.; Hilf, R.; Donnelly, D. J.; Detty, M. R. *Bioorg. Med. Chem.*, **2004**, *12*, 4625-4631.

(132) Scotto, K. W. *Oncogene*, **2003**, *22*, 7496-7511.

(133) Tiberghien, F.; Loor, F. *Anti-Cancer Drugs*, **1996**, *7*, 568-578.

(134) Hirsch-Ernst, K. I.; Ziemann, C.; Rustenbeck, I.; Kahl, G. F. *Toxicology*, **2001**, *167*, 47-57.

(135) Loor, F.; Tiberghien, F.; Wenandy, T.; Didier, A.; Traber, R. *J. Med. Chem.*, **2002**, *45*, 4598-4612.

(136) Kessel, D. *Cancer Commun.*, **1989**, *1*, 145-149.

(137) Chatchanok Loetchutinat, C. S. C. M.-G. A. G.-S. *Eur. J. Biochem.*, **2003**, *270*, 476-485.

(138) Eytan, G. D.; Regev, R.; Oren, G.; Hurwitz, C. D.; Assaraf, Y. G. *Eur. J. Biochem.*, **1997**, *248*, 104-112.

(139) Shapiro, A. B.; Ling, V. *Eur. J. Biochem.*, **1998**, *254*, 189-193.

(140) Johnson, L. V.; Walsh, M. L.; Chen, L. B. *Proc. Natl. Acad. Sci. U. S. A.*, **1980**, *77*, 990-994.

(141) McBride, H. M.; Neuspiel, M.; Wasiak, S. *Curr. Biol.*, **2006**, *16*, R551-R560.

(142) Bragadin, M. *Mitochondrial Pharmacol. Toxicol.*, **2006**, 1-22.

(143) Murphy, M. P. *Biochim. Biophys. Acta, Bioenerg.*, **2008**, *1777*, 1028-1031.

(144) Rottenberg, H. *J. Membr. Biol.*, **1984**, *81*, 127-138.

(145) Singh, G.; Shaughnessy, S. G. *Can. J. Physiol. Pharmacol.*, **1988**, *66*, 243-245.

(146) Davis, S.; Weiss, M. J.; Wong, J. R.; Lampidis, T. J.; Chen, L. B. *J. Biol. Chem.*, **1985**, *260*, 13844-13850.

(147) Nadakavukaren, K. K.; Nadakavukaren, J. J.; Chen, L. B. *Cancer Res.*, **1985**, *45*, 6093-6099.

(148) Modica-Napolitano, J. S.; Aprille, J. R. *Cancer Res.*, **1987**, *47*, 4361-4365.

(149) Lampidis, T. J.; Bernal, S. D.; Summerhayes, I. C.; Chen, L. B. *Cancer Res.*, **1983**, *43*, 716-720.

(150) Lampidis, T. J.; Bernal, S. D.; Summerhayes, I. C.; Chen, L. B. *Ann. N. Y. Acad. Sci.*, **1982**, *397*, 299-302.

(151) Bernal, S. D.; Lampidis, T. J.; Summerhayes, I. C.; Chen, L. B. *Science*, **1982**, *218*, 1117-1119.

(152) Bernal, S. D.; Lampidis, T. J.; McIsaac, R. M.; Chen, L. B. *Science*, **1983**, *222*, 169-172.

(153) Zhang, Y.-Z.; Haugland, R. P., U.S. Patent 5,686,261, **1997**

(154) Mao, F.; Leung, W.-Y.; Haugland, R. P., WO 99/15517, **1999**

(155) Wang, Z.-Q.; Diwu, Z.; Francisco-Reyes, J.; Yi, G. G. *Chem. Lett.*, **2005**, *34*, 404-405.

(156) Jiao, G.-S.; Castro, J. C.; Thoresen, L. H.; Burgess, K. *Org. Lett.*, **2003**, *5*, 3675-3677.

(157) Adamczyk, M.; Grote, J. *Bioorg. Med. Chem. Lett.*, **2000**, *10*, 1539-1541.

(158) Beach, S. F.; Hepworth, J. D.; Mason, D.; Swarbrick, E. A. *Dyes and Pigments*, **1999**, *42*, 71-77.

(159) Gannon, M. K., II; Detty, M. R. *J. Org. Chem.*, **2007**, *72*, 2647-2650.

(160) Brennan, N. K.; Donnelly, D. J.; Detty, M. R. *J. Org. Chem.*, **2003**, *68*, 3344-3347.

(161) Del Valle, D. J.; Donnelly, D. J.; Holt, J. J.; Detty, M. R. *Organometallics*, **2005**, *24*, 3807-3810.

(162) Holt, J. J.; Calitree, B. D.; Vincek, J.; Gannon, M. K., II; Detty, M. R. *J. Org. Chem.*, **2007**, *72*, 2690-2693.

(163) Holt, J. J.; Gannon, M. K.; Tombline, G.; McCarty, T. A.; Page, P. M.; Bright, F. V.; Detty, M. R. *Bioorg. Med. Chem.*, **2006**, *14*, 8635-8643.

(164) Gibson, S. L.; Holt, J. J.; Ye, M.; Donnelly, D. J.; Ohulchanskyy, T. Y.; You, Y.; Detty, M. R. *Bioorg. Med. Chem.*, **2005**, *13*, 6394-6403.

(165) Wagner, S. J.; Skripchenko, A.; Donnelly, D. J.; Ramaswamy, K.; Detty, M. R. *Bioorg. Med. Chem.*, **2005**, *13*, 5927-5935.

(166) Ahn, Y.-H.; Lee, J.-S.; Chang, Y.-T. *J. Am. Chem. Soc.*, **2007**, *129*, 4510-4511.

(167) Parish, E. J. *e-EROS Encycl. Reagents Org. Synth.*, **2001**.

(168) Farrell, P. G.; Moskowitz, D.; Terrier, F. *Synth. Commun.*, **1993**, *23*, 231-235.

(169) Smith, G. A.; Watson, D. W.; Voorheis, P., WO: 2005098437, **2005**

(170) Awad, W. I.; El-Bieh, A.; Abdel Latif, T. M. *Egypt. J. Chem.*, **1983**, *25*, 415-419.

(171) Gorvin, J. H.; Whalley, D. P. *J. Chem. Soc., Perkin Trans. 1*, **1979**, 1364-1370.

(172) Detty, M. R.; Prasad, P. N.; Donnelly, D. J.; Ohulchanskyy, T.; Gibson, S. L.; Hilf, R. *Bioorg. Med. Chem.*, **2004**, *12*, 2537-2544.

(173) Chalkley, L. *J. Org. Chem.*, **1961**, *26*, 408-412.

(174) P. Ehrlich, L. B. *Ber. Dtsch. Chem. Ges.*, **1913**, *46*, 1931-1951.

(175) Kenmoku, S.; Urano, Y.; Kojima, H.; Nagano, T. *J. Am. Chem. Soc.*, **2007**, *129*, 7313-7318.

(176) Holt, J. J.; Calitree, B. D.; Vincek, J.; Gannon, M. K., II; Detty, M. R. *J. Org. Chem.*, **2007**, *72*, 2690-2693.

(177) Del Valle, D. J.; Donnelly, D. J.; Holt, J. J.; Detty, M. R. *Organometallics*, **2005**, *24*, 3807-3810.

(178) Brennan, N. K.; Donnelly, D. J.; Detty, M. R. *J. Org. Chem.*, **2003**, *68*, 3344-3347.

(179) Miller, E. W.; Albers, A. E.; Pralle, A.; Isacoff, E. Y.; Chang, C. J. *J. Am. Chem. Soc.*, **2005**, *127*, 16652-16659.

(180) Wu, L.; Burgess, K. *Org. Lett.*, **2008**, *10*, 1779-1782.

(181) DeSilva, A. P.; Gunaratne, H. G. N.; Gunnlaugsson, T.; Huxley, A. J. M.; McCoy, C. P.; Rademacher, J. T.; Rice, T. E. *Chem. Rev.*, **1997**, *97*, 1515-1566.

(182) Saha, S.; Samanta, A. *J. Phys. Chem. A*, **2002**, *106*, 4763-4771.

(183) Bhattacharya, B.; Samanta, A. *Chem. Phys. Lett.*, **2007**, *442*, 316-321.

(184) Zachariasse, K. A.; Grobys, M.; von der Haar, T.; Hebecker, A.; Il'ichev, Y. V.; Morawski, O.; Rueckert, I.; Kuehnle, W. *J. Photochem. Photobiol., A*, **1997**, *105*, 373-383.

(185) Martin, M. M.; Lindqvist, L. *J. Luminescence*, **1975**, *10*, 381-390.

(186) Urano, Y.; Kamiya, M.; Kanda, K.; Ueno, T.; Hirose, K.; Nagano, T. *J. Am. Chem. Soc.*, **2005**, *127*, 4888-4894.

(187) Karstens, T.; Kobs, K. *J. Phys. Chem.*, **1980**, *84*, 1871-1872.

(188) Magde, D.; Brannon, J. H.; Cremers, T. L.; Olmsted III, J. *J. Phys. Chem.*, **1979**, *83*, 696-699.

(189) Mosmann, T. *J. Immunol. Methods*, **1983**, *65*, 55-63.

(190) El Baraka, M.; Deumie, M.; Viallet, P.; Lampidis, T. J. *J. Photochem. Photobio., A*, **1991**, *56*, 295-311.

(191) Onganer, Y.; Quitevis, E. L. *J. Phys. Chem.*, **1992**, *96*, 7996-8001.

(192) Acemioglu, B.; Arik, M.; Onganer, Y. *J. Lumin.*, **2002**, *97*, 153-160.

(193) Drexhage, K. H. *Laser Focus*, **1973**, *9*, 35.

(194) Vogel, M.; Rettig, W.; Sens, R.; Drexhage, K. H. *Chem. Phys. Lett.*, **1988**, *147*, 461-465.

(195) Vogel, M.; Rettig, W.; Sens, R.; Drexhage, K. H. *Chem. Phys. Lett.*, **1988**, *147*, 452-460.

(196) Wolfgang, R. *Angew. Chem. Int. Ed.*, **1986**, *25*, 971-988.

(197) Lopez Arbeloa, F.; Aguirresacona, I. U.; Arbeloa, I. L. *Chem. Phys.*, **1989**, *130*, 371-378.

(198) Singh, A. K.; Darshi, M.; Kanvah, S. *J. Phys. Chem. A*, **2000**, *104*, 464-471.

(199) Lopez Arbeloa, F.; Lopez Arbeloa, T.; Tapia Estevez, M. J.; Lopez Arbeloa, I. *J. Phys. Chem.*, **1991**, *95*, 2203-2208.

(200) Lopez Arbeloa, T.; Tapia Estevez, M. J.; Lopez Arbeloa, F.; Urretxa Aguirresacona, I.; Lopez Arbeloa, I. *J. Lumin.*, **1991**, *48-49*, 400-404.

(201) Arik, M.; Meral, K.; Onganer, Y. *J. Lumin.*, **2009**, *129*, 599-604.

(202) Lim, H.; Cheong, H.; Lee, J. S.; Yoon, K. B. *J. Korean Phys. Soc.*, **2008**, *53*, 2328-2331.

(203) Busby, M.; Blum, C.; Tibben, M.; Fibikar, S.; Calzaferri, G.; Subramaniam, V.; De Cola, L. *J. Am. Chem. Soc.*, **2008**, *130*, 10970-10976.

(204) Arik, M.; Onganer, Y. *Chem. Phys. Lett.*, **2003**, *375*, 126-133.

(205) Ramamurthy, V.; Sanderson, D. R.; Eaton, D. F. *J. Am. Chem. Soc.*, **1993**, *115*, 10438-10439.

(206) Gormally, J.; Higson, S. *J. Chem. Soc., Faraday Trans. 1*, **1986**, *82*, 157-160.

(207) Acemioglu, B.; ArIk, M.; Efeoglu, H.; Onganer, Y. *J. Mol. Struct. (Theochem)*, **2001**, *548*, 165-171.

(208) Shapiro, H. M. *Cytometry*, **1981**, *2*, 143-150.

(209) Darzynkiewicz, Z.; Kapuscinski, J.; Traganos, F.; Crissman, H. A. *Cytometry*, **1987**, *8*, 138-145.

(210) Darzynkiewicz, Z.; Juan, G.; Srour Edward, F. *Curr. Protoc. Cytom.*, **2004**, *Chapter 7*, Unit 7.3.

(211) Kapuscinski, J.; Darzynkiewicz, Z. *Cytometry*, **1987**, *8*, 129-137.

(212) Feng, S.; Tang, A.; Fan, J. *Fenxi Huaxue*, **2001**, *29*, 1315-1317.

(213) Feng, S.; Tang, A.; Fan, J. *Fenxi Huaxue*, **2001**, *29*, 558-560.

(214) Feng, S.-l.; Tang, J.-m.; Fan, J. *Guangpuxue Yu Guangpu Fenxi*, **2003**, *23*, 322-324.

(215) Dare-Doyen, S.; Doizi, D.; Guilbaud, P.; Djedaieni-Pilard, F.; Perly, B.; Millie, P. *J. Phys. Chem. B*, **2003**, *107*, 13803-13812.

(216) Scala-Valero, C., *Thesis: Synthese et caracterisation de nouveaux colorants laser*, University Of Orleans: Orleans, France, **1997**.

(217) Magde, D.; Brannon, J. H.; Cremers, T. L.; Olmsted, J., III *J. Phys. Chem.*, **1979**, *83*, 696-699.

(218) Han, J.; Jose, J.; Mei, E.; Burgess, K. *Angew. Chem., Int. Ed.*, **2007**, *46*, 1684-1687.

(219) Jiao, G.-S.; Kim, T. G.; Topp, M. R.; Burgess, K. *Org. Lett.*, **2004**, *6*, 1701-1704.

(220) Jiao, G.-S.; Loudet, A.; Lee, H. B.; Kalinin, S.; Johansson, L. B. A.; Burgess, K. *Tetrahedron*, **2003**, *59*, 3109-3116.

(221) Jiao, G.-S.; Thoresen, L. H.; Kim, T. G.; Haaland, W. C.; Gao, F.; Topp, M. R.; Hochstrasser, R. M.; Metzker, M. L.; Burgess, K. *Chem.--Eur. J.*, **2006**, *12*, 7816-7826, S7816/7811-S7816/7838.

(222) Bandichhor, R.; Petrescu, A. D.; Vespa, A.; Kier, A. B.; Schroeder, F.; Burgess, K. *J. Am. Chem. Soc.*, **2006**, *128*, 10688-10689.

(223) Wu, L.; Burgess, K. *J. Am. Chem. Soc.*, **2008**, *130*, 4089-4096.

(224) Li, L.; Han, J.; Nguyen, B.; Burgess, K. *J. Org. Chem.*, **2008**, *73*, 1963-1970.

(225) Lewis Warren, G.; Green Luke, G.; Grynszpan, F.; Radic, Z.; Carlier Paul, R.; Taylor, P.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.*, **2002**, *41*, 1053-1057.

(226) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.*, **2001**, *40*, 2004-2021.

(227) Demko, Z. P.; Sharpless, K. B. *Angew. Chem., Int. Ed.*, **2002**, *41*, 2113-2116.

(228) Chen, J.; Burghart, A.; Derecskei-Kovacs, A.; Burgess, K. *J. Org. Chem.*, **2000**, *65*, 2900-2906.

(229) Burghart, A.; Kim, H.; Welch, M. B.; Thoresen, L. H.; Reibenspies, J.; Burgess, K.; Bergstroem, F.; Johansson, L. B. A. *J. Org. Chem.*, **1999**, *64*, 7813-7819.

(230) Umezawa, K.; Nakamura, Y.; Makino, H.; Citterio, D.; Suzuki, K. *J. Am. Chem. Soc.*, **2008**, *130*, 1550-1551.

(231) Kralovicova, E.; Krutosikova, A.; Kovac, J.; Dandarova, M. *Collect. Czech. Chem. Commun.*, **1986**, *51*, 106-111.

(232) Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Org. Lett.*, **2004**, *6*, 2853-2855.

(233) Bayer, E. A.; Wilchek, M. *Trends Biochem. Sci. (Pers. Ed.)*, **1978**, *3*, N257-N259.

(234) Adamczyk, M.; Moore, J. A.; Shreder, K. *Org. Lett.*, **2001**, *3*, 1797-1800.

(235) Dong, D.; Zheng, D.; Wang, F.-Q.; Yang, X.-Q.; Wang, N.; Li, Y.-G.; Guo, L.-H.; Cheng, J. *Anal. Chem.*, **2004**, *76*, 499-501.

(236) Green, N. M. *Methods Enzymol.*, **1970**, *18*, 418-424.

(237) Loudet, A.; Han, J.; Barhoumi, R.; Pellois, J.-P.; Burghardt, R. C.; Burgess, K. *Org. Biomol. Chem.*, **2008**, *6*, 4516-4522.

(238) Ulrich, G.; Goeb, S.; De Nicola, A.; Retailleau, P.; Ziessel, R. *Synlett*, **2007**, 1517-1520.

(239) Goodman, S. N.; Jacobsen, E. N. *Adv. Synth. Catal.*, **2002**, *344*, 953-956.

(240) Silverstein, R.; Ryskiewicz, E.; Willard, C. *Org. Synth.*, **1963**, *4*, 831-833.

APPENDIX A
GENERAL EXPERIMENTAL PROCEDURES

General Experimental Methods

All reactions were carried out under an atmosphere of dry nitrogen. Glasswares were oven-dried prior to use. Unless otherwise indicated, common reagents or materials were obtained from commercial source and used without further purification. All the solvents were used after appropriate distillation or purification.

Flash column chromatography was performed using silica gel 60 (230-400 mesh). Analytical thin layer chromatography (TLC) was carried out on Merck silica gel plates with QF-254 indicator and visualized by UV. IR spectra were recorded on a Bruker Tensor 27 spectrometer. Calculations were run with the Gaussian 03 program. TD-DFT was used to calculate the UV absorbance spectra.

^{1}H and ^{13}C spectra were recorded on a Varian 300 (300 MHz ^{1}H; 75 MHz ^{13}C) or Varian 500 (500 MHz ^{1}H; 125 MHz ^{13}C) spectrometer at room temperature. Chemical shifts were reported in ppm relative to the residual $CDCl_3$ (δ 7.24 ppm ^{1}H; δ 77.0 ppm ^{13}C) CD_3OD (δ 3.31 ppm ^{1}H; δ 49.0 ppm ^{13}C) or d^6-DMSO (δ 2.49 ppm ^{1}H; δ 39.5 ppm ^{13}C). ^{19}F NMR were acquired on a Varian 300 (300 MHz ^{1}H; 282 MHz ^{19}F) spectrometer. $CFCl_3$ was used as an external reference for the ^{19}F NMR spectra. The gated ^{1}H-decoupling technique was used to generate proton-coupled ^{13}C NMR. Coupling constants (*J*) were reported in Hertz.

Photophysical Properties and Determination of Quantum Yields

Absorption spectra were recorded on a Varian 100 Bio UV-Vis spectrophotometer at room temperature. Measurements for the molar extinction coefficients were performed according to the following protocol. 2-3 mg of dye was weighed using a digital scale ($\Delta w = 0.1$ mg), and dissolved into 10 mL of appropriate solvent. A number of further diluted solutions with different dye concentrations (*ca*: 10^{-6}-10^{-5} M) were prepared from this stock solution. The absorption spectra of these diluted solutions were measured, and the absorbance (A) and the concentration (C) were plotted on a graph of A vs C. The slope of the above line was the molar extinction coefficient according to the Beer-Lambert law ($A = \varepsilon \bullet C \bullet L$).

Steady-state fluorescence spectroscopic studies were performed on a Cary Eclipse fluorometer at room temperature. The slit width was 5 nm for both excitation and emission. The relative quantum yields of the samples were obtained by comparing the area under the corrected emission spectrum of the test sample with that of a standard. The quantum efficiencies of fluorescence were obtained from three measurements with the following equation:

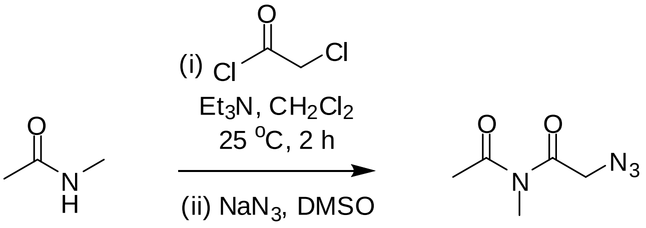
$$\boldsymbol{\Phi_x = \Phi_{st} (I_x/I_{st}) (A_{st}/A_x) (\eta_x^2/\eta_{st}^2)}$$

where $\boldsymbol{\Phi_{st}}$ is the reported quantum yield of the standard, **I** is the area under the emission spectra, **A** is the absorbance at the excitation wavelength and η is the refractive index of the solvent used, measured on a pocket refractometer from ATAGO. The **X** subscript denotes the test sample, and **st** denotes standard.

APPENDIX B

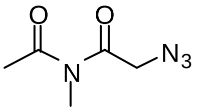
EXPERIMENTAL DATA FOR CHAPTER II

Synthesis of compound 2a



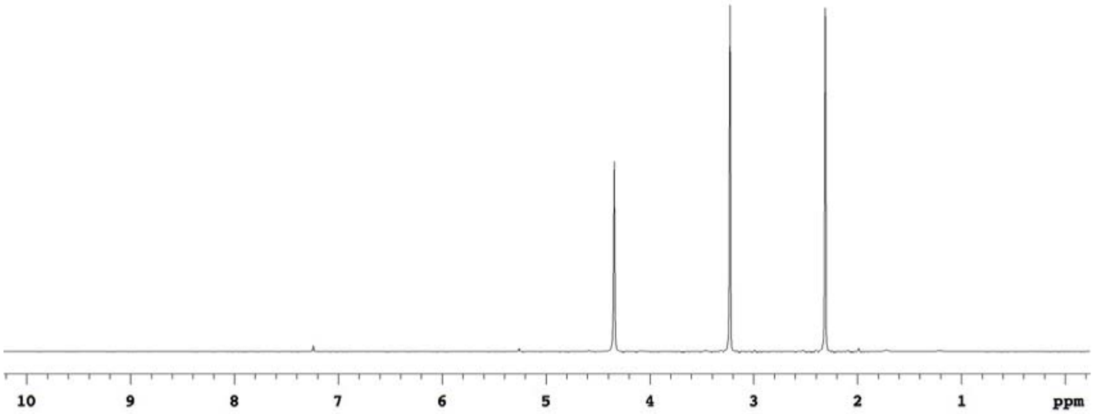
2a

2-Chloroacetyl chloride (8.8 mL, 0.11 mol) was added dropwise to the solution of N-methylacetamide (7.3 g, 0.1 mol) and Et_3N (15.2 mL, 0.11 mol) in 100 mL CH_2Cl_2. The solution was stirred at 25°C for 2h, then washed with water (3 x 30 mL) and brine (1 x 30 mL). The organic layer was dried over Na_2SO_4 and evaporated under reduced pressure. The residue was dissolved in 100 mL DMSO and NaN_3 (13 g, 0.2 mol) was added to the resulting solution in one portion. The reaction mixture was stirred at 25 °C for 12 h and then poured into ice water and extracted with EtOAc (3 x 50 mL). The combined extracts were washed with H_2O (3 x 30 mL) and brine (1 x 30 mL). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by flash chromatography eluting with 1:4 EtOAc/hexane to afford **2a** (3.0 g, 19 %) as a yellow oil. ^1H NMR (300 MHz, $CDCl_3$) δ 4.34 (s, 2H), 3.23 (s, 3H), 2.31 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.1, 171.1, 55.5, 31.9, 25.4. MS (ESI) m/z calcd for $(M+H)^+$ $C_5H_9N_4O_2$ 157.07; found 157.07.

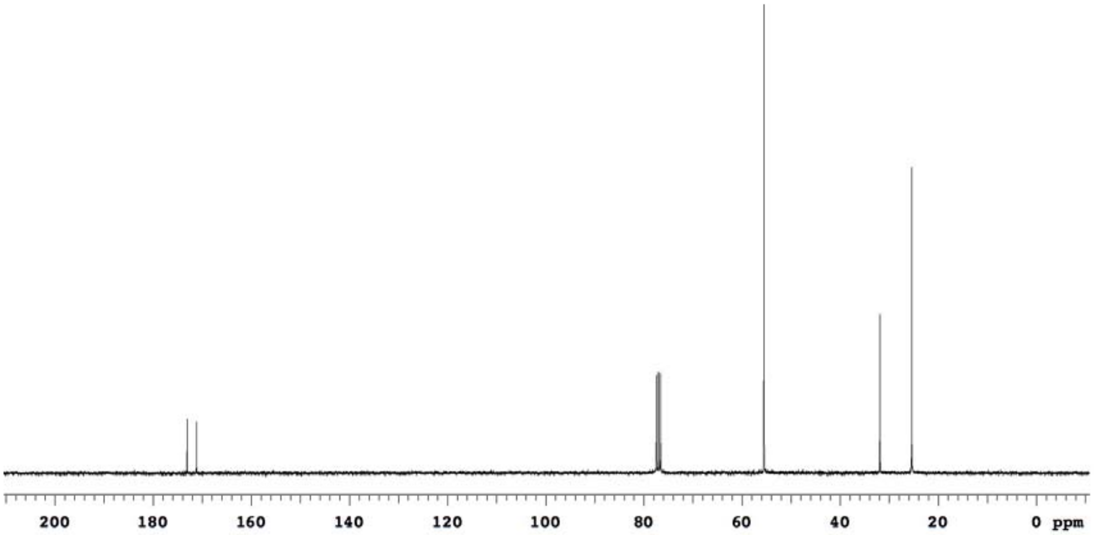


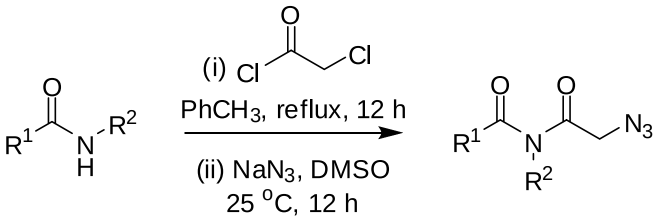
2a

^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

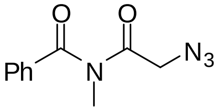




2b - d

General procedure for the synthesis of α-azido substituted imides 2b-d[90]

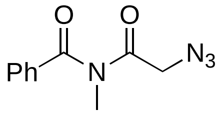
2-Chloroacetyl chloride (1.1 eq.) was added dropwise to a suspension of the amide (1.0 eq.) in toluene (0.3 M) at 25 °C. The mixture was heated to reflux for 12 h and then evaporated to dryness under reduced pressure. DMSO was added to the residue to give a solution (0.5 M), to which was added NaN_3 (2 eq.) in one portion. After stirring at 25 °C for 12 h, the mixture was poured into water and extracted with EtOAc (3 x 30 mL). The combined extracts were washed with H_2O (3 x 20 mL) and brine (1 x 20 mL). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by flash chromatography to afford **2**.



2b

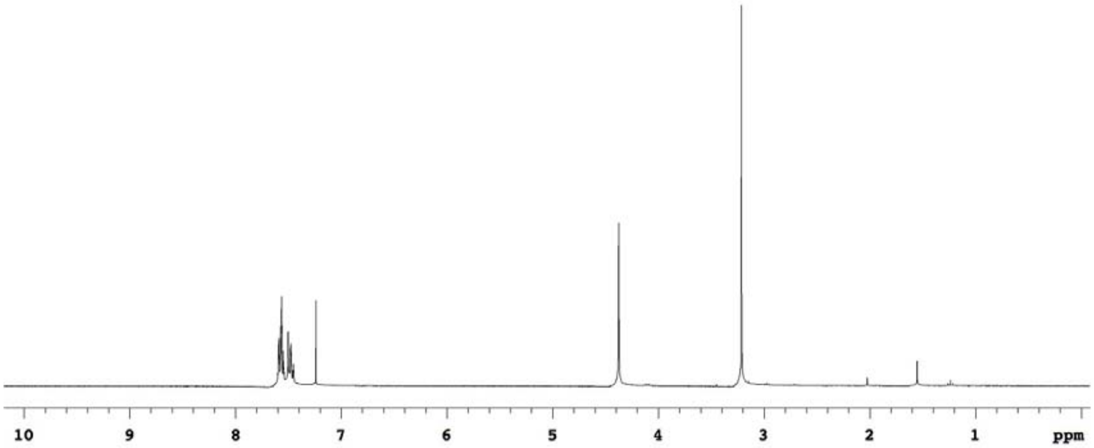
N-(2-Azidoacetyl)-*N*-methylbenzamide (2b)

Prepared according to literature procedure,[90] as a white solid (3.66 g, 78 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.60-7.45 (m, 5H), 4.38 (s, 2H), 3.21 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.6, 171.9, 133.9, 132.6, 128.9, 128.3, 54.5, 35.2.

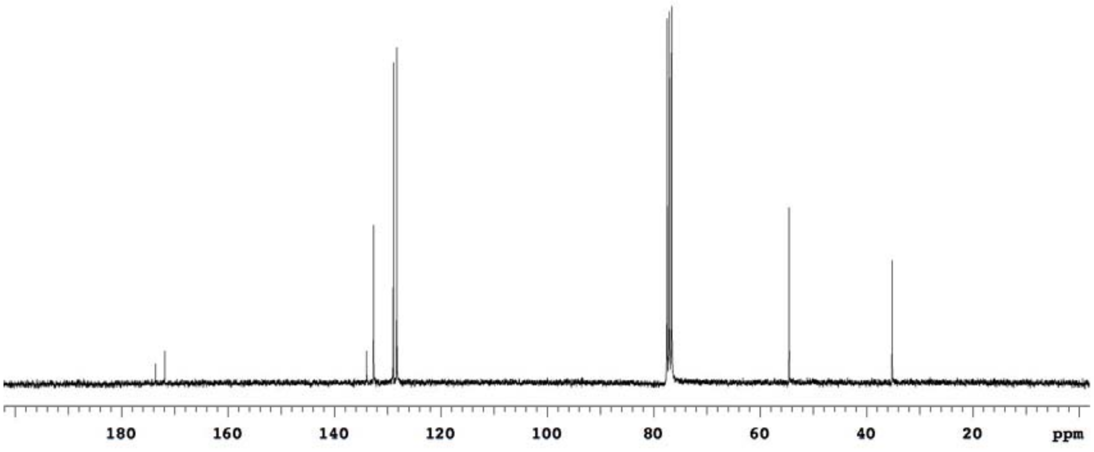


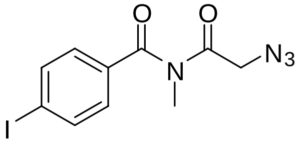
2b

^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

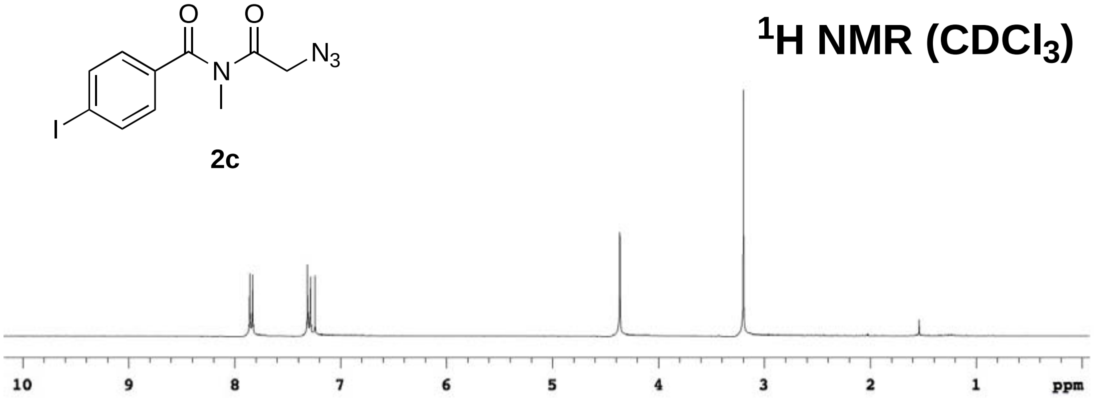


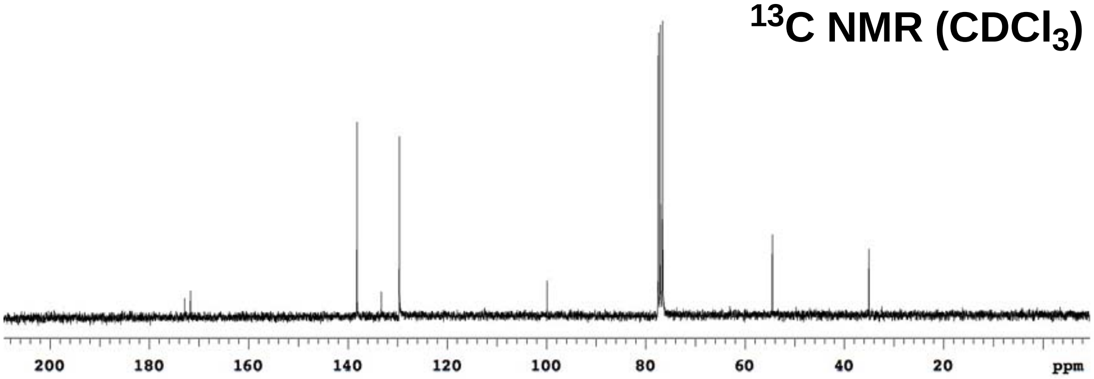


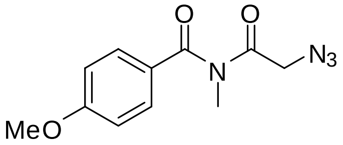
2c

N-(2-Azidoacetyl)-4-iodo-N-methylbenzamide (2c)

Purified by flash chromatography (4:1 hexanes/EtOAc) to afford **2c** (1.48 g, 66 %) as a white solid. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.84 (d, 2H, *J* = 8.6 Hz), 7.30 (d, 2H, *J* = 8.6 Hz), 4.36 (s, 2H), 3.20 (s, 3H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 172.9, 171.7, 138.2, 133.3, 129.6, 99.9, 54.5, 35.1.



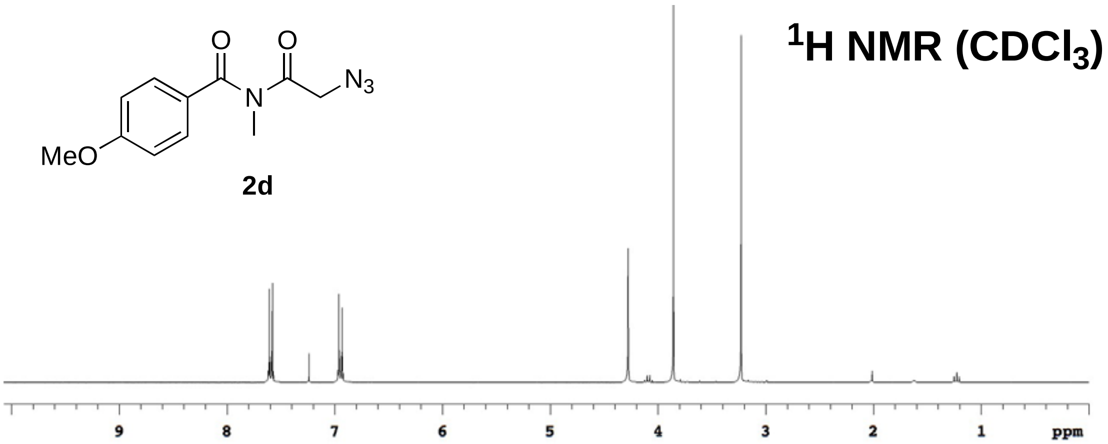


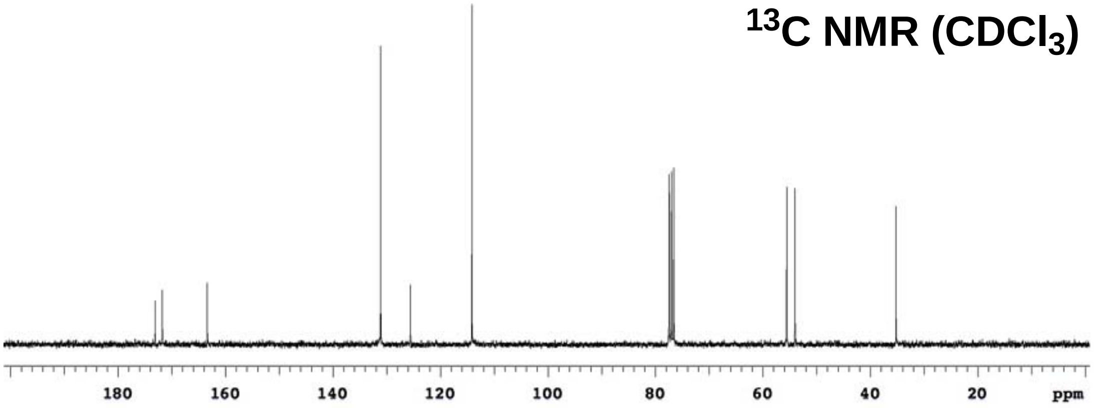


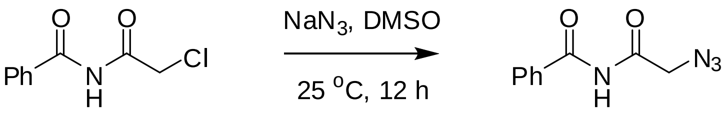
2d

N-(2-Azidoacetyl)-4-methoxy-*N*-methylbenzamide (2d)

Purified by flash chromatography (9:1 hexanes/EtOAc) to afford **2d** (1.71 g, 69 %) as a yellow oil. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.59 (d, 2H, *J* = 9.0 Hz), 6.95 (d, 2H, *J* = 9.0 Hz), 4.28 (s, 2H), 3.86 (s, 3H), 3.23(s, 3H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 173.1, 171.7, 163.4, 131.1, 125.5, 114.1, 55.5, 54.0, 35.2. MS (ESI) m/z calcd for $(M+H)^+$ $C_{11}H_{13}N_4O_3$ 249.10; found 249.10.



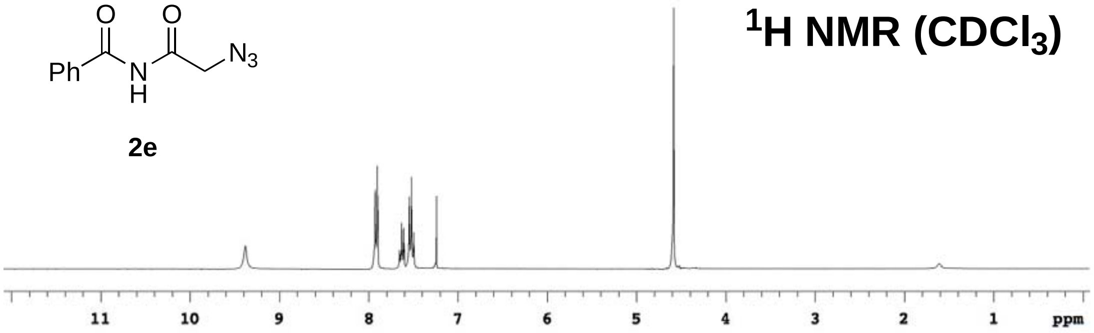


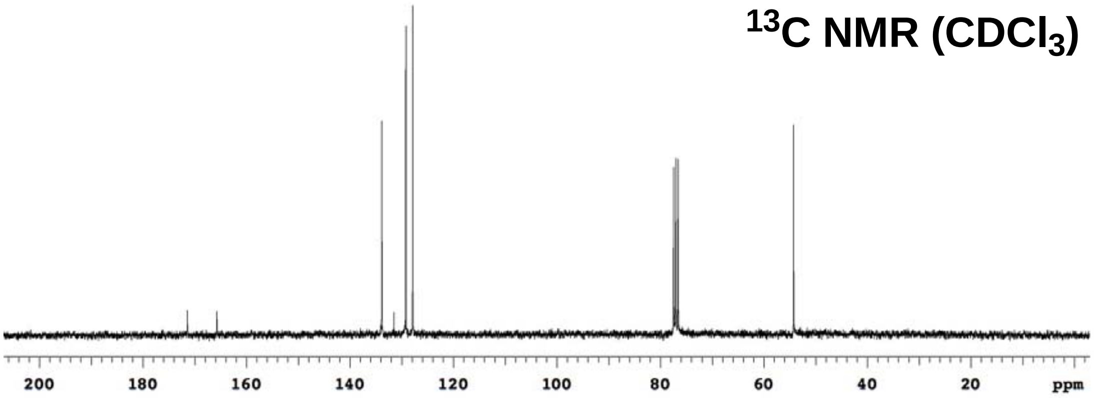


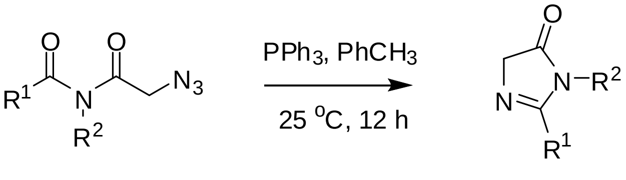
2e

N-(2-Azidoacetyl)benzamide (2e)

NaN_3 (4.1 g, 62.6 mmol) was added to the solution of *N*-(2-chloroethanoyl)benzamide (6.2 g, 31.3 mmol) (prepared as a white solid with 77 % yield, according to literature procedure.[239]) in 30 mL DMSO. The mixture was stirred at 25 °C for 12 h and then poured into ice water. The resulting precipitate was filtered and washed with diethyl ether to afford **2e** (5.9 g, 92 %) as a white solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.38 (br, 1H), 7.93-7.90 (m, 2H), 7.65-7.50 (m, 3H), 4.58 (s, 2H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.4, 165.7, 133.8, 131.5, 129.2, 127.9, 54.3.



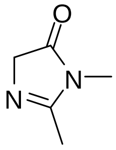




2 **3**

General procedure for the synthesis of imidazolinones 3[90]

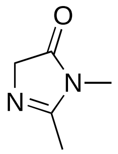
Triphenylphosphine (1.1 eq.) was added to a solution of imide derivatives **2** (1.0 eq.) in toluene (0.2 M). The mixture was stirred at 25 °C for 12 h and then evaporated under reduced pressure. The residue was purified by flash chromatography to afford **3**.



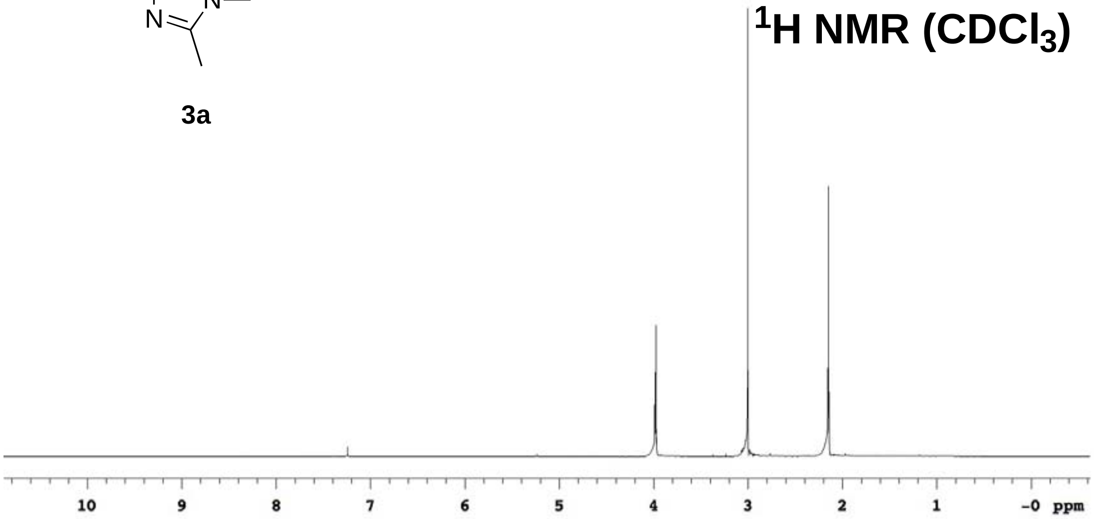
3a

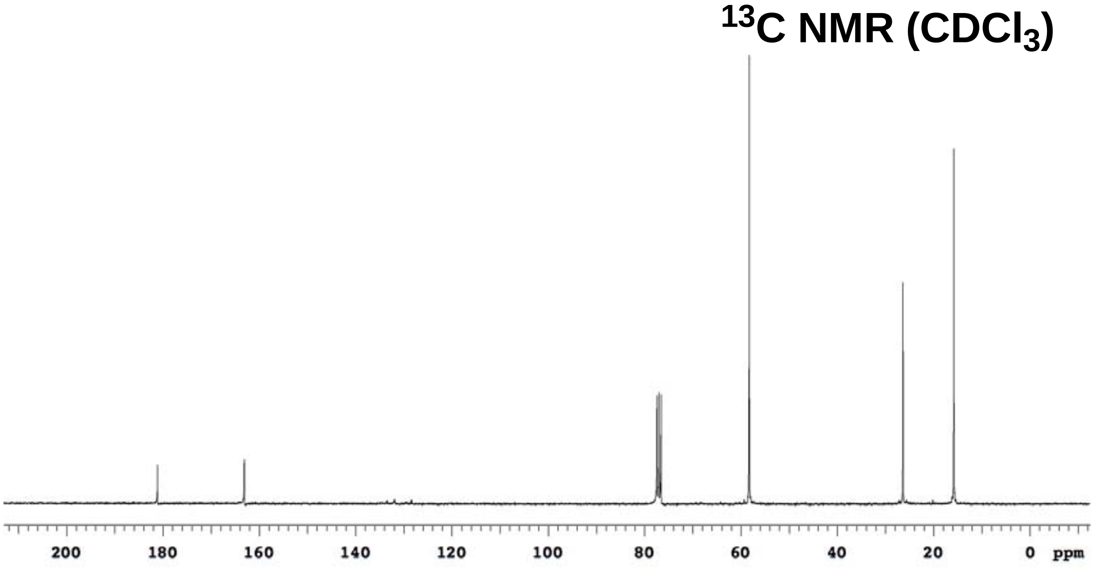
1,2-Dimethyl-1*H*-imidazol-5(4*H*)-one (3a)

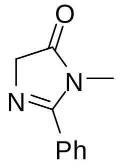
Purified by flash chromatography (1:19 MeOH/EtOAc) to afford **3a** (2.5 g, 87 %) as a yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 3.98 (q, 2H, *J* = 2.4 Hz), 3.00 (s, 3H), 2.15 (t, 3H, *J* = 2.4 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 181.1, 163.1, 58.3, 26.3, 15.8. MS (ESI) m/z calcd for $(M+H)^+$ $C_5H_9N_2O$ 113.07; found 113.07.


O
N
N

3a


^{1}H NMR ($CDCl_3$)
10
9
8
7
6
5
4
3
2
1
-0
ppm

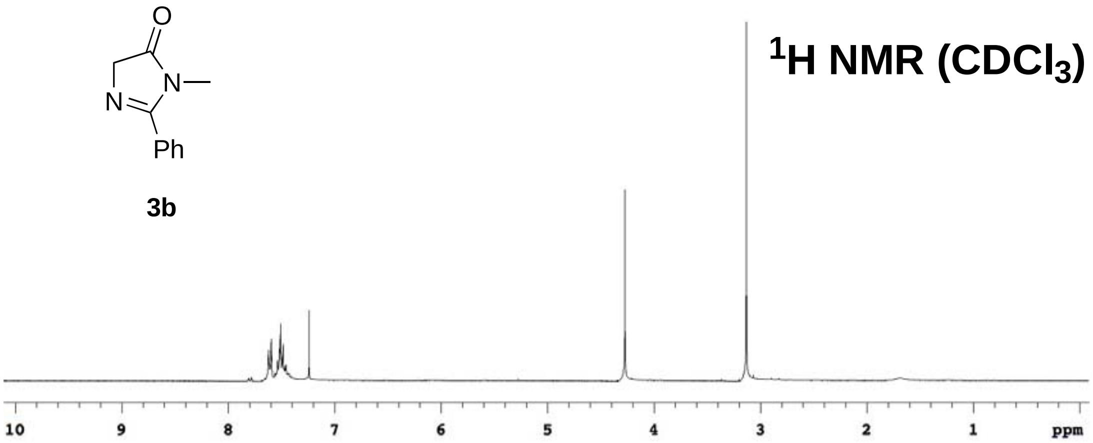

^{13}C NMR ($CDCl_3$)
200
180
160
140
120
100
80
60
40
20
0
ppm

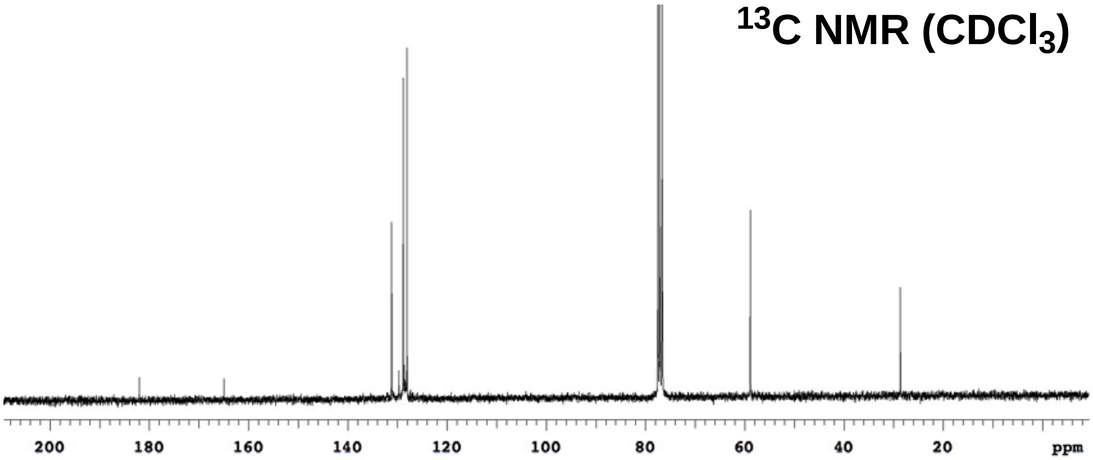


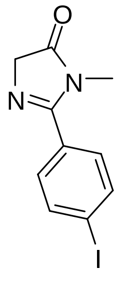
3b

1-Methyl-2-phenyl-1*H*-imidazol-5(4*H*)-one (3b)

Prepared according to literature procedure,[90] as a red solid (1.0 g, 89 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.63-7.48 (m, 5H), 4.27 (2, 2H), 3.13 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 182.0, 164.9, 131.2, 129.7, 128.8, 128.0, 58.9, 28.6.



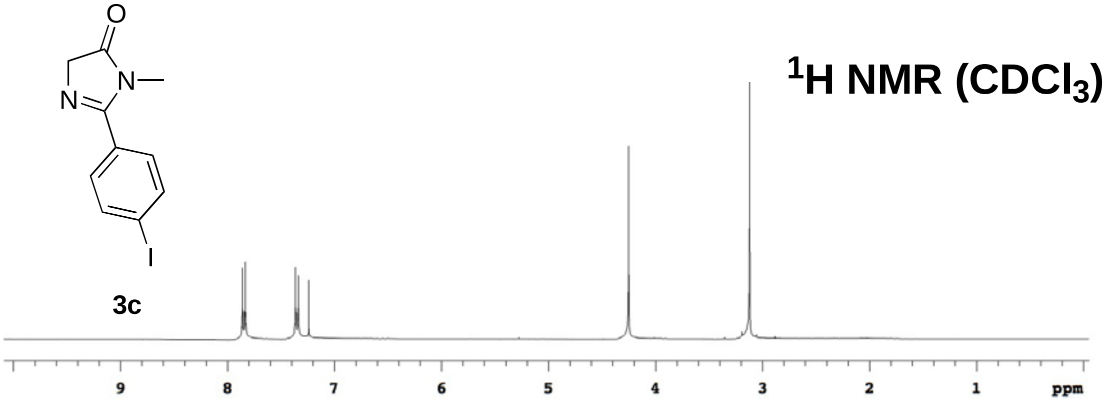


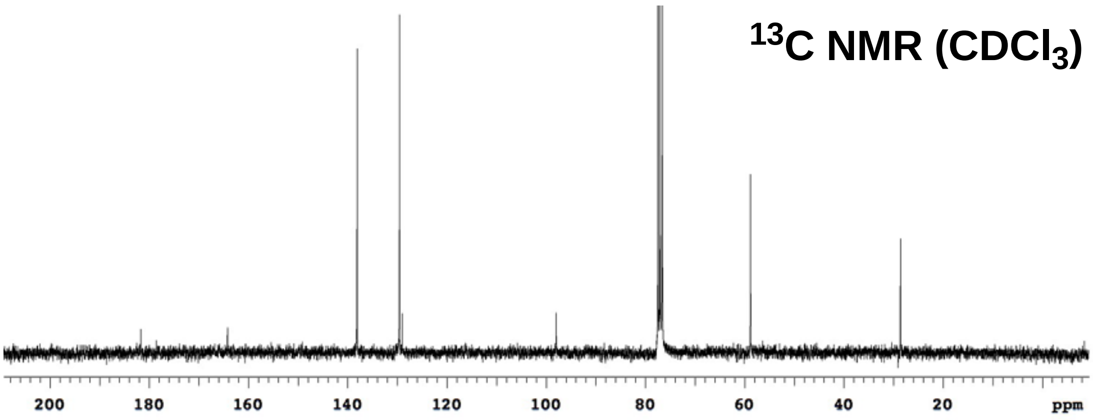


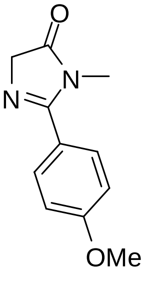
3c

2-(4-iodophenyl)-1-methyl-1*H*-imidazol-5(4*H*)-one (3c)

Purified by flash chromatography (9:1 CH_2Cl_2/EtOAc) to afford **3c** (1.1 g, 83 %) as a yellow solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.85 (d, 2H, *J* = 8.5 Hz), 7.35 (d, 2H, *J* = 8.5 Hz), 4.25 (s, 2H), 3.12 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 181.7, 164.2, 138.1, 129.6, 129.0, 98.0, 58.8, 28.6. MS (ESI) m/z calcd for $(M+H)^+$ $C_{10}H_{10}IN_2O$ 300.98; found 300.98.



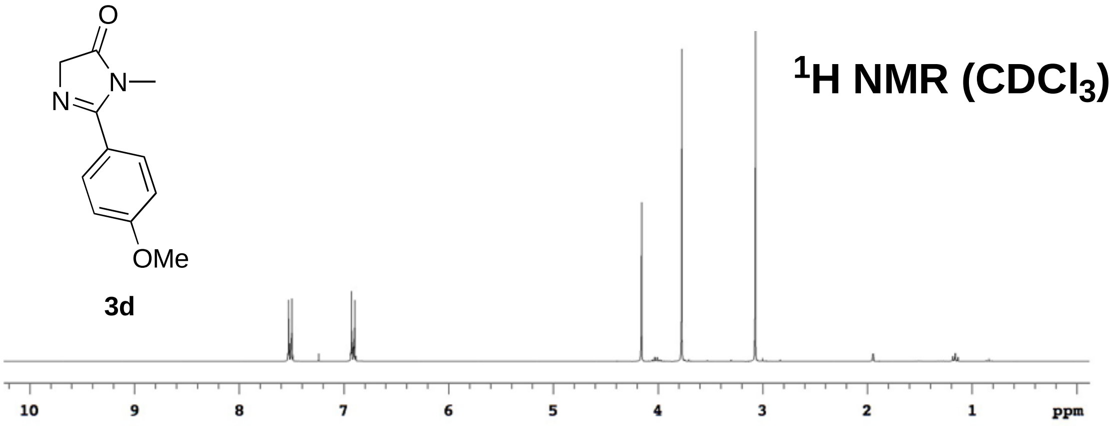


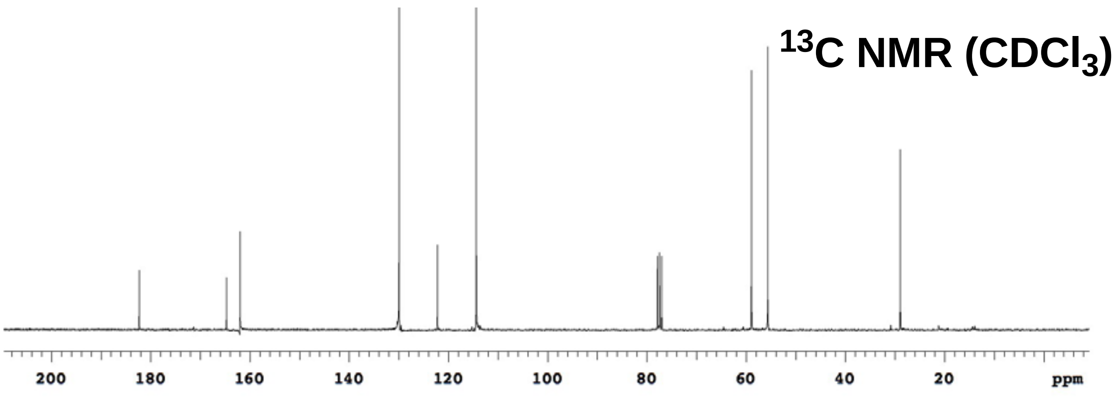


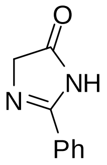
3d

2-(4-Methoxyphenyl)-1-methyl-1*H*-imidazol-5(4*H*)-one (3d)

Purified by flash chromatography (3:2 CH_2Cl_2/EtOAc) to afford **3d** (890 mg, 65 %) as a red oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.51 (d, 2H, *J* = 9.0 Hz), 6.91 (d, 2H, *J* = 9.0 Hz), 4.16 (s, 2H), 3.77 (s, 3H), 3.07 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 181.9, 164.3, 161.6, 129.5, 121.8, 114.0, 58.5, 55.2, 28.5. MS (ESI) m/z calcd for $(M+H)^+$ $C_{11}H_{13}N_2O_2$ 205.10; found 205.09.



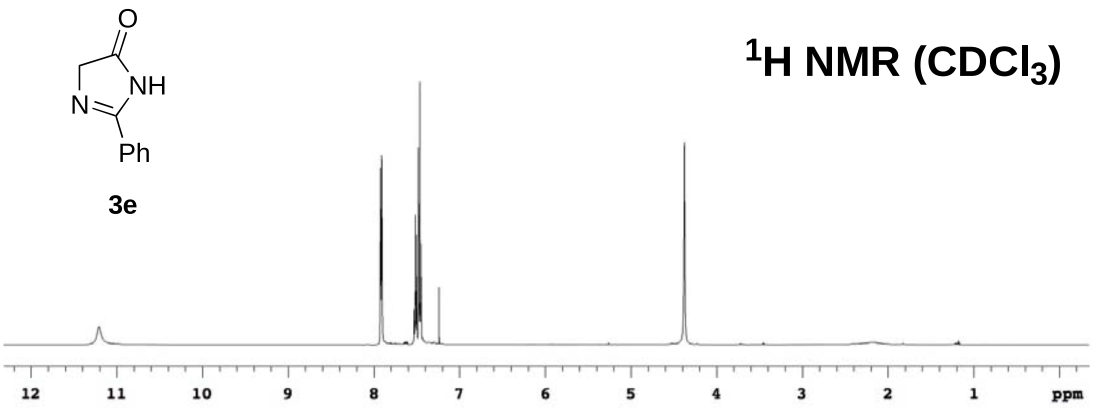


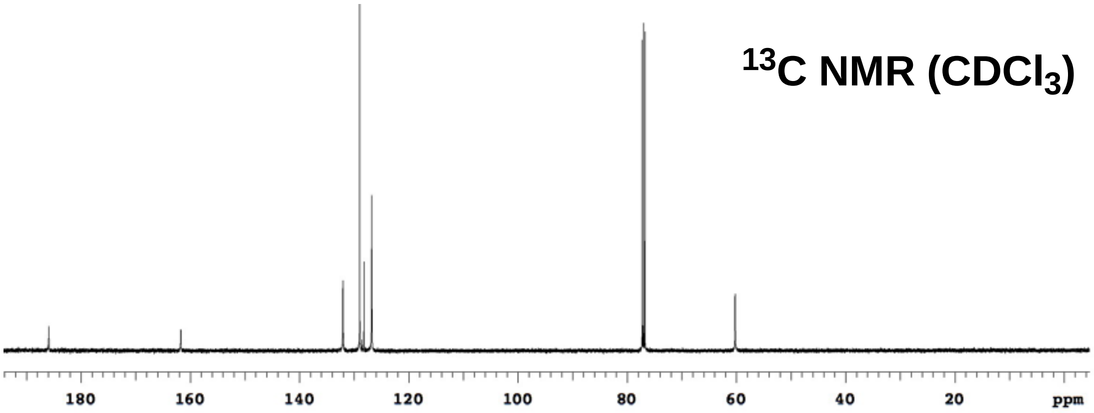


3e

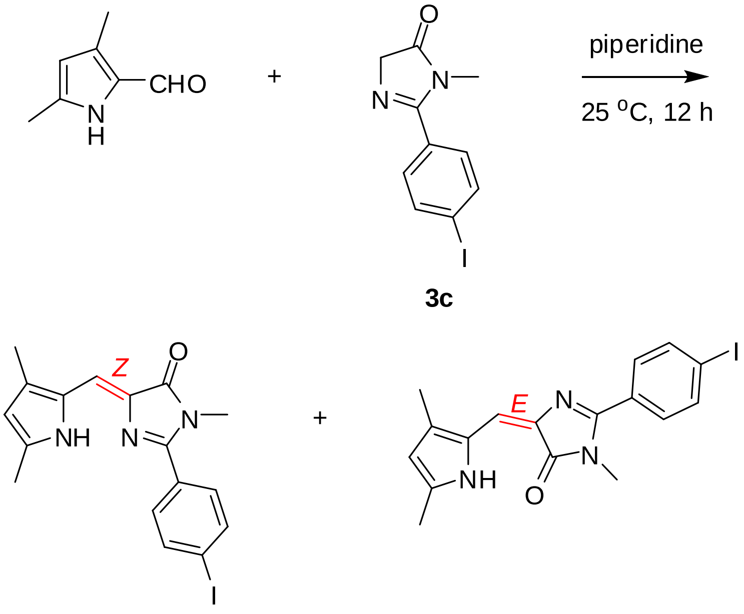
2-Phenyl-1*H*-imidazol-5(4*H*)-one (3e)

The crude product was washed with diethyl ether to afford pure product **3e** (1.5 g, 55 %) as a red solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.2 (br, 1H), 7.92-7.91 (m, 2H), 7.53-7.45 (m, 3H), 4.38 (s, 2H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 185.9, 161.7, 132.0, 129.0, 128.2, 126.8, 60.2. MS (ESI) m/z calcd for $(M+H)^+$ $C_9H_9N_2O$ 161.07; found 161.07.





Synthesis of compounds *Z* - 4c and *E* - 4c

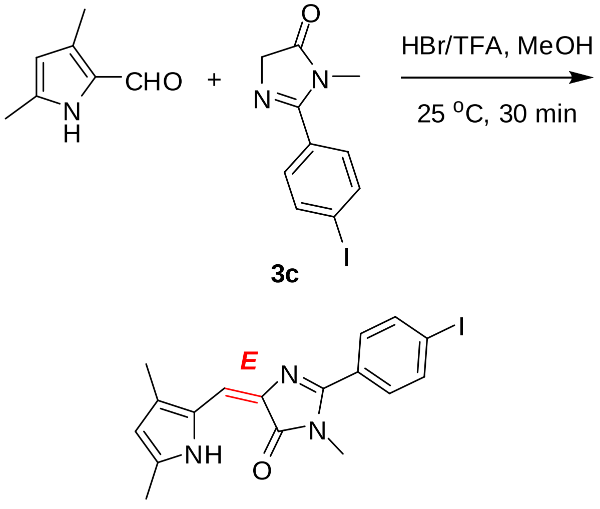


3c

***Z* - 4c**

***E* - 4c**

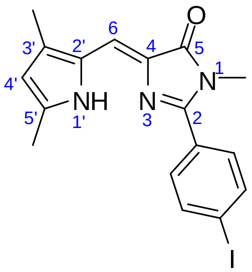
A solution of 3,5-dimethyl-1H-pyrrole-2-carbaldehyde (1.0 eq.) and imidazolinone **3c** (1.1 eq.) in piperidine (0.2 M) was stirred at room temperature for 12 h. The reaction mixture was diluted with CH_2Cl_2, washed with water, brine and dried over Na_2SO_4. The solvents were removed under reduced pressure. The *E* and *Z* isomer can be separated by flash chromatography (1:19 to 1:9 EtOAc/Hexanes) and further purified by recrystallization from EtOAc/Hexanes.



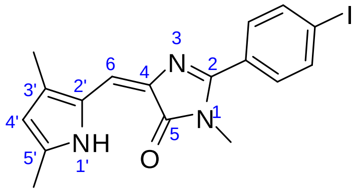
3c

***E*-4c**

3,5-Dimethyl-1H-pyrrole-2-carbaldehyde (25 mg, 0.2 mmol) and imidazolinone **3c** (60 mg, 0.2 mmol) were dissolved in 1 mL MeOH then 2 mL TFA and 0.2 mL HBr were added. The solution was stirred at 25 °C for 30 min and a red precipitate formed. The mixture was filtered and the solid was washed with $NaHCO_{3(aq.)}$ then taken into EtOAc. The organic layer was washed with $NaHCO_{3(aq.)}$, water and dried over Na_2SO_4. The solvent was removed under reduced pressure to afford the pure product ***E*-4c** as a red solid (80 mg, *quant.*).



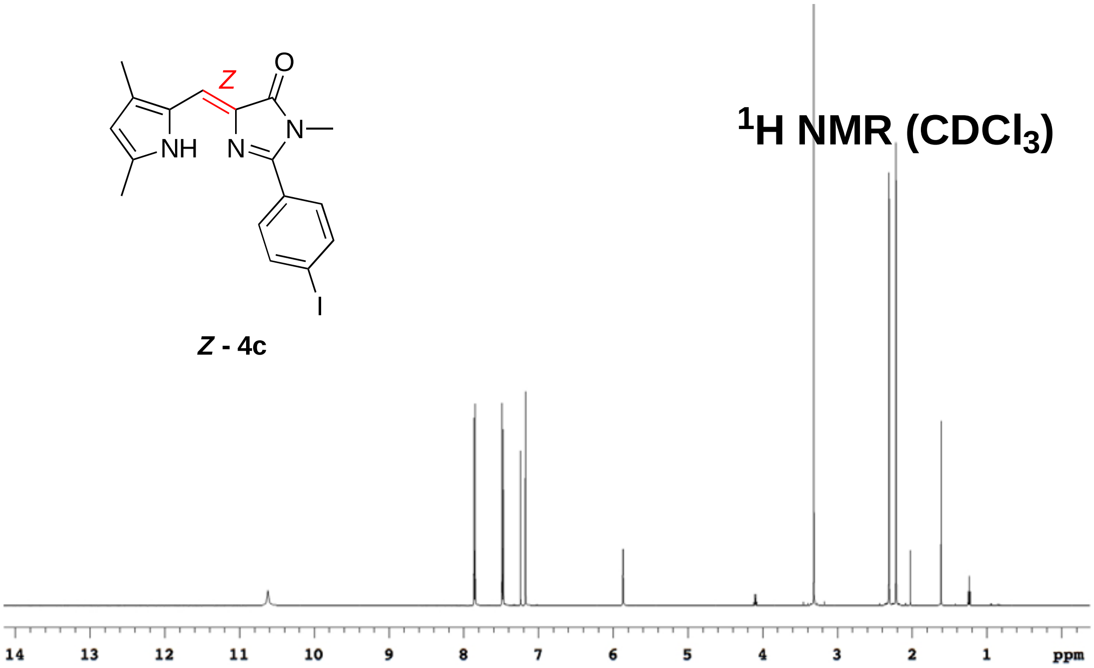
***Z* -4c**

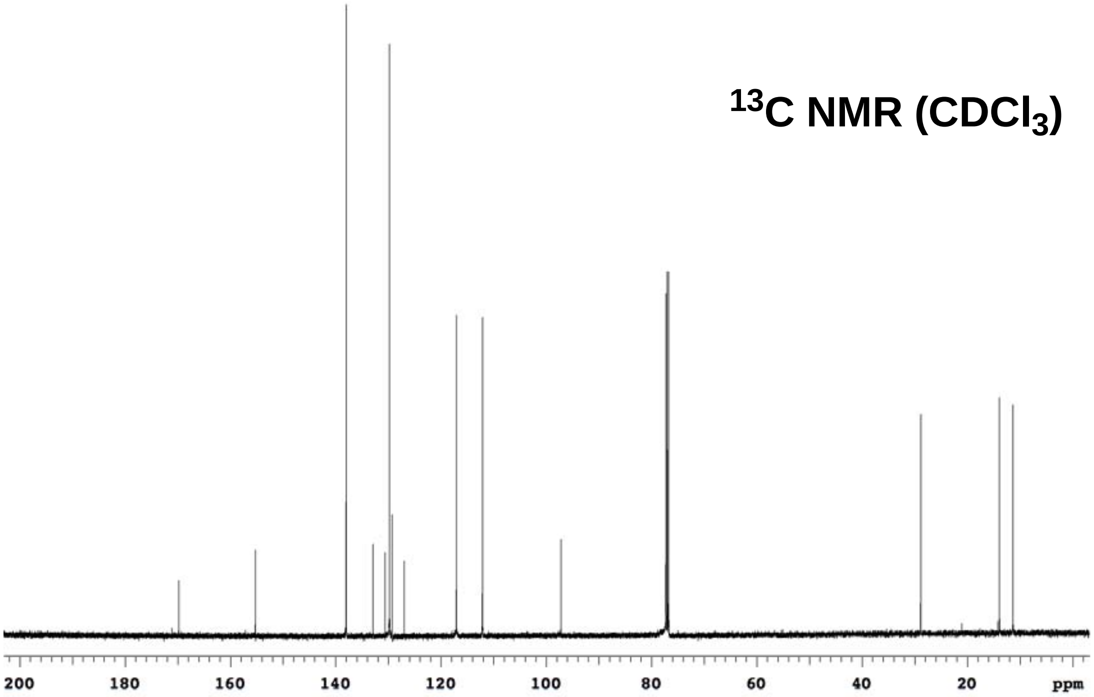


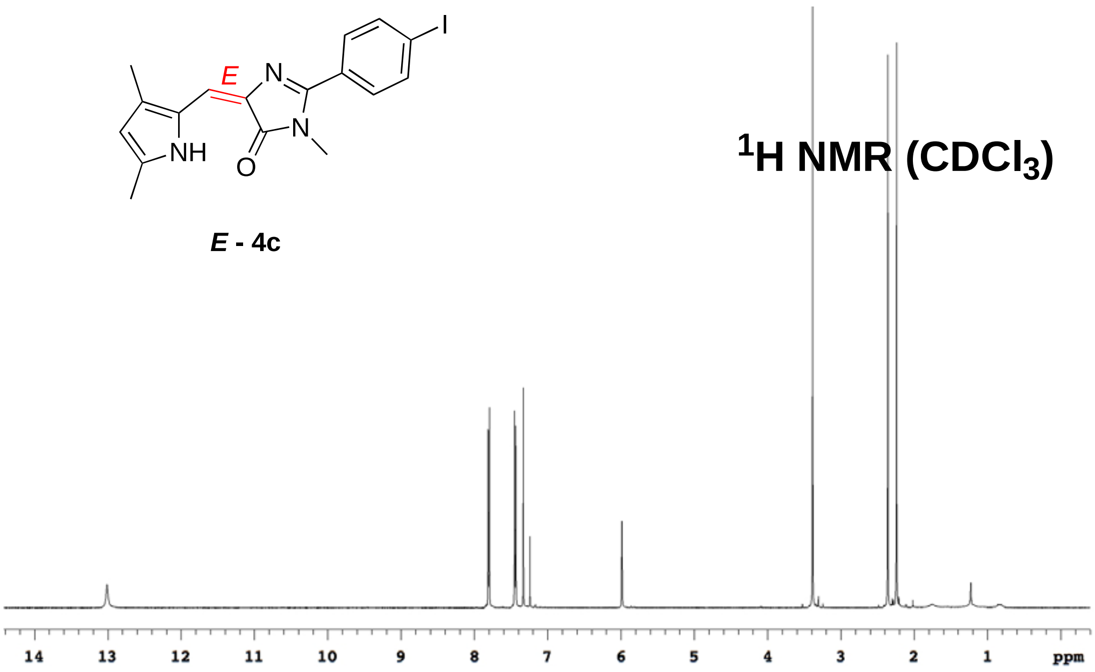
***E*-4c**

***Z* – 4c**: ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.62 (br, 1H), 7.85 (d, 2H, *J* = 8.6 Hz), 7.48 (d, 2H, *J* = 8.6 Hz), 7.17 (s, 1H), 5.87 (d, 1H, *J* = 2.7 Hz), 3.32 (s, 3H), 2.31 (s, 3H), 2.21 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 169.8, 155.3, 138.1, 138.0, 132.9, 130.6, 129.8, 129.3, 127.0, 117.1, 112.1, 97.2, 28.9, 13.9, 11.3. From proton-coupled ^{13}C NMR: $^3J_{(H6-C5)}$ = 3.8 Hz. IR (neat) $\nu_{C=O}$ 1686 cm^{-1}; $\lambda_{max\ abs}$ 458 nm (ε = 40196 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 515 nm, Φ = 0.0005 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{17}IN_3O$ 406.04; found 406.04.

***E* – 4c:** ^{1}H NMR (500 MHz, $CDCl_3$) δ 13.01 (br, 1H), 7.80 (d, 2H, *J* = 8.5 Hz), 7.44 (d, 2H, *J* = 8.5 Hz), 7.33 (s, 1H), 5.99 (d, 1H, *J* = 2.2 Hz), 3.39 (s, 3H), 2.36 (s, 3H), 2.24 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.6, 151.7, 138.2, 137.9, 134.4, 129.5, 129.3, 129.2, 127.2, 125.6, 113.5, 96.6, 29.1, 14.0, 11.5. From proton-coupled ^{13}C NMR: $^3J_{(H6-C5)}$ = 8.9 Hz. IR (neat) $\nu_{C=O}$ 1663 cm^{-1}; $\lambda_{max\ abs}$ 458 nm (ε = 48034 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 515 nm, Φ = 0.0003 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{17}IN_3O$ 406.04; found 406.04.

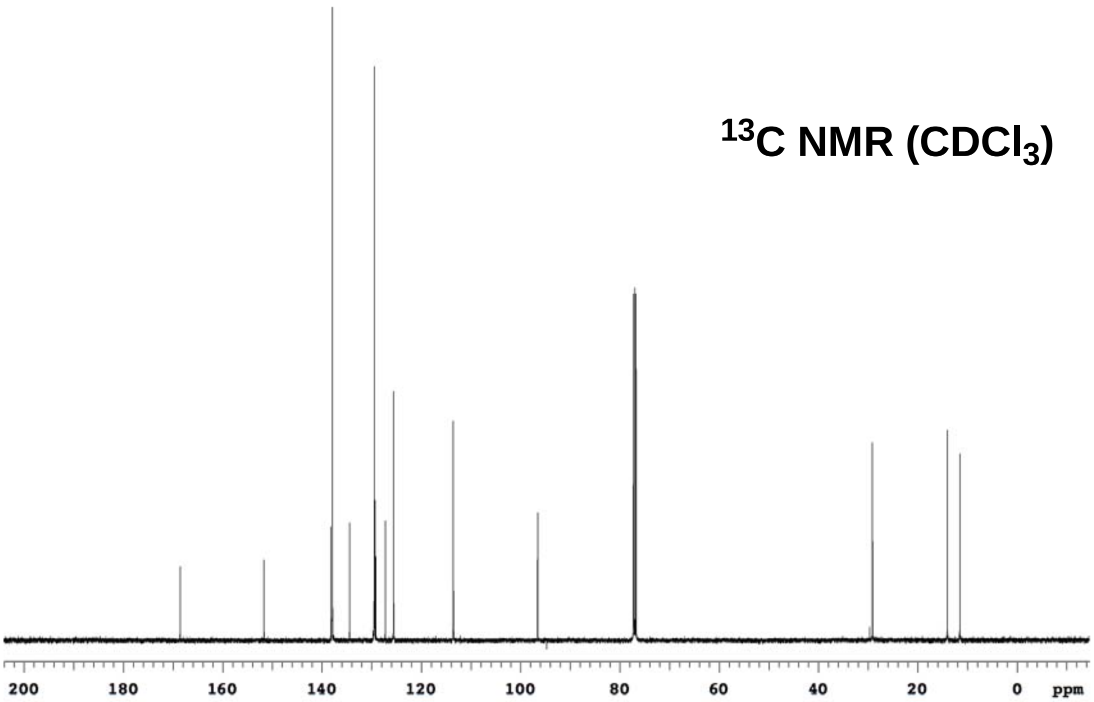

O
N
NH
N
I
Z
Z - 4c
^{1}H NMR (CDCl$_3$)
14
13
12
11
10
9
8
7
6
5
4
3
2
1
ppm


^{13}C NMR (CDCl$_3$)
200
180
160
140
120
100
80
60
40
20
ppm



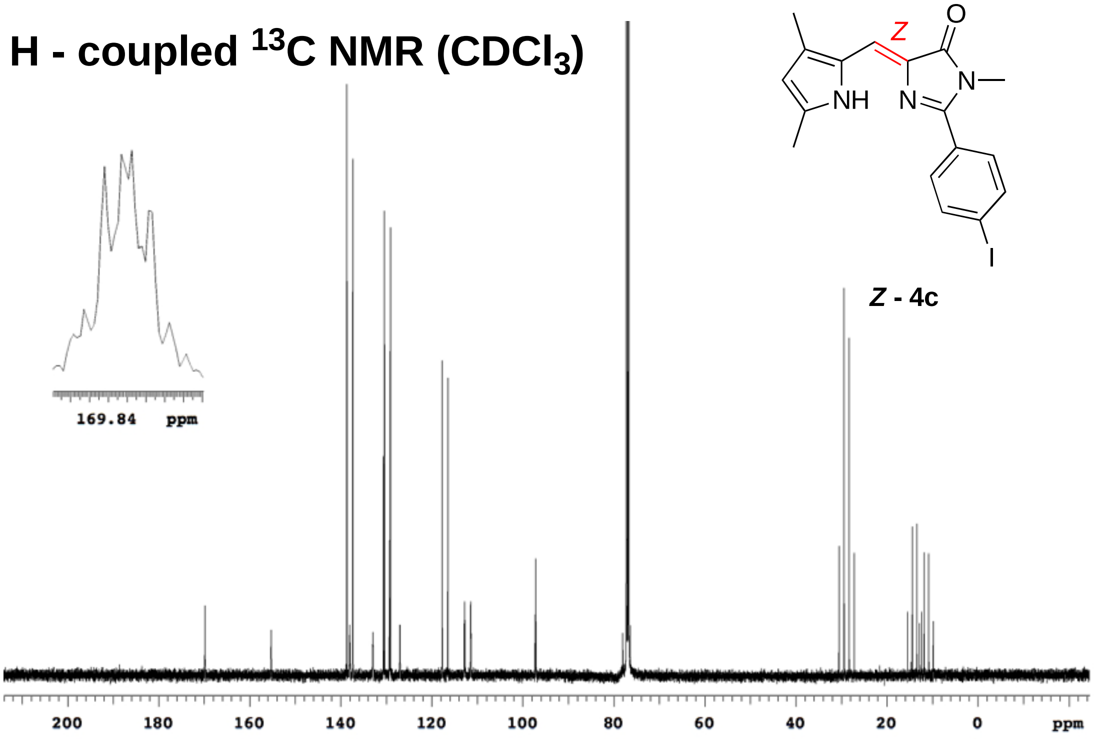
***E* - 4c**

E
N
N
NH
O
I
^{1}H NMR (CDCl$_3$)
14
13
12
11
10
9
8
7
6
5
4
3
2
1
ppm



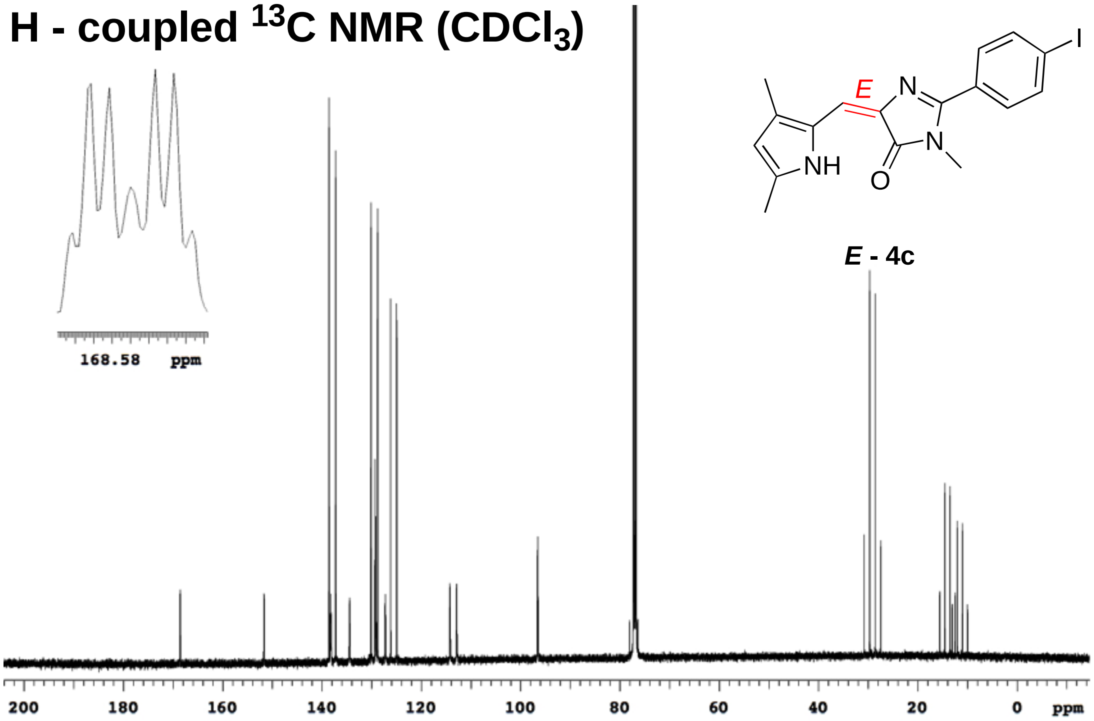
^{13}C NMR (CDCl$_3$)
200
180
160
140
120
100
80
60
40
20
0
ppm

H - coupled ^{13}C NMR ($CDCl_3$)



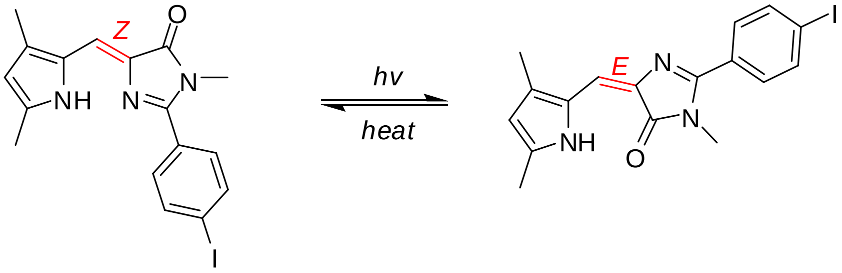
***Z* - 4c**

H - coupled ^{13}C NMR ($CDCl_3$)



***E* - 4c**

Photo and thermal isomerization

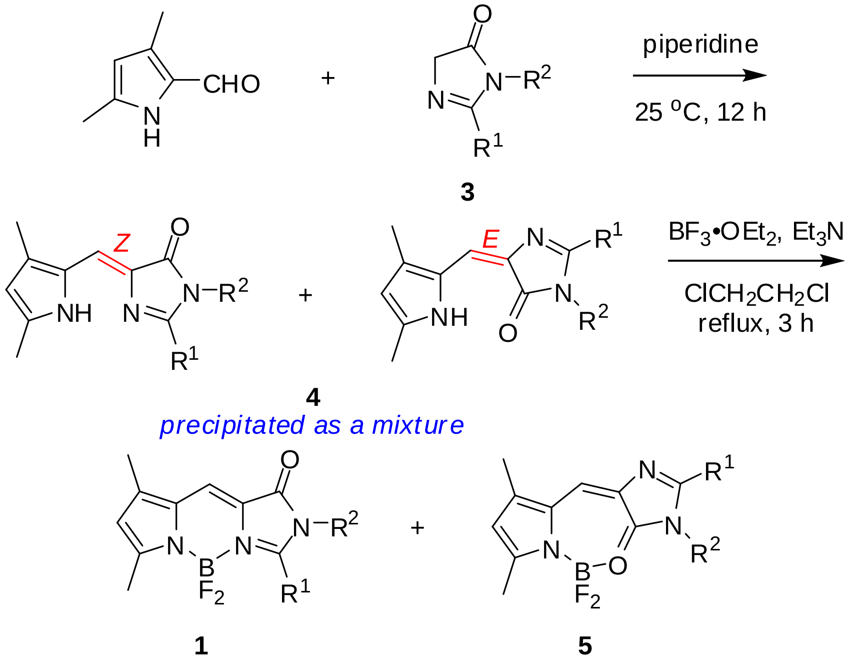

Z
hv
heat
E
NH
N
O
I

***Z* - 4c** ***E* - 4c**

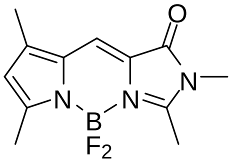
A solution of 15 mg of the mixture of ***E* – 4c** and ***Z* – 4c** in 1 mL of $CDCl_3$ was irradiated at 360 nm at room temperature. NMR was taken periodically.

A solution of 15 mg of the mixture of ***E* – 4c** and ***Z* – 4c** in 1 mL of d^6-DMSO was heated to 100°C. NMR was taken periodically after cooled to room temperature.

General procedure for the synthesis of GFP-chromophore analogs 1 and 5


CHO
+
piperidine
25 °C, 12 h
3
Z
E
4
precipitated as a mixture
BF3•OEt2, Et3N
ClCH2CH2Cl
reflux, 3 h
R1
R2
1
5

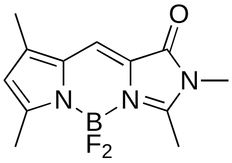
A solution of 3,5-dimethyl-1H-pyrrole-2-carbaldehyde (1.0 eq.) and imidazolinone **3** (1.1 eq.) in piperidine (0.2 M) was stirred at room temperature for 12 h. Water was added to the solution and yellow precipitate formed. The mixture was stirred for 10 min at room temperature. The precipitate was collected, washed with water and dried *in* *vacuo* to afford a yellow solid. The solid was dissolved in 1,2-dichloroethane (0.2 M) then Et_3N (3 eq.) was added. The reaction mixture was stirred at 25 °C for 15 min then BF_3•OEt_2 (5 eq.) was added. The mixture was heated to reflux for 3 h and cooled to room temperature. After diluting with dichloromethane, the mixture was washed with water and brine. The organic layer was dried over Na_2SO_4 and evaporated under reduced pressure to give the crude product, which was purified by flash chromatography to afford the pure products **1** and **5**.



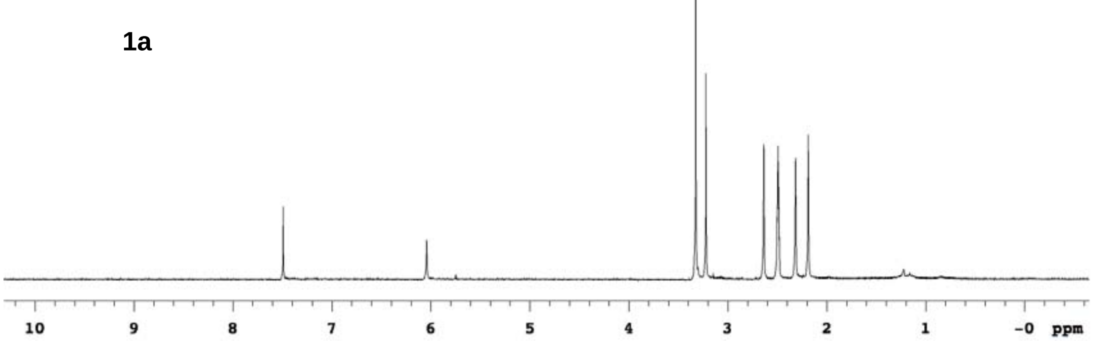
1a

GFP-chromophore analog 1a

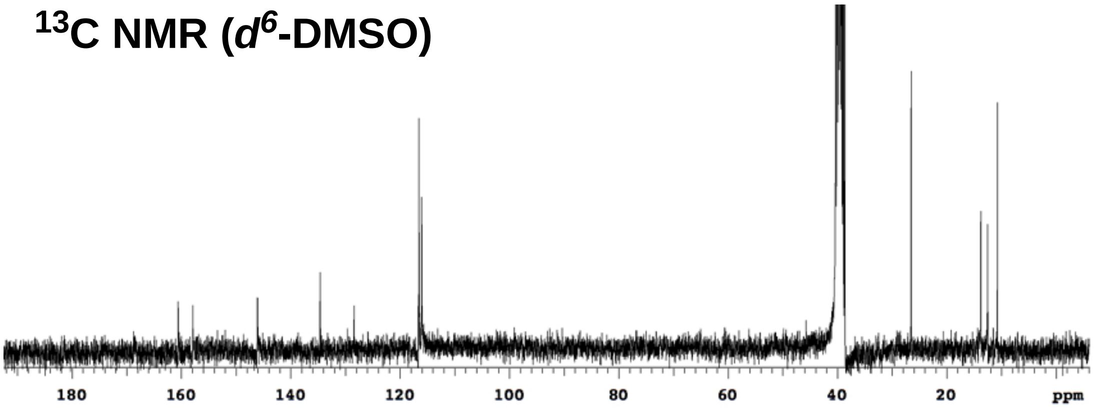
Purified by flash chromatography (3:2 EtOAc/hexanes) to afford **1a** (153 mg, 24 %) as a yellow solid. ^{1}H NMR (300 MHz, *d*^{6}-DMSO) δ 7.49 (s, 1H), 6.04 (s, 1H), 3.22 (s, 3H), 2.64 (s, 3H), 2.31 (s, 3H), 2.19 (s, 3H); ^{13}C NMR (75 MHz, *d*^{6}-DMSO) δ 160.6, 157.9, 146.1, 134.6, 128.4, 116.7, 116.5, 116.1, 26.5, 13.8, 12.5, 10.7; ^{19}F NMR (282 MHz, *d*^{6}-DMSO) δ 39.43 (q, *J* = 31.5 Hz); IR (neat) $\nu_{C=O}$ 1710 cm^{-1}; $\lambda_{max\ abs}$ 477 nm (ε = 57286 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 485 nm, Φ = 0.89 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{12}H_{15}BF_2N_3O$ 266.13; found 266.13.



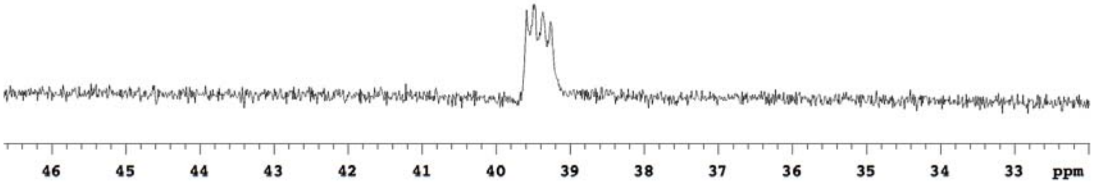
1a



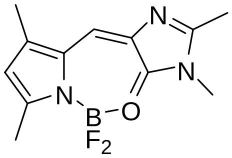
^{1}H NMR (d^6-DMSO)



^{13}C NMR (d^6-DMSO)



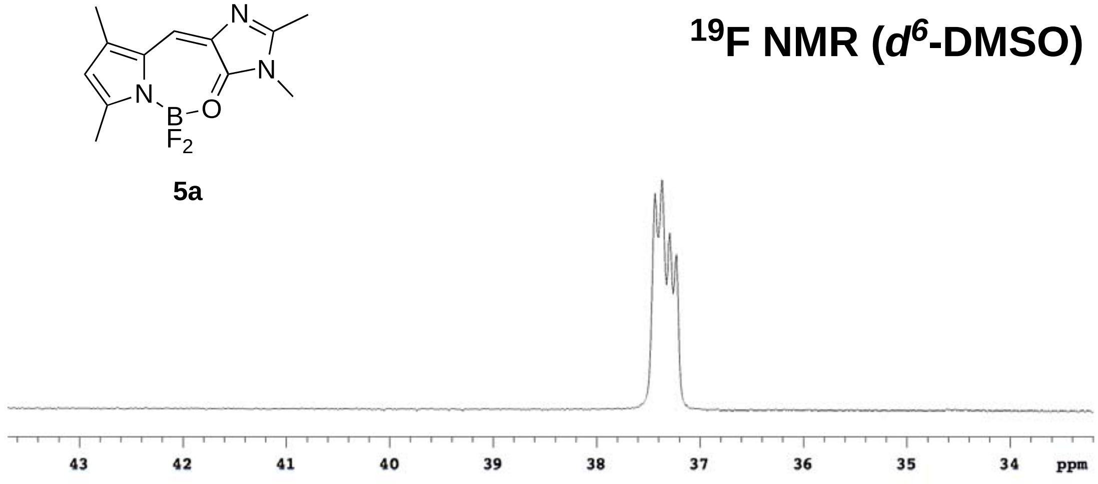
^{19}F NMR (d^6-DMSO)

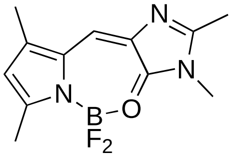


5a

GFP-chromophore analog 5a

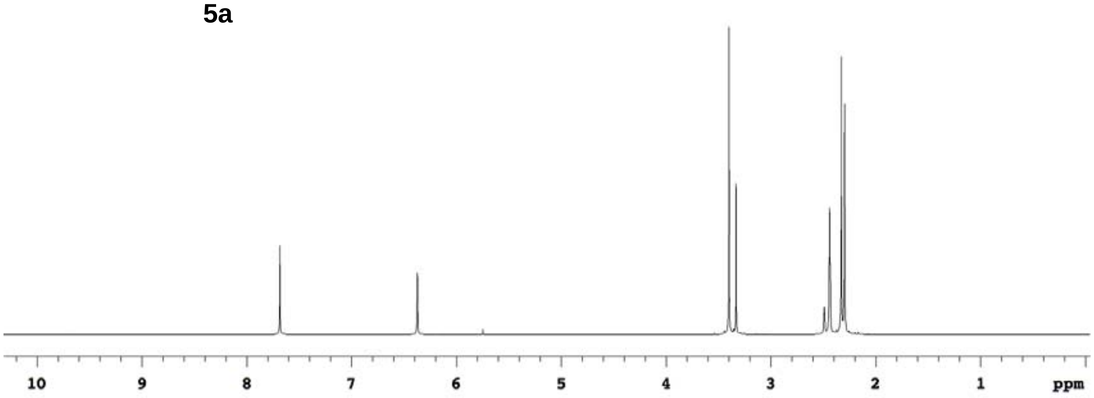
Purified by flash chromatography (3:2 EtOAc/hexanes) to afford **5a** (350 mg, 54 %) as a yellow solid. ^{1}H NMR (500 MHz, d^{6}-DMSO) δ 7.68 (s, 1H), 6.37 (s, 1H), 3.40 (s, 3H), 2.44 (t, 3H, J = 3.0 Hz), 2.33 (s, 3H), 2.30 (s, 3H); ^{13}C NMR (125 MHz, d^{6}-DMSO) δ 158.7 (t, J = 5.3 Hz), 156.4, 147.6, 145.1, 133.8, 130.9, 120.9 (2C), 28.4, 16.0 (t, J = 6.2 Hz), 13.7, 11.5; ^{19}F NMR (282 MHz, d^{6}-DMSO) δ 37.33 (q, J = 18.4 Hz); IR (neat) $\nu_{C=O}$ 1622 cm^{-1}; $\lambda_{max\ abs}$ 426 nm (ε = 37009 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 485 nm, Φ = 0.0007 in MeOH; MS (ESI) m/z calcd for $(M+H)^{+}$ $C_{12}H_{15}BF_{2}N_{3}O$ 266.13; found 266.13.



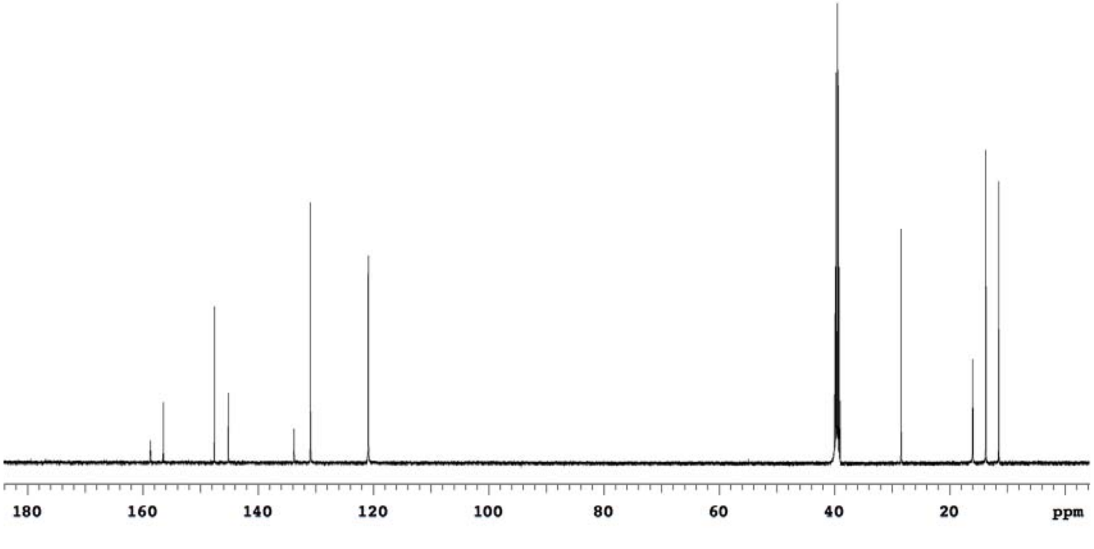


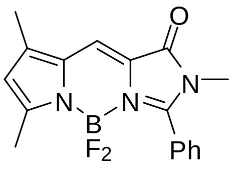
5a

^{1}H NMR (d^{6}-DMSO)



^{13}C NMR (d^{6}-DMSO)

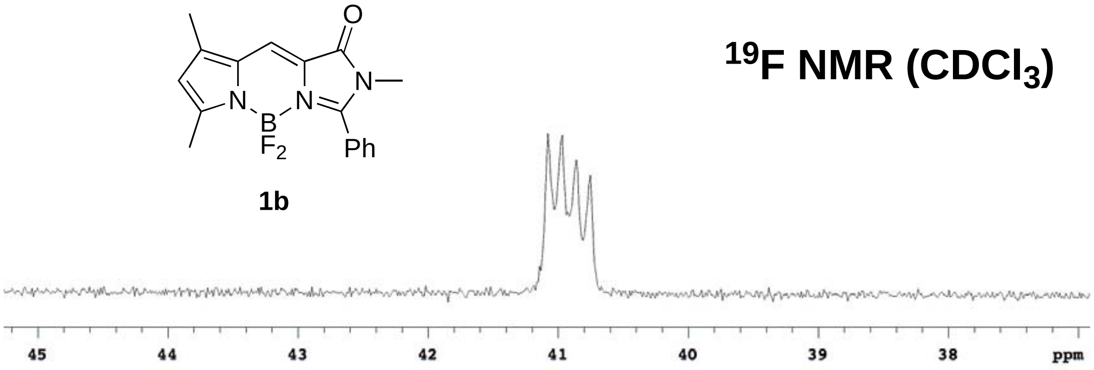


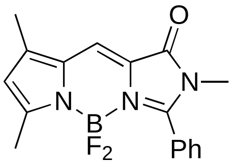


1b

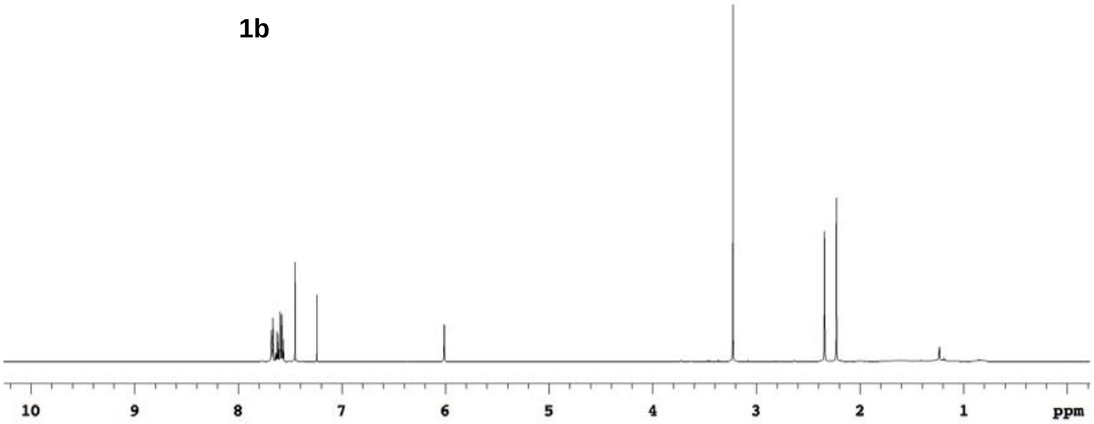
GFP-chromophore analog 1b

Purified by flash chromatography (1:9 EtOAc/hexanes) to afford **1b** (152 mg, 22 %) as a red solid. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.68-7.56 (m, 5H), 7.45 (s, 1H), 6.01 (s, 1H), 3.22 (s, 3H), 2.34 (s, 3H), 2.23 (s, 3H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 161.0, 153.0, 151.6, 138.4, 132.1, 130.4, 129.4, 128.8, 124.4, 118.6, 117.7, 116.4, 28.3, 14.4, 11.2; ^{19}F NMR (282 MHz, CDCl$_3$) δ 40.92 (q, *J* = 30.6 Hz); IR (neat) $\nu_{C=O}$ 1687 cm^{-1}; $\lambda_{max\ abs}$ 490 nm (ε = 58684 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 521 nm, Φ = 0.87 in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ $C_{17}H_{17}BF_2N_3O$ 328.14; found 328.14.

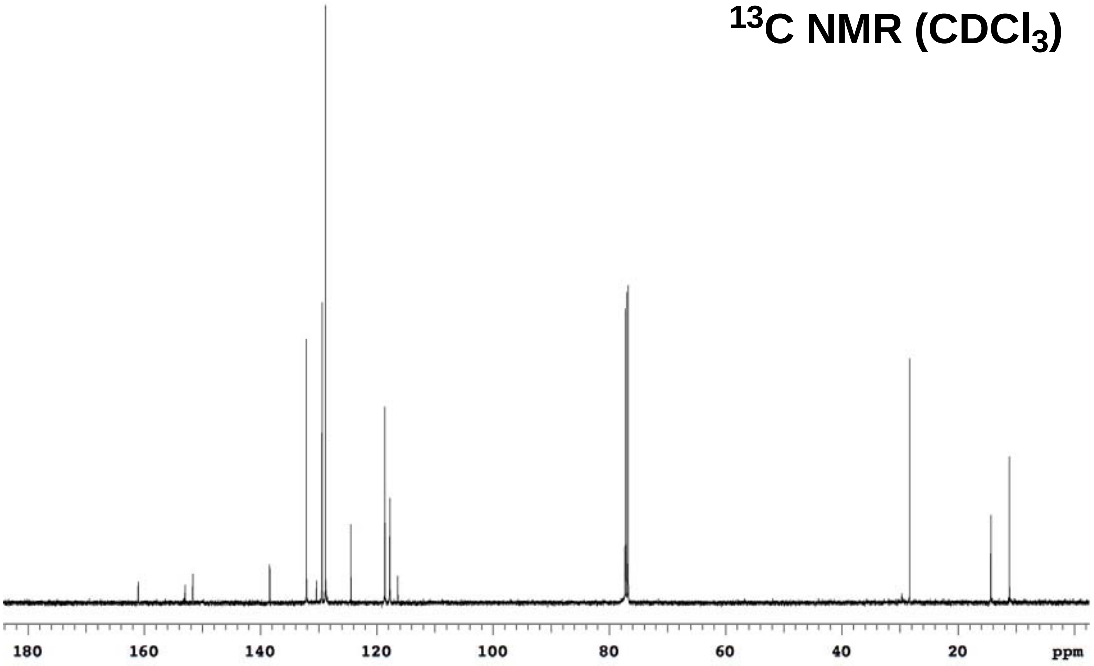




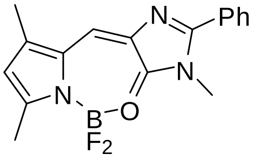
1b



^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

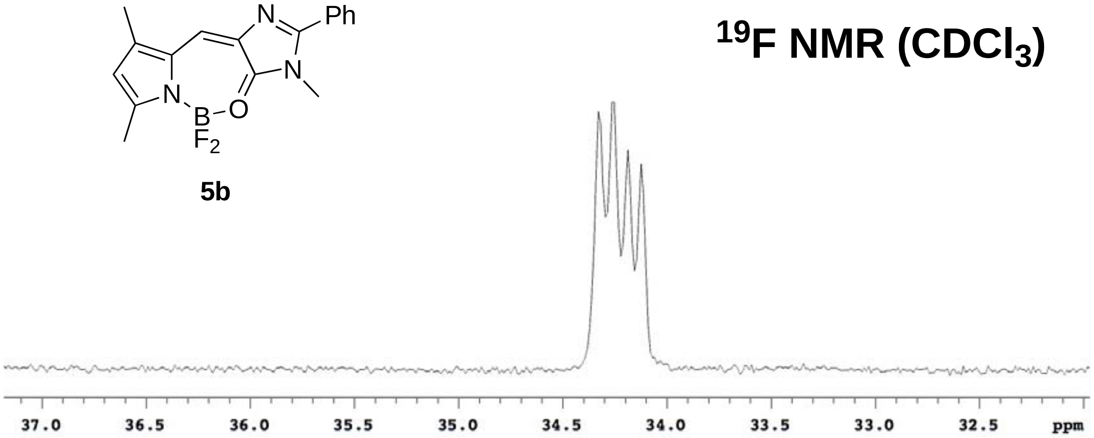


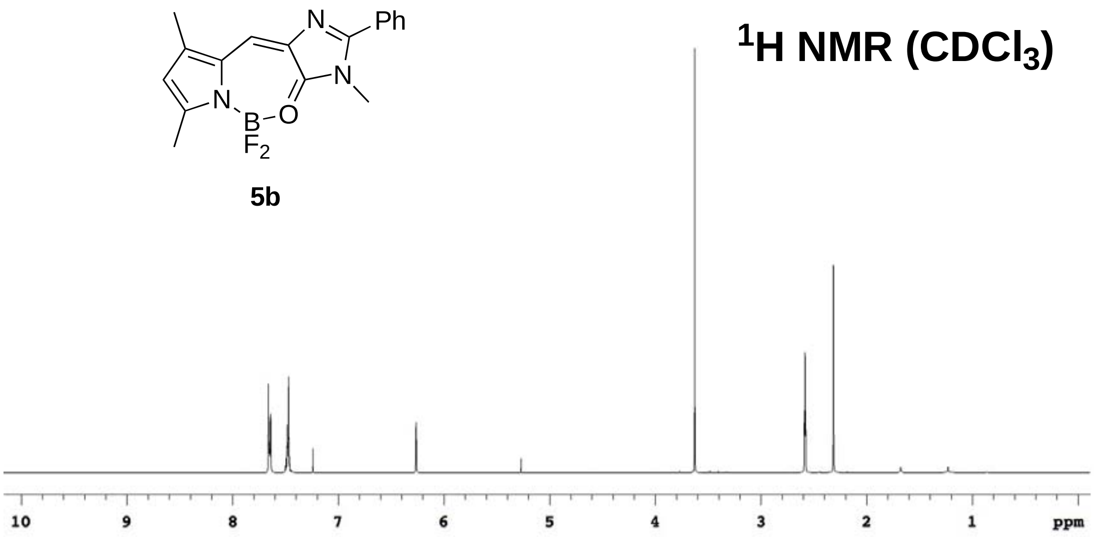
5b

GFP-chromophore analog 5b

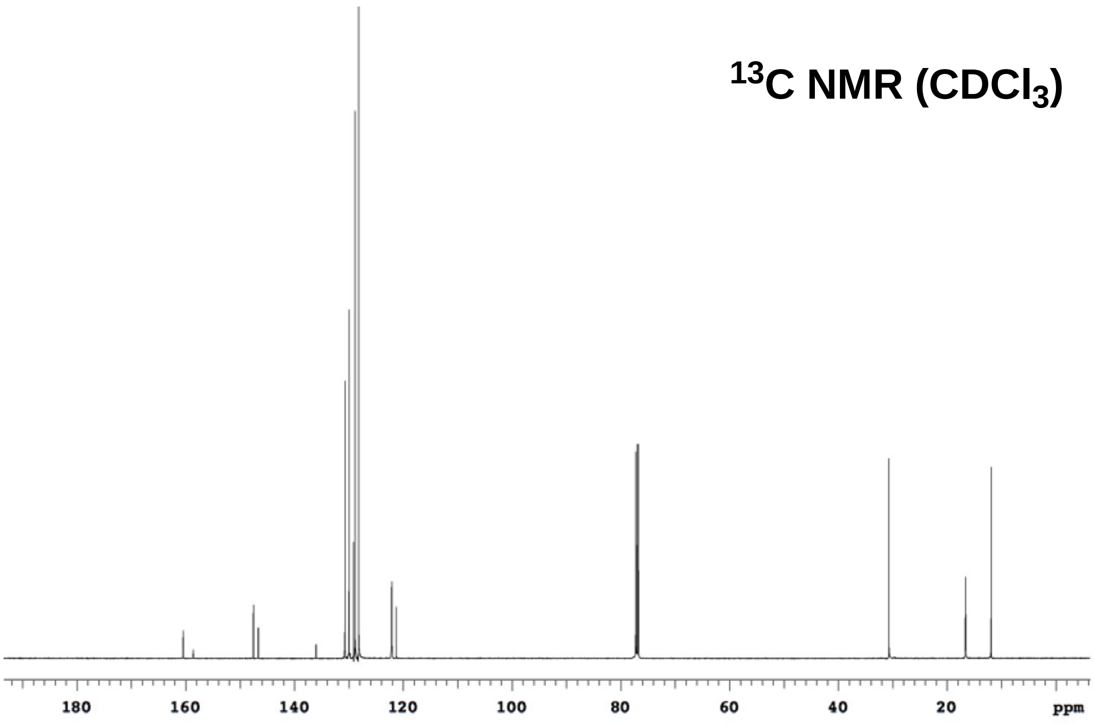
Purified by flash chromatography (1:9 EtOAc/hexanes) to afford **5b** (430 mg, 63 %) as a red solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.66-7.64 (m, 3H), 7.50-7.46 (m, 3H), 6.26 (s, 1H), 3.63 (s, 3H), 2.58 (t, 3H, *J* = 3.0 Hz), 2.31 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 160.5, 158.6 (t, *J* = 5.3 Hz), 147.5, 146.6, 136.0, 130.7, 130.0, 129.1, 128.9, 128.2, 122.1, 121.3, 30.7, 16.6 (t, *J* = 6.4 Hz), 11.9; ^{19}F NMR (282 MHz, $CDCl_3$) δ 34.22 (q, *J* = 19.3 Hz) ; IR (neat) $\nu_{C=O}$ 1640 cm^{-1}; $\lambda_{max\ abs}$ 463 nm (ε = 50452 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 517 nm, Φ = 0.006 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{17}BF_2N_3O$ 328.14; found 328.13.

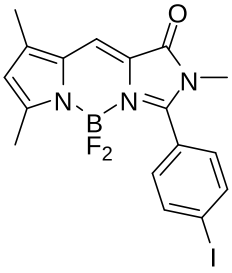
5b




N
Ph
N
N
B
O
F2
1H NMR (CDCl3)
10
9
8
7
6
5
4
3
2
1
ppm

5b

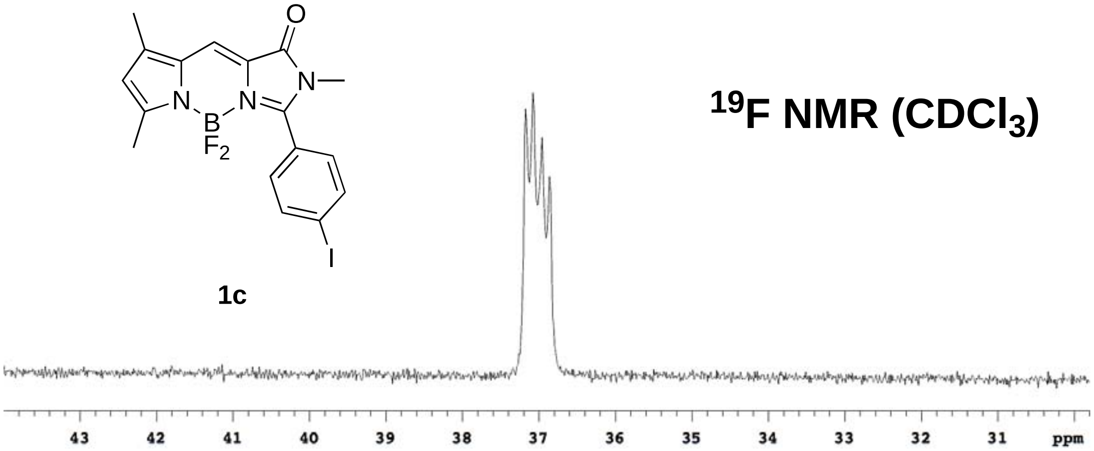

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm

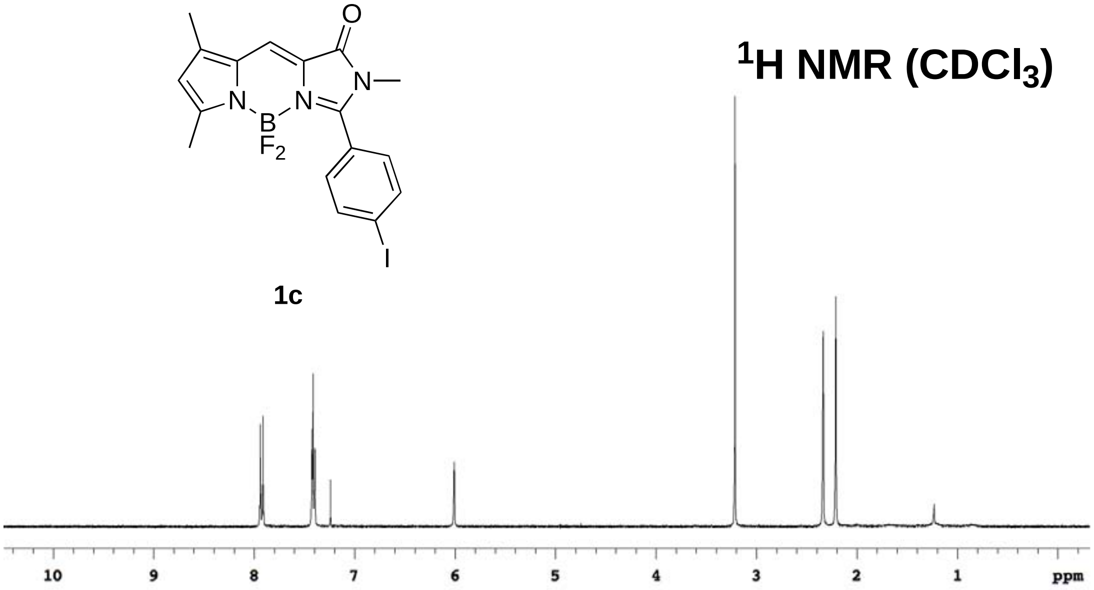


1c

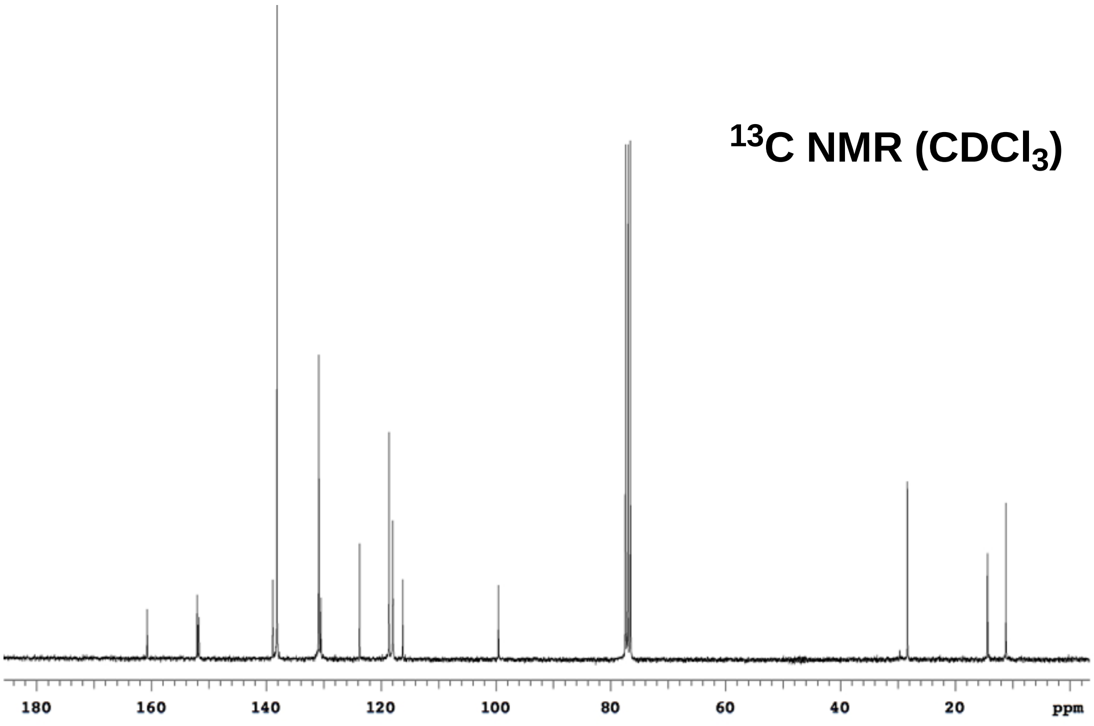
GFP-chromophore analog 1c

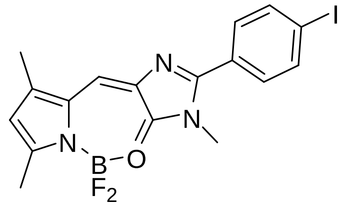
Purified by flash chromatography (1:9 EtOAc/hexanes) to afford **1c** (672 mg, 24 %) as an orange solid. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.93 (d, 2H, *J* = 8.8 Hz), 7.42-7.40 (m, 3H), 6.01 (s, 1H), 3.21 (s, 3H), 2.34 (s, 3H), 2.21 (s, 3H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 160.7, 152.0, 151.7, 138.9, 138.1, 130.8, 130.5, 123.8, 118.6, 118.0, 116.2, 99.6, 28.3, 14.4, 11.2; ^{19}F NMR (282 MHz, CDCl$_3$) δ 37.02 (q, *J* = 27.5 Hz); IR (neat) $\nu_{C=O}$ 1705 cm^{-1}; $\lambda_{max\ abs}$ 492 nm (ε = 57303 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 532 nm, Φ = 0.86 in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ $C_{17}H_{16}BF_2IN_3O$ 454.04; found 454.04.




O
N
N
B
F2
N
I
1H NMR (CDCl3)
10
9
8
7
6
5
4
3
2
1
ppm

1c

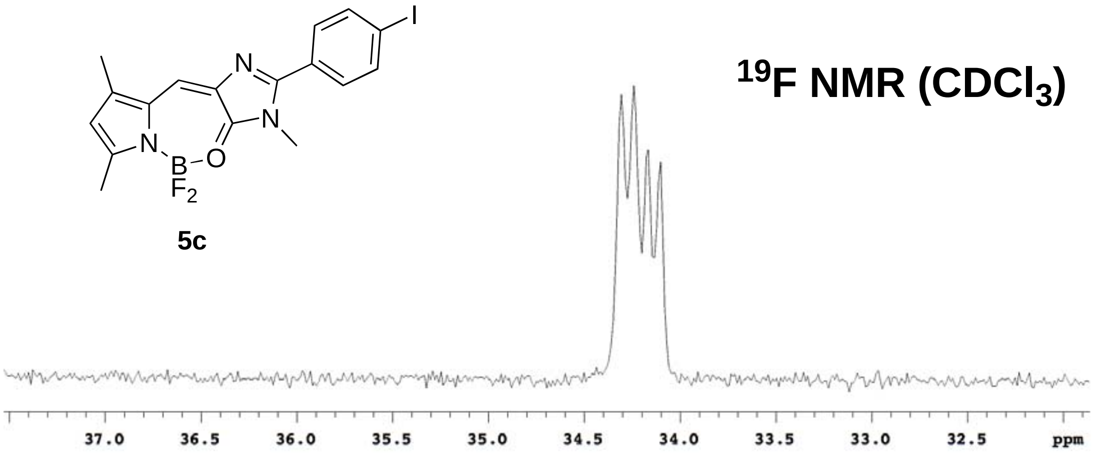

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm

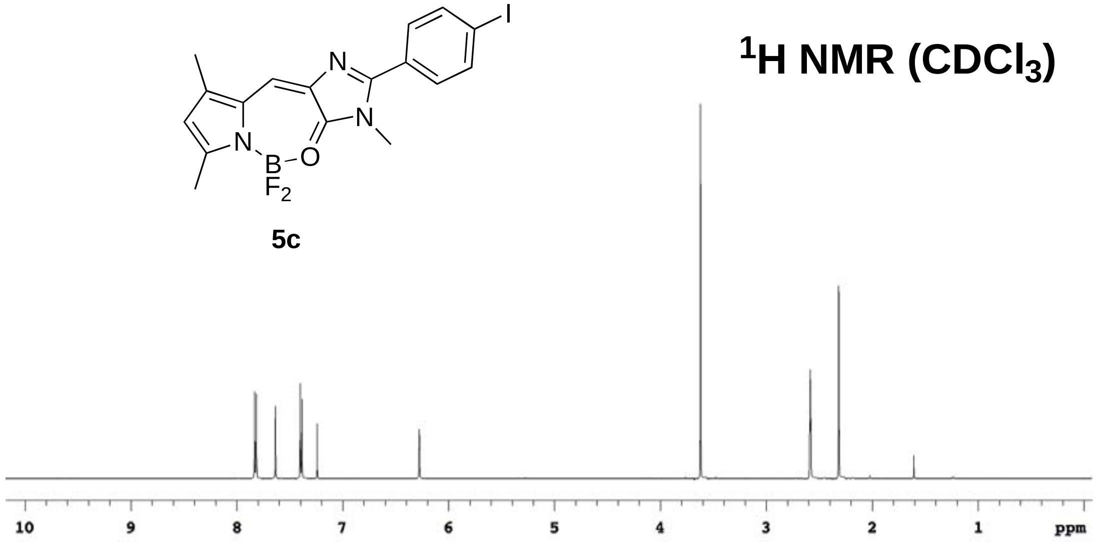


5c

GFP-chromophore analog 5c

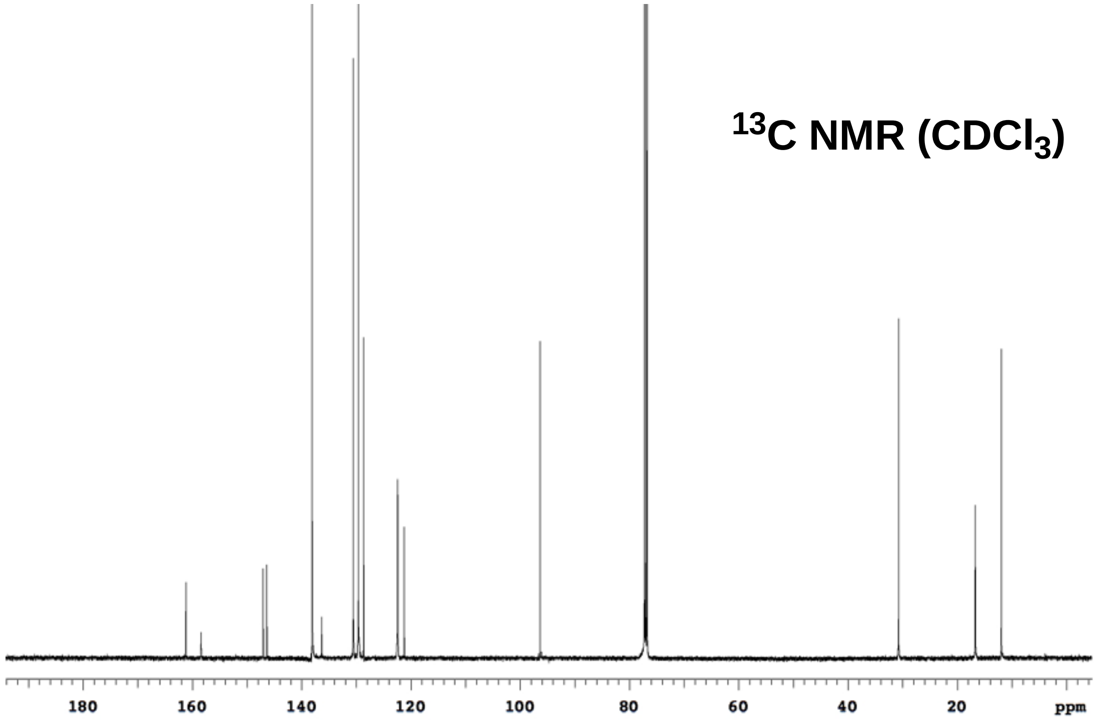
Purified by flash chromatography (1:9 EtOAc/hexanes) to afford **5c** (1.64 g, 58 %) as a red solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.82 (d, 2H, *J* = 8.5 Hz), 7.64 (s, 1H), 7.39 (d, 2H, *J* = 8.5 Hz), 6.28 (s, 1H), 3.62 (s, 3H), 2.58 (t, 3H, *J* = 3.0 Hz), 2.32 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 161.2, 158.4 (t, *J* = 6.1 Hz), 147.0, 146.3, 138.1, 136.3, 130.5, 129.6, 128.6, 122.4, 121.2, 96.4, 30.7, 16.7 (t, *J* = 6.7 Hz), 11.9; ^{19}F NMR (282 MHz, $CDCl_3$) δ 34.20 (q, *J* = 18.9 Hz); IR (neat) $\nu_{C=O}$ 1610 cm^{-1}; $\lambda_{max\ abs}$ 472 nm (ε = 42669 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 520 nm, Φ = 0.01 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{16}BF_2IN_3O$ 454.04; found 454.03.

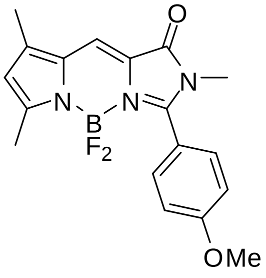



I
N
N
N
B
O
O
F2

5c

^{1}H NMR (CDCl$_3$)
10
9
8
7
6
5
4
3
2
1
ppm

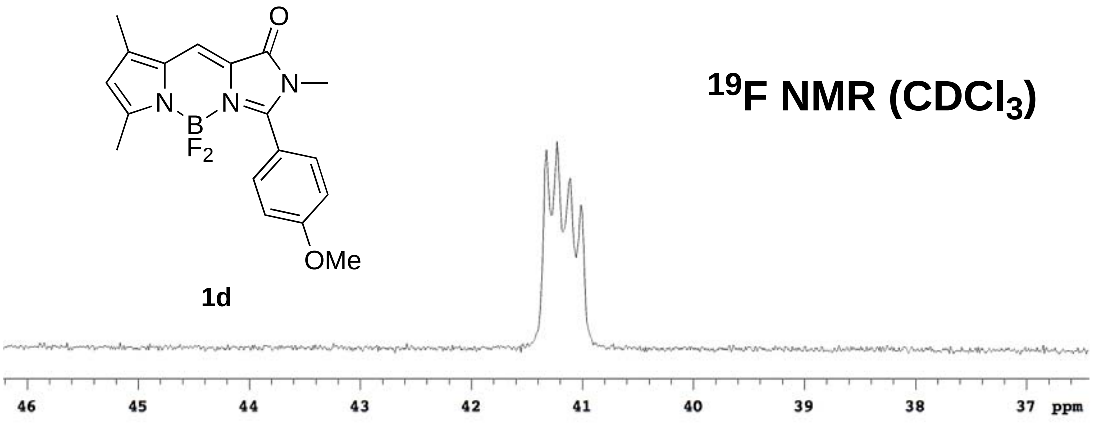

^{13}C NMR (CDCl$_3$)
180
160
140
120
100
80
60
40
20
ppm

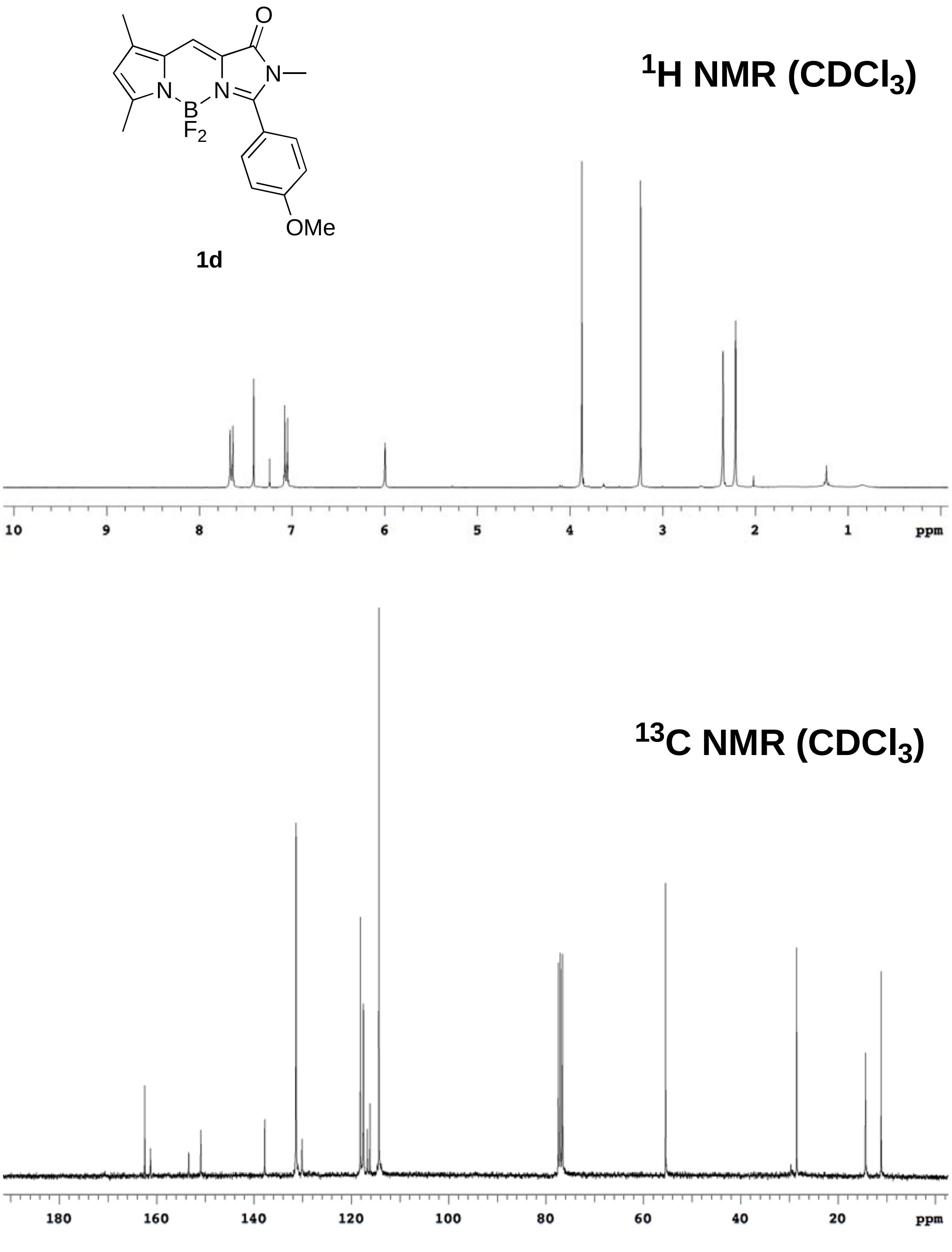


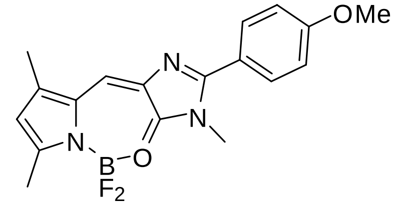
1d

GFP-chromophore analog 1d

Purified by flash chromatography (1:1:3 EtOAc/CH_2Cl_2/hexanes) to afford **1d** (48 mg, 13 %) as a red solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.65 (d, 2H, *J* = 8.9 Hz), 7.41 (s, 1H), 7.06 (d, 2H, *J* = 8.9 Hz), 5.99 (s, 1H), 3.87 (s, 3H), 3.24 (s, 3H), 2.35 (s, 3H), 2.21 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.4, 161.3, 153.4, 150.9, 137.8, 131.3, 130.1, 118.1, 117.5, 116.7, 116.2, 114.3, 55.4, 28.5, 14.3, 11.1; ^{19}F NMR (282 MHz, $CDCl_3$) δ 41.17 (q, *J* = 27.4 Hz) ; IR (neat) $\nu_{C=O}$ 1664 cm^{-1}; $\lambda_{max\ abs}$ 493 nm (ε = 51269 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 521 nm, Φ = 0.85 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{18}H_{19}BF_2N_3O_2$ 358.15; found 358.15.



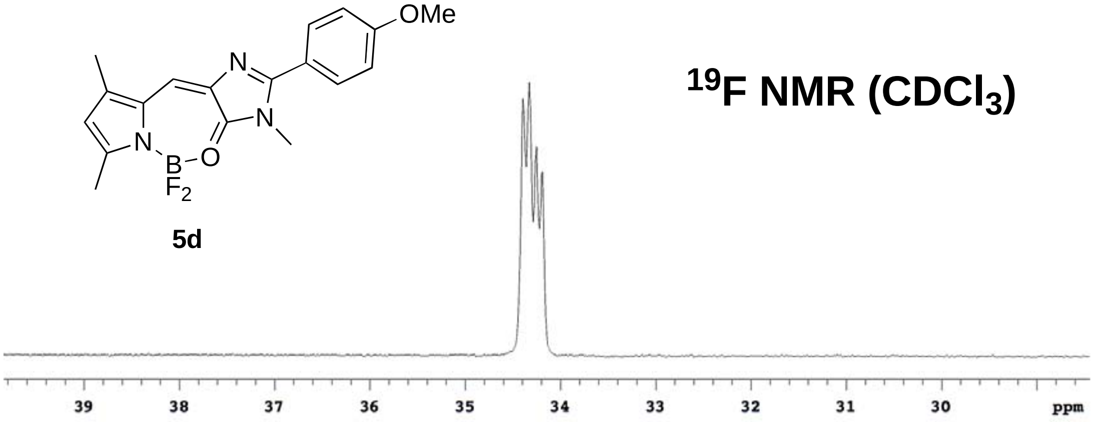

O
N
N
B
F2
N
OMe
1d
1H NMR (CDCl3)
10
9
8
7
6
5
4
3
2
1
ppm
13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm

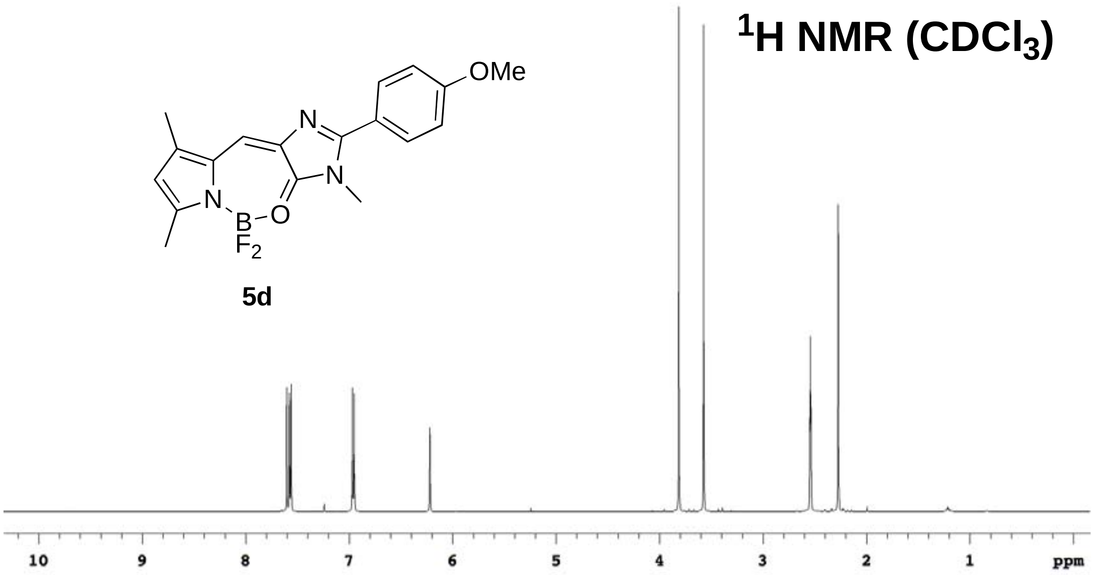


5d

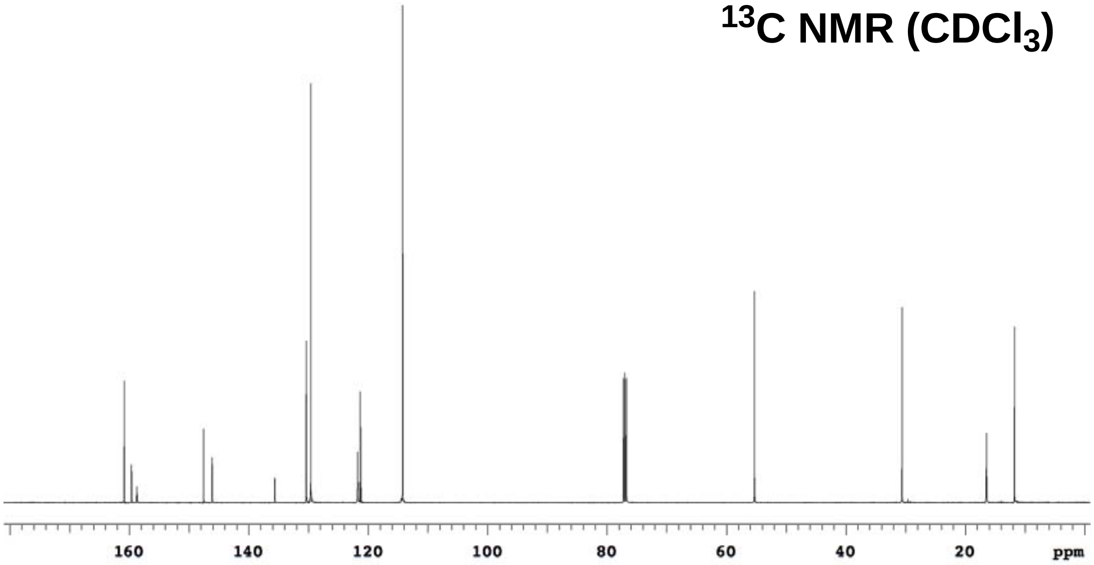
GFP-chromophore analog 5d

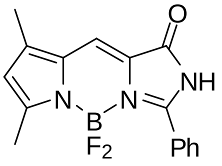
Purified by flash chromatography (1:1:3 EtOAc/CH_2Cl_2/hexanes) to afford **5d** (144 mg, 40 %) as a red solid. ^1H NMR (500 MHz, $CDCl_3$) δ 7.60 (s, 1H), 7.57 (d, 2H, *J* = 10.0 Hz), 6.96 (d, 2H, *J* = 10.0 Hz), 6.22 (s, 1H), 3.81 (s, 3H), 3.57 (s, 3H), 2.54 (t, 3H, *J* = 3.0 Hz), 2.27 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 160.9, 159.6, 158.7 (t, *J* = 5.5 Hz), 147.6, 146.1, 135.6, 130.4, 129.6, 121.8, 121.4, 121.2, 114.2, 55.3, 30.6, 16.5 (t, *J* = 6.4 Hz) , 11.8; ^{19}F NMR (282 MHz, $CDCl_3$) δ 34.29 (q, *J* = 18.7 Hz) ; IR (neat) $\nu_{C=O}$ 1616 cm^{-1}; $\lambda_{max\ abs}$ 455 nm (ε = 38642 $M^{-1}cm^{-1}$), $\lambda_{max\ emiss}$ 528 nm, Φ = 0.03 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{18}H_{19}BF_2N_3O_2$ 358.15; found 358.15.




^{1}H NMR (CDCl$_3$)
OMe
N
N
N
B
O
F$_2$
10
9
8
7
6
5
4
3
2
1
ppm

5d

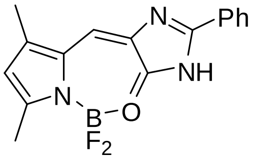

^{13}C NMR (CDCl$_3$)
160
140
120
100
80
60
40
20
ppm



1e

GFP-chromophore analog 1e

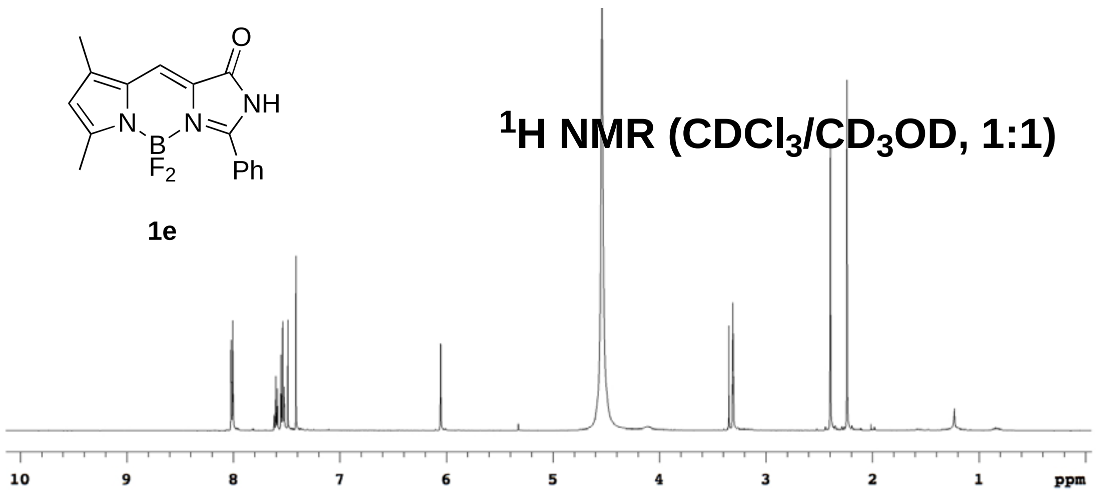
Purified by flash chromatography (1:4 EtOAc/hexanes) to afford **1e** (82 mg, 32 %) as a red solid. ^{1}H NMR (500 MHz, 1:1 $CDCl_3/CD_3OD$) δ 8.02-8.00 (m, 2H), 7.62-7.59 (m, 1H), 7.55-7.52 (m, 2H), 7.41 (s, 1H), 6.05 (s, 1H), 2.40 (s, 3H), 2.24 (s, 3H); ^{13}C NMR (125 MHz, 1:1 $CDCl_3/CD_3OD$) δ 163.6, 153.6, 152.0, 138.8, 133.1, 130.8, 129.9 (t, J = 3.8 Hz), 129.2, 125.4, 119.5, 118.5, 118.4, 14.5 (t, J = 2.4 Hz), 11.2; ^{19}F NMR (282 MHz, 1:1 $CDCl_3/CD_3OD$) δ 41.35 (q, J = 29.1 Hz) ; IR (neat) $\nu_{C=O}$ 1670 cm^{-1}; $\lambda_{max\ abs}$ 495 nm (ε = 37867 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 531 nm, Φ = 0.80 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{16}H_{15}BF_2N_3O$ 314.13; found 314.13.



5e

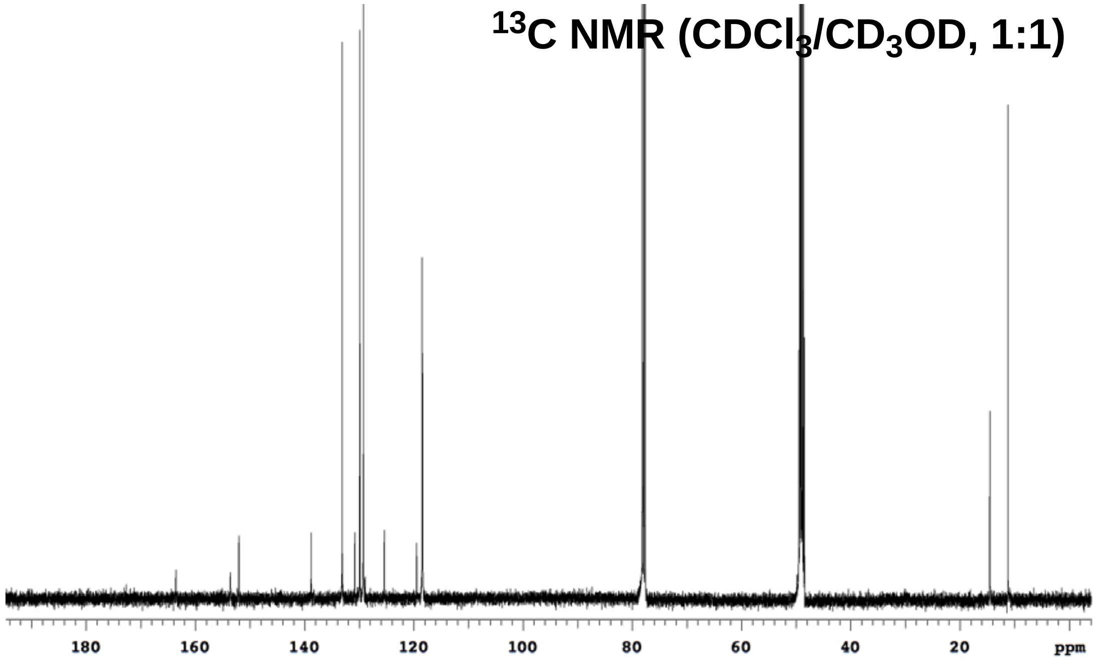
GFP-chromophore analog 5e

Purification was unsuccessful.

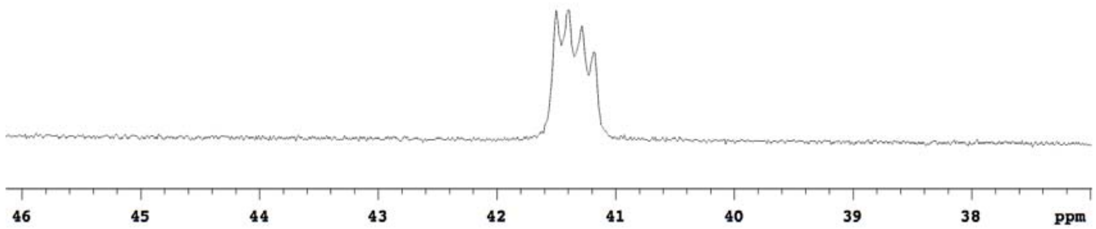


1e

^{1}H NMR ($CDCl_3$/CD_3OD, 1:1)

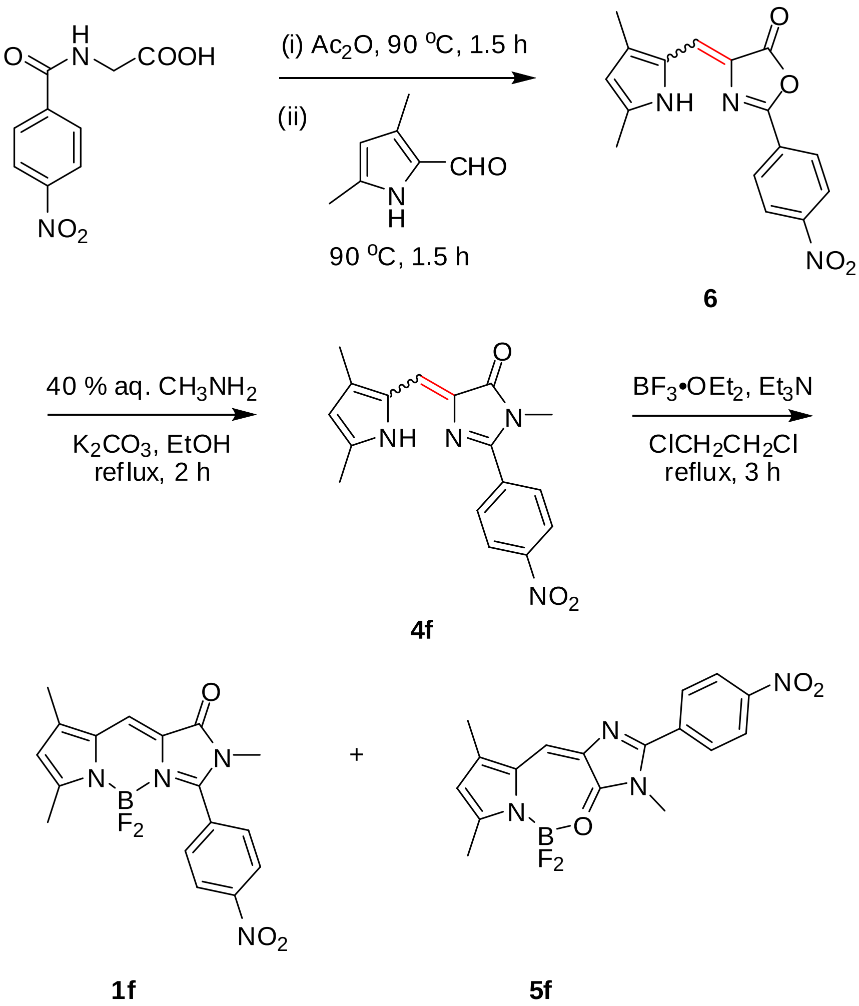


^{13}C NMR ($CDCl_3$/CD_3OD, 1:1)



^{19}F NMR ($CDCl_3$/CD_3OD, 1:1)

Synthesis of GFP-chromophore analogs 1f and 5f


(i) Ac2O, 90 °C, 1.5 h
(ii)
CHO
90 °C, 1.5 h
6
40 % aq. CH3NH2
K2CO3, EtOH
reflux, 2 h
4f
BF3•OEt2, Et3N
ClCH2CH2Cl
reflux, 3 h
+
1f
5f

A solution of 2-(4-nitrobenzamido) acetic acid (1.48 g, 6.6 mmol) in 10 mL Ac_2O was heated to 90 °C and stirred for 1.5 h. To the mixture was added 3,5-dimethyl-1H-pyrrole-2-carbaldehyde (739 mg, 6 mmol) in one portion at room temperature. After stirring at 90 °C for another 1.5 h, the reaction mixture was cooled to room temperature and evaporated to dryness under reduced pressure. The residue was purified by flash chromatography (1:1 CH_2Cl_2/Hexanes) to give product **6** (860 mg, 46 %) as a red solid. NMR analysis of the product indicated that both the ***E*** and ***Z*** isomers were present in a 40:60 ratio. The product was used in the next step without further purification. ^{1}H

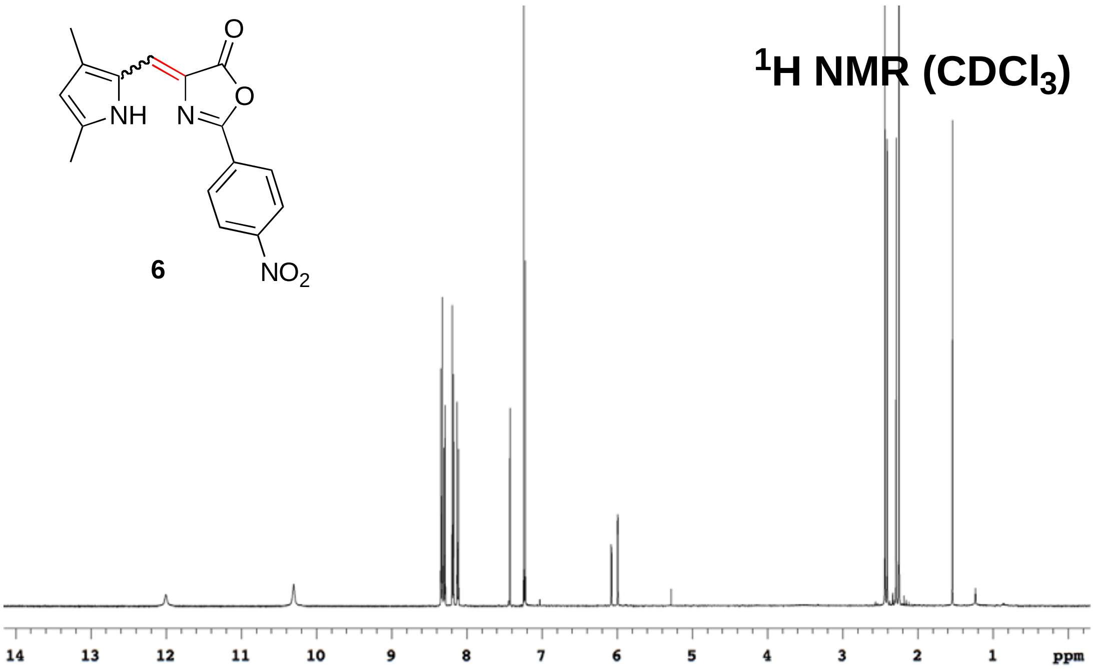
NMR (500 MHz, $CDCl_3$) ***Z*-6**: δ10.30 (s, 1H), 8.33 (d, 2H, *J* = 9.1 Hz), 8.18 (d, 2H, *J* = 9.1 Hz), 7.22 (s, 1H), 5.99 (d, 1H, *J* = 2.4 Hz), 2.44 (s, 3H), 2.25 (s, 3H); ***E*-6**: δ 12.00 (s, 1H), 8.30 (d, 2H, *J* = 9.1 Hz), 8.12 (d, 2H, *J* = 9.1 Hz), 7.42 (s, 1H), 6.07 (d, 1H, *J* = 2.2 Hz), 2.41 (s, 3H), 2.29 (s, 3H); MS (ESI) m/z calcd for $(M+H)^+$ $C_{16}H_{14}N_3O_4$ 312.10; found 312.09.

CH_3NH_2 (40 % in H_2O, 1.9 mL, 22 mmol) was added to the solution of **6** (685 mg, 2.2 mmol) and K_2CO_3 (456 mg, 3.3 mmol) in 25 mL EtOH. The reaction mixture was heated to reflux for 2 h, and then cooled to room temperature. After removal of solvents under reduced pressure, the product was purified by flash chromatography (CH_2Cl_2->1:19 EtOAc/CH_2Cl_2) to afford compound **4f** (243 mg, 34 %) as a red solid. NMR analysis of the product indicated that both the ***E*** and ***Z*** isomers were present in a 48:52 ratio. The product was used in the next step without further purification. ^{1}H NMR (500 MHz, $CDCl_3$) ***Z*-4f**: δ 10.58 (s, 1H), 8.36 (d, 2H, *J* = 8.9 Hz), 7.94 (d, 2H, *J* = 8.9 Hz), 7.23 (s, 1H), 5.91 (d, 1H, *J* = 2.5 Hz), 3.37 (s, 3H), 2.39 (s, 3H), 2.23 (s, 3H); ***E*-4f**: δ 13.12 (s, 1H), 8.32 (d, 2H, *J* = 8.9 Hz), 7.92 (d, 2H, *J* = 8.9 Hz), 7.38 (s, 1H), 6.04 (d, 1H, *J* = 2.2 Hz), 3.47 (s, 3H), 2.34 (s, 3H), 2.27 (s, 3H); MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{17}N_4O_3$ 325.13; found 325.13.

Et_3N (0.34 mL, 2.45 mmol) was added to the solution of **4f** (158 mg, 0.49 mmol) in 10 mL 1,2-dichloroethane. The resulting solution was stirred at room temperature for 10 min and $BF_3 \cdot OEt_2$ (0.49 mL, 3.92 mmol) was added. The reaction mixture was heated to reflux for 3 h then cooled to room temperature. After removal of solvents under reduced pressure, the residue was purified by flash chromatography (1:19 EtOAc/CH_2Cl_2) to afford compound **1f** (84 mg, 46 %) as a red solid and **5f** (79 mg, 44 %) as a red solid.

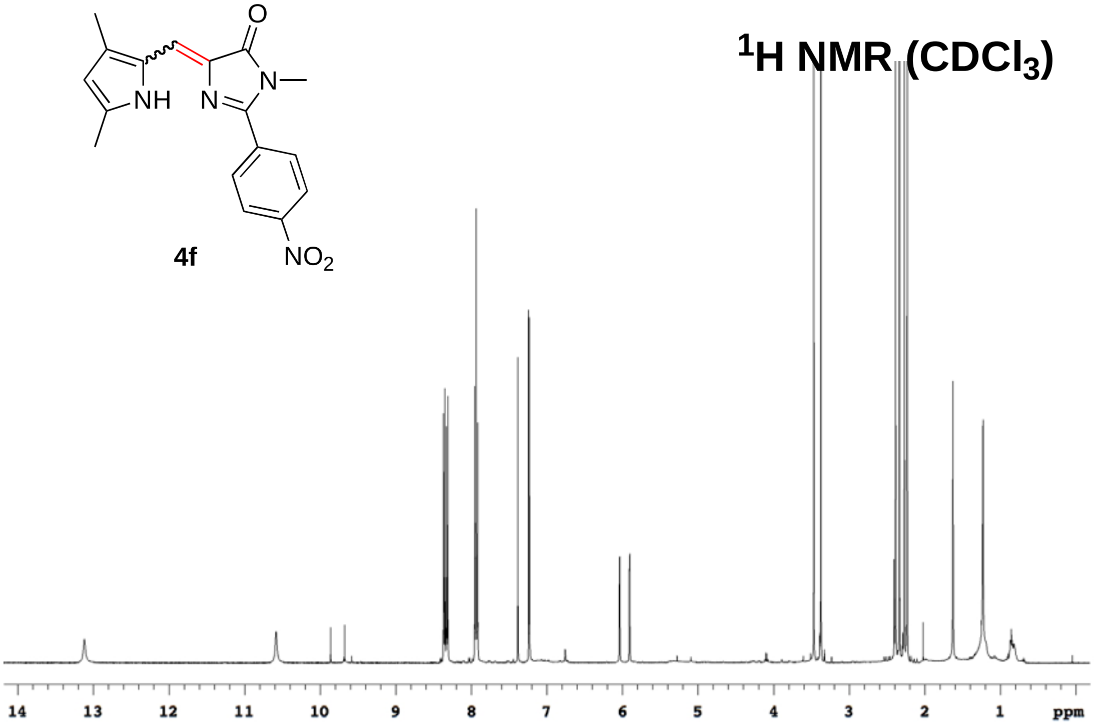
For **1f**: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.43 (d, 2H, *J* = 8.8 Hz), 7.92 (d, 2H, *J* = 8.8 Hz), 7.44 (s, 1H), 6.03 (s, 1H), 3.25 (s, 3H), 2.33 (s, 3H), 2.22 (s, 3H); ^{13}C NMR (500 MHz, $CDCl_3$) δ 160.3, 153.3, 149.7, 149.4, 140.0, 131.1 (t, *J* = 2.8 Hz), 130.9, 130.4, 124.0, 119.0, 118.5, 115.9, 28.3, 14.4 (t, *J* = 2.0 Hz), 11.2; ^{19}F NMR (282 MHz, $CDCl_3$) δ 41.58 (q, *J* = 28.1 Hz) ; IR (neat) $\nu_{C=O}$ 1678 cm^{-1}; $\lambda_{max\ abs}$ 492 nm (ε = 44003 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 598 nm, Φ = 0.0004 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{16}BF_2N_4O_3$ 373.13; found 373.11.

For **5f**: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.34 (d, 2H, *J* = 8.9 Hz), 7.88 (d, 2H, *J* = 8.9 Hz), 7.65 (s, 1H), 6.32 (s, 1H), 3.71 (s, 3H), 2.61 (t, 3H, *J* = 2.9 Hz), 2.34 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 162.9, 158.1 (t, *J* = 5.0 Hz), 148.1, 148.0, 144.4, 137.1, 135.2, 130.4, 128.7, 124.2, 123.1, 121.4, 31.0, 16.8 (t, *J* = 6.5 Hz), 12.0; ^{19}F NMR (282 MHz, $CDCl_3$) δ 34.31 (q, *J* = 18.5 Hz); IR (neat) $\nu_{C=O}$ 1612 cm^{-1}; $\lambda_{max\ abs}$ 472 nm (ε = 35502 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 525 nm, Φ = 0.005 in MeOH; MS (ESI) m/z calcd for $(M+H)^+$ $C_{17}H_{16}BF_2N_4O_3$ 373.13; found 373.12.



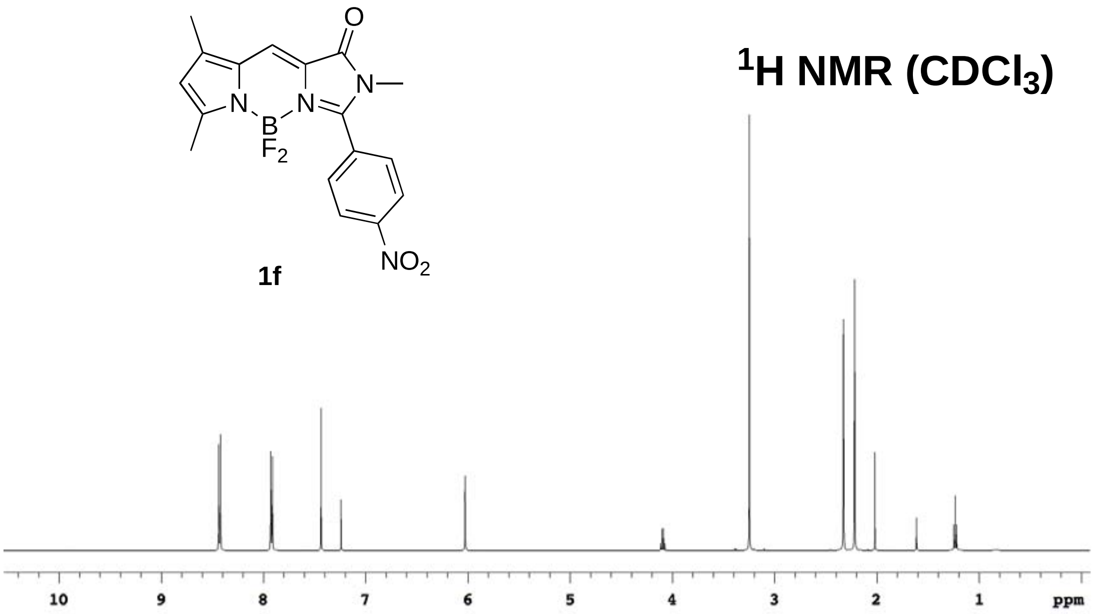
6

^{1}H NMR ($CDCl_3$)

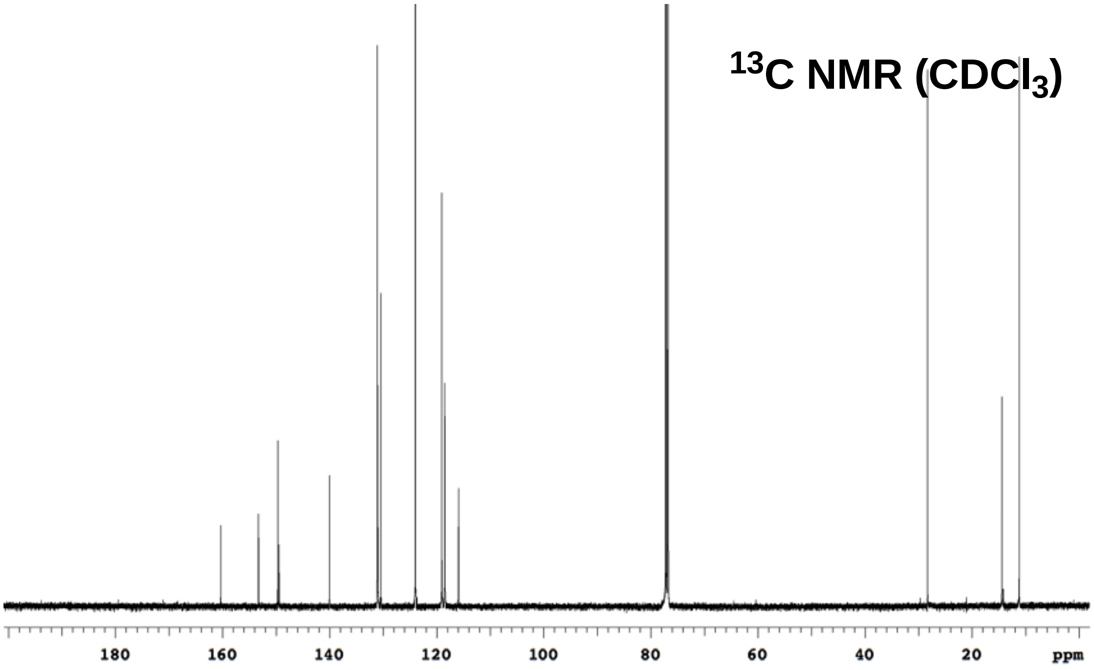


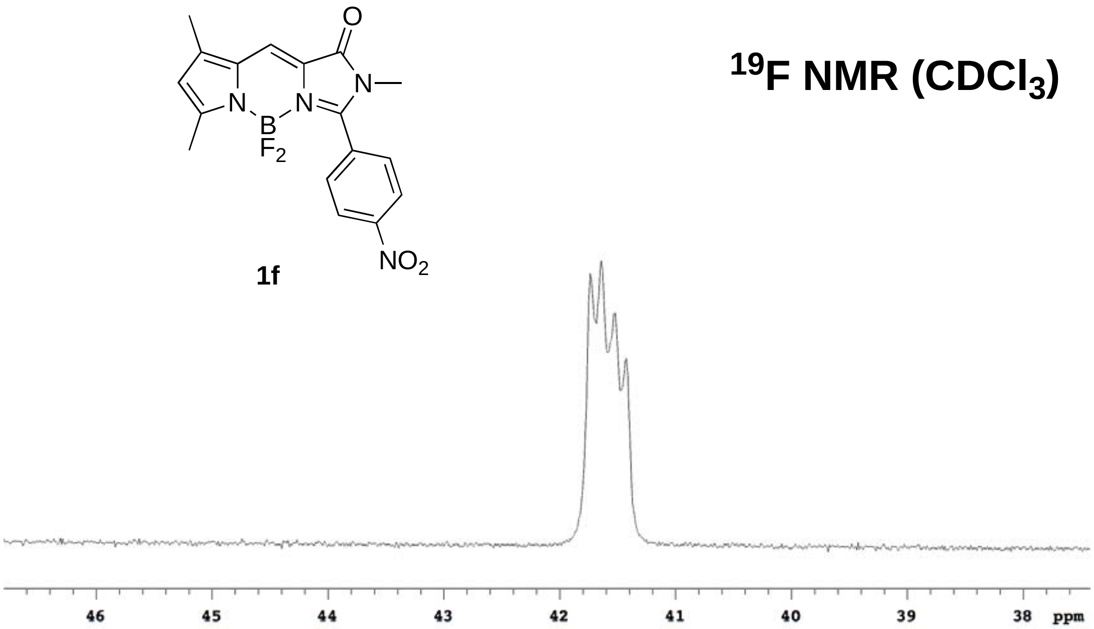
4f

^{1}H NMR ($CDCl_3$)


O
N
N
B
N
F2
NO2
1H NMR (CDCl3)
10
9
8
7
6
5
4
3
2
1
ppm

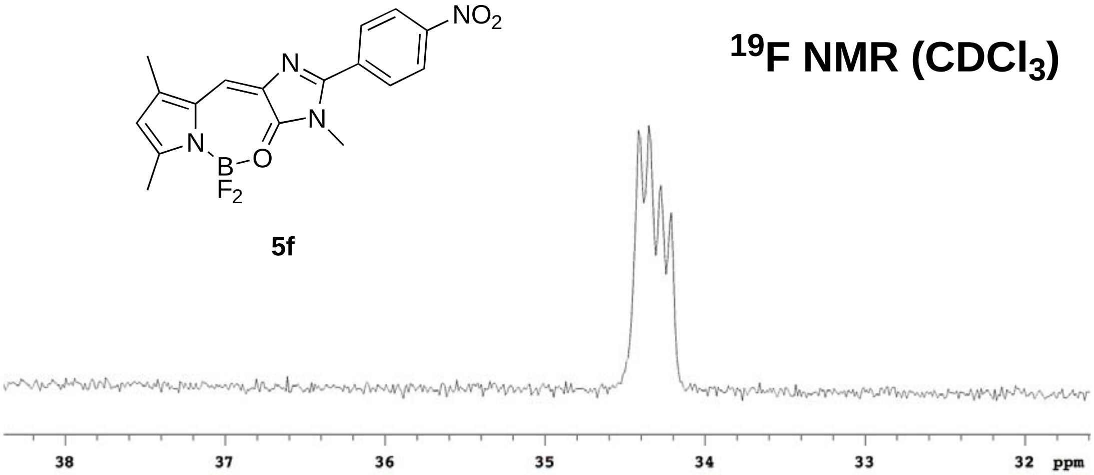
1f


13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm



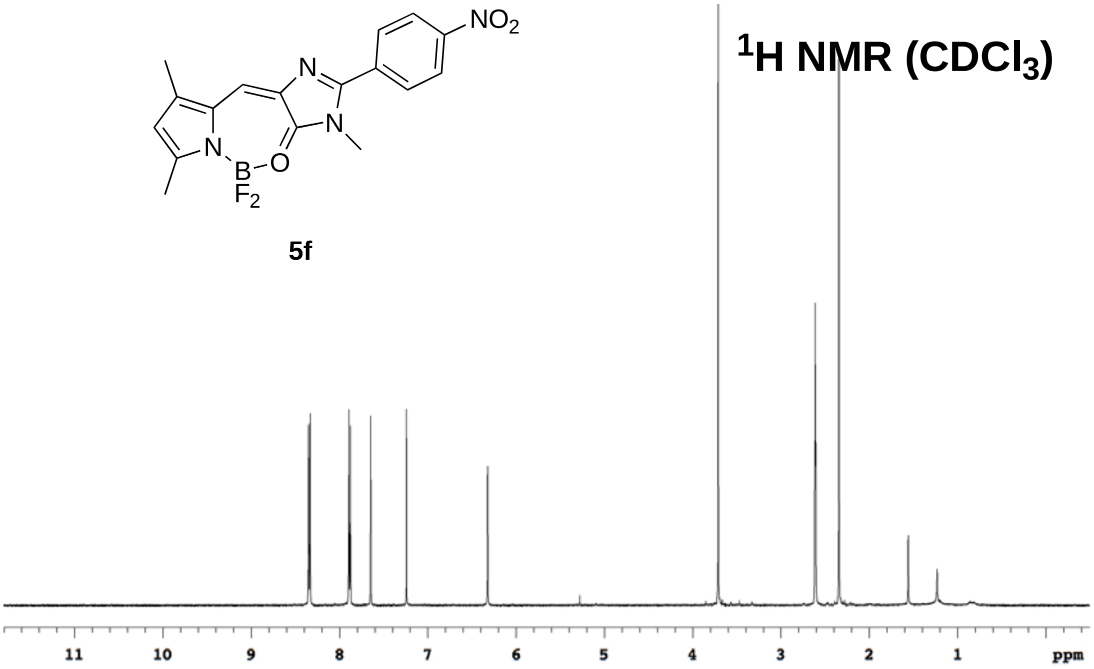
1f

^{19}F NMR (CDCl$_3$)



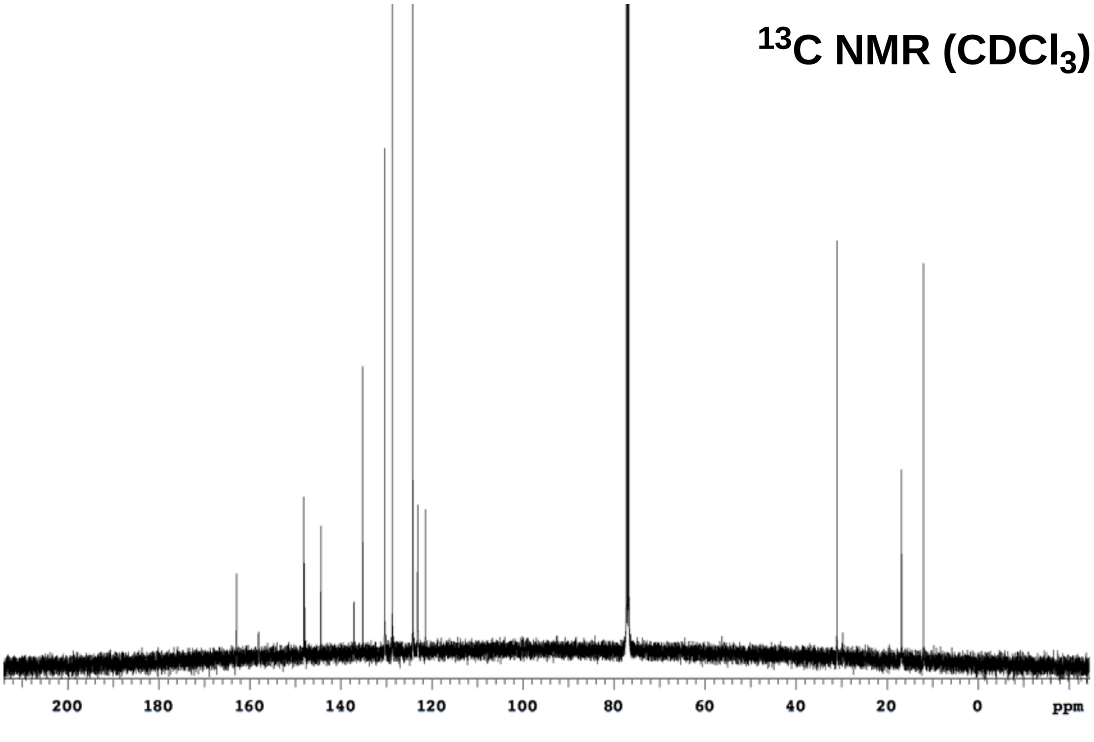
5f

^{19}F NMR (CDCl$_3$)



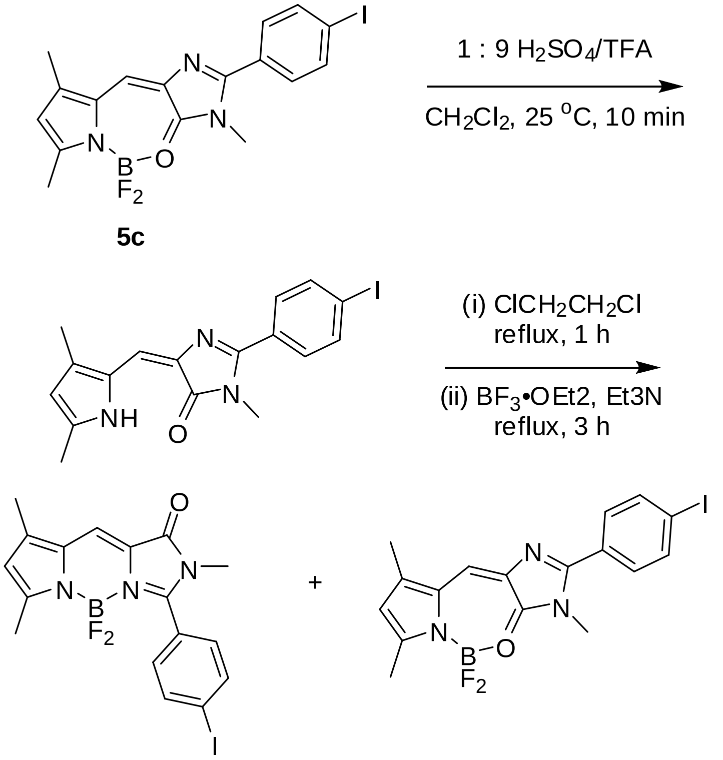
5f

^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

Recycle of undesired GFP-chromophore analog 5



5c

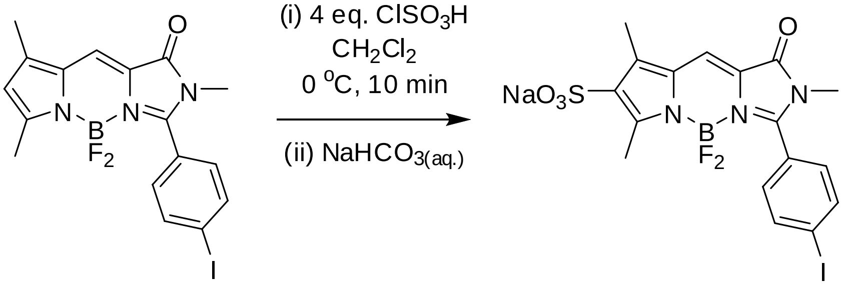
1c

5c

Compound **5c** (227 mg, 0.5 mmol) was dissolved in 5 mL CH_2Cl_2 and 5 mL 1:9 H_2SO_4/TFA was added. The solution was stirred at 25 °C for 10 min then poured into 20 mL ice-water and extracted with CH_2Cl_2 (3 x 50 mL). The organic layers were combined, washed carefully with saturated $NaHCO_3$ (3 x 30 mL), H_2O (1 x 30 mL), brine (1 x 30 mL) and dried over Na_2SO_4. The solvent was removed under reduced pressure to afford an orange solid, which was dried *in vacuo* and used in the next step without further purification.

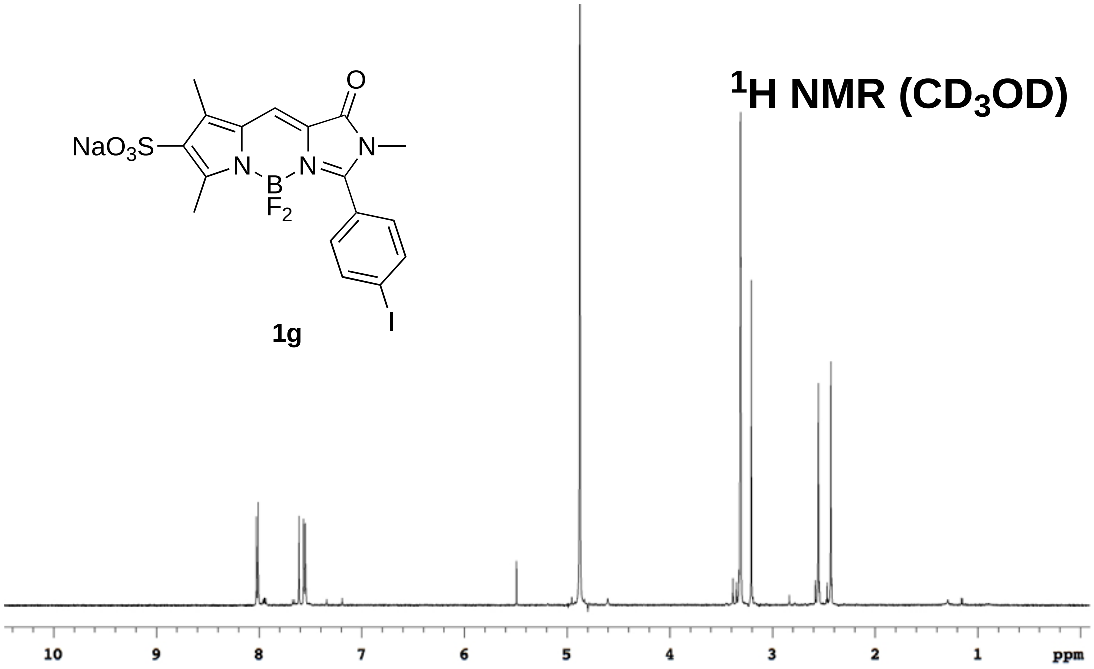
The orange solid was dissolved in 10 mL $ClCH_2CH_2Cl$. After refluxing for 1 h, the solution was cooled to room temperature and Et_3N (0.35 mL, 2.5 mmol) was added, followed by BF_3•OEt_2 (0.51 mL, 4.0 mmol) after 10 min. The reaction mixture was heated to reflux for 3 h. After cooled to room temperature, the solution was diluted with 30 mL CH_2Cl_2 and washed with 30 mL H_2O, 30 mL brine and dried over Na_2SO_4. The solvents were removed under reduced pressure and the residue was purified by flash chromatography (EtOAc/Hexanes, 1:9) to afford **1c** (64 mg, 28 %) as an orange solid and **5c** (127 mg, 56 %) as a red solid.

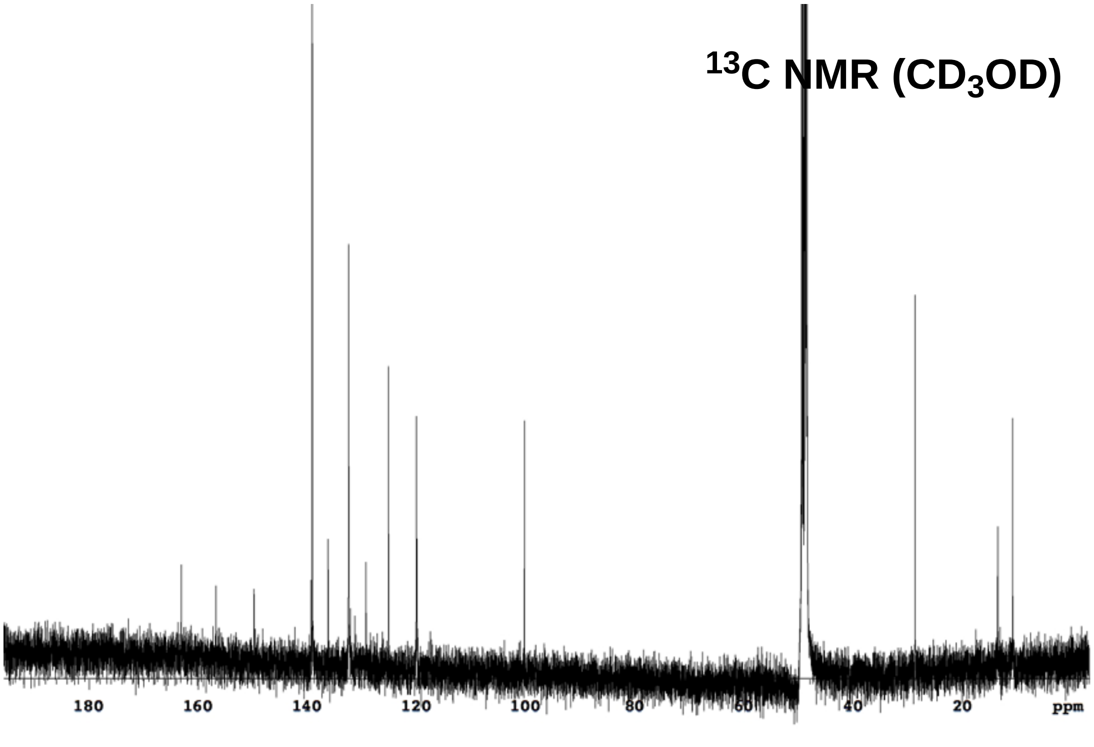
Synthesis of GFP-chromophore analog 1g



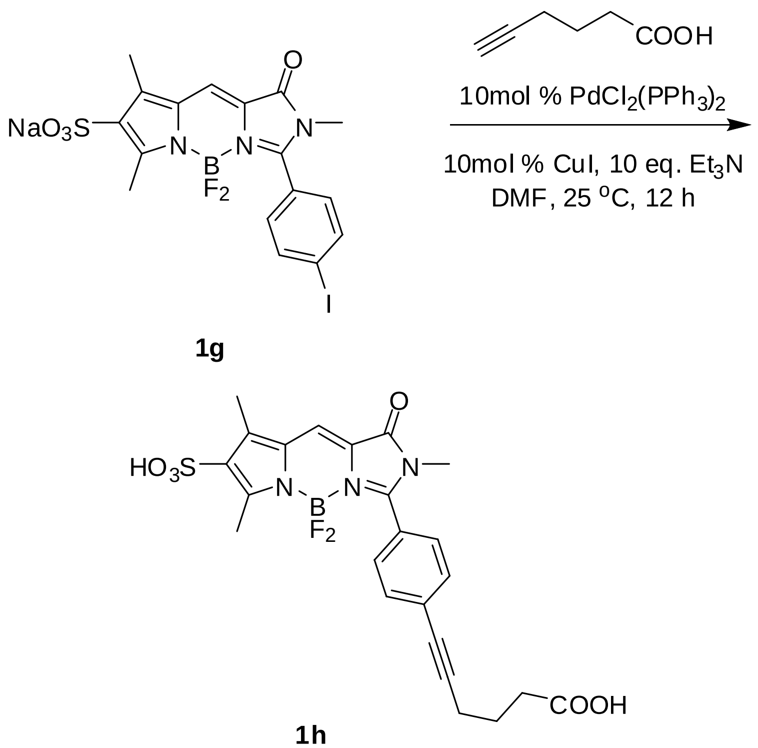
1c **1g**

$ClSO_3H$ (53 μL, 0.8 mmol) was added dropwise to the solution of **1c** (91 mg, 0.2 mmol) in 2 mL CH_2Cl_2 at 0 °C. After stirring at 0 °C for 10 min, the reaction mixture was poured into 20 mL ice-water, neutralized with saturated $NaHCO_{3(aq.)}$ and extracted with $^{i}PrOH/CH_2Cl_2$ (1:1) (3 x 20 mL). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatography (1:9 $MeOH/CH_2Cl_2$) to afford the product **1g** (102 mg, 92 %) as an orange solid. ^{1}H NMR (500 MHz, CD_3OD) δ 8.02 (d, 2H, J = 8.4 Hz), 7.61 (s, 1H), 7.56 (d, 2H, J = 8.4 Hz), 3.20 (s, 3H), 2.55 (s, 3H), 2.43 (s, 3H); ^{13}C NMR (125 MHz, CD_3OD) δ 163.1, 156.7, 149.8, 139.1, 136.2, 132.5, 131.3, 129.3, 125.2, 120.1, 119.9, 100.3, 28.7, 13.6, 10.9. IR (neat) $\nu_{C=O}$ 1688cm^{-1}; $\lambda_{max\ abs}$ 488 nm (ε = 48132 M^{-1}cm^{-1}), $\lambda_{max\ emiss}$ 531 nm, Φ = 0.88 in MeOH; MS (ESI) m/z calcd for $(M\text{-}Na)^{-}$ $C_{17}H_{14}BF_2IN_3O_4S$ 531.98; found 531.88.


NaO$_3$S
O
N
N
B
N
F$_2$
I
1g
^{1}H NMR (CD$_3$OD)
10
9
8
7
6
5
4
3
2
1
ppm


^{13}C NMR (CD$_3$OD)
180
160
140
120
100
80
60
40
20
ppm

Synthesis of GFP-chromophore analog 1h

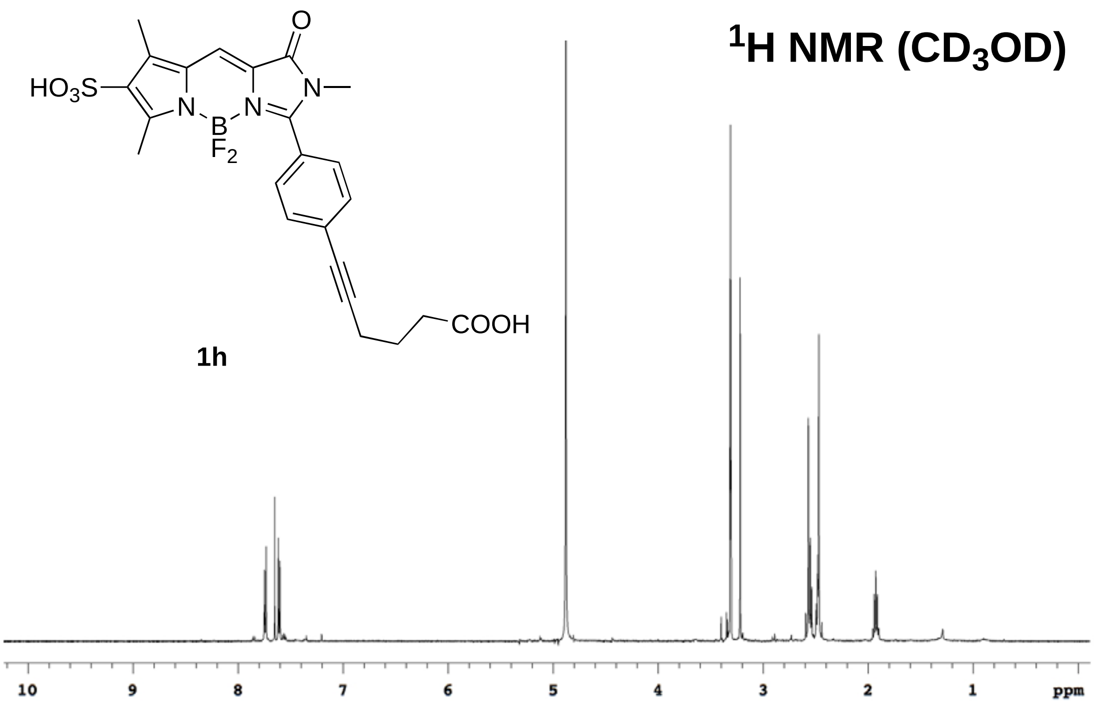


1g

1h

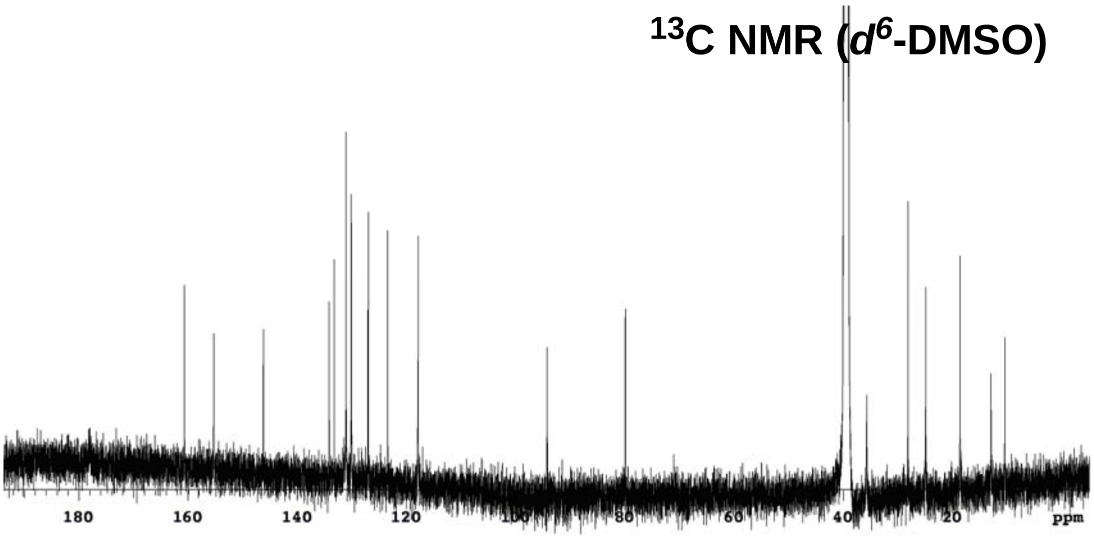
Compound **1g** (56 mg, 0.1 mmol), $PdCl_2(PPh_3)_2$ (7 mg, 0.01 mmol) and CuI (2 mg, 0.01 mmol) were dissolved in 2 mL DMF in a Schlenk tube. The mixture was freeze-pump-thawed three times, and then 5-hexynoic acid (56 μL, 0.5mmol) and Et_3N (140 μL, 1.0 mmol) were added. After stirring at 25 °C for 12 h, the reaction mixture was concentrated under reduced pressure. The residue was dissolved in 20 mL H_2O and acidified to pH ≈ 2.0 with 1N HCl. The solution was extracted with ^{i}PrOH/CH_2Cl_2 (1:1) (3 x 30 mL). The organic layer was dried over Na_2SO_4 and concentrated *in vacuo*. The crude product was purified by flash chromatography (1:4 MeOH/CH_2Cl_2) to give the desired product **1h** as an orange solid (37 mg, 71 %). ^{1}H NMR (500 MHz, CD_3OD) δ 7.74 (d, 2H, *J* = 8.4 Hz), 7.65 (s, 1H), 7.61 (d, 2H, *J* = 8.4 Hz), 3.22 (s, 3H), 2.57-2.47 (m, 10H), 1.96-1.90 (m, 2H); ^{13}C NMR (125 MHz, d^6-DMSO) δ 178.0, 160.7, 155.3, 146.2, 134.2, 133.2, 131.1, 130.1, 127.1, 127.0, 123.4, 117.9, 117.8, 94.3, 79.9, 35.7, 28.2, 24.9, 18.6, 13.0, 10.4. IR (neat) $\nu_{C=O}$ 1702, 1593cm^{-1}; $\lambda_{max\ abs}$ 488 nm (ε = 35024

$M^{-1}cm^{-1}$), $\lambda_{max\ emiss}$ 538 nm, $\Phi = 0.88$ in MeOH; MS (ESI) m/z calcd for $(M\text{-}H)^-$ $C_{23}H_{21}BF_2N_3O_6S$ 516.12; found 516.05.



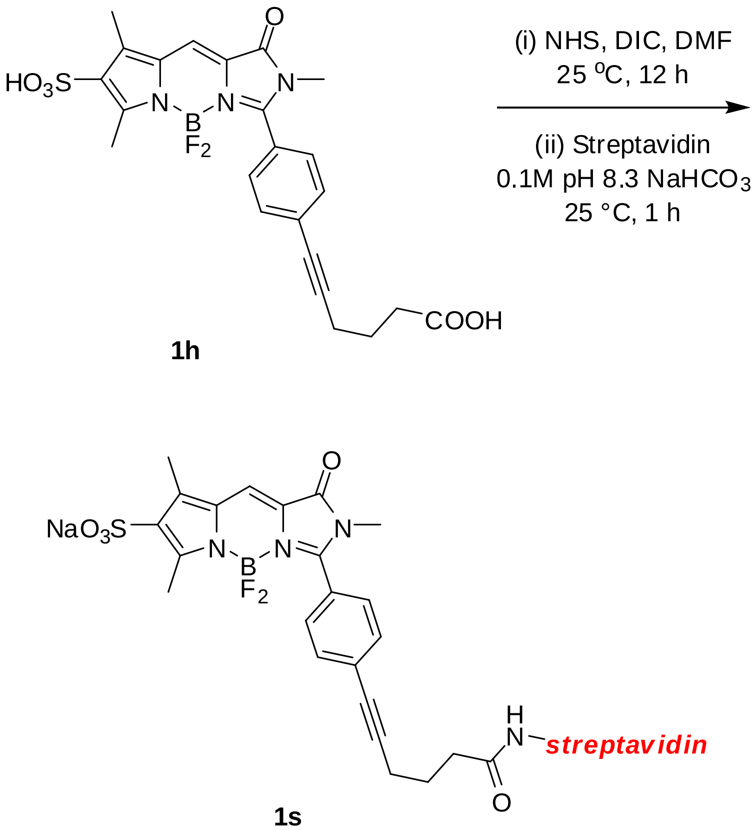
1h

^{1}H NMR (CD_3OD)



^{13}C NMR (d^6-DMSO)

Labeling of streptavidin with 1h


HO3S
N
B
N
F2
O
N
COOH
(i) NHS, DIC, DMF
25 °C, 12 h
(ii) Streptavidin
0.1M pH 8.3 NaHCO3
25 °C, 1 h
NaO3S
H
N
streptavidin

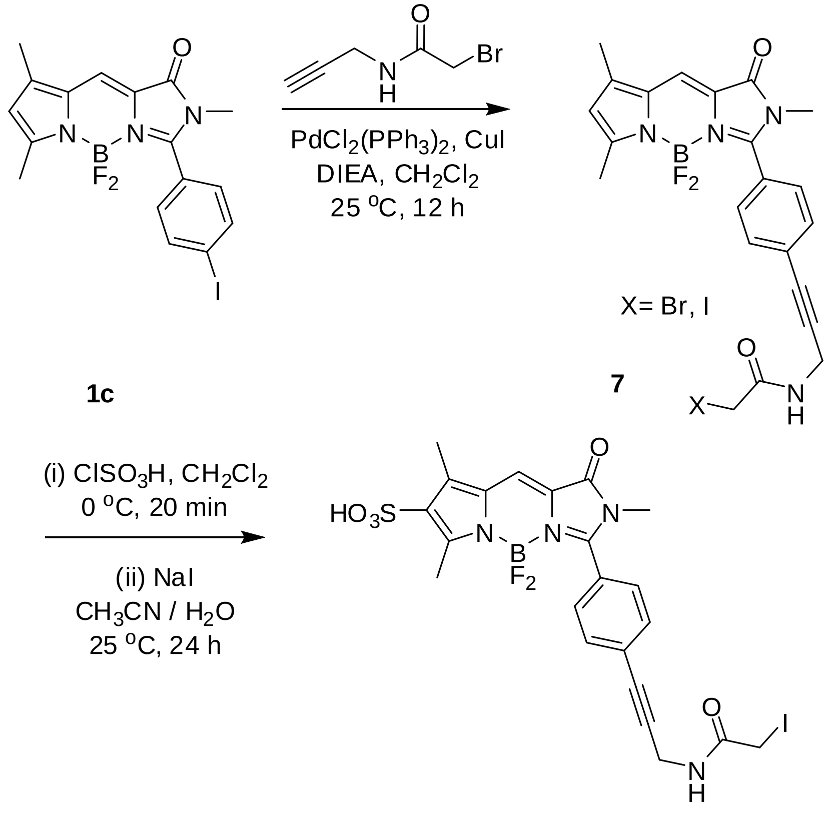
1h

1s

dye:protein 3.0:1.0

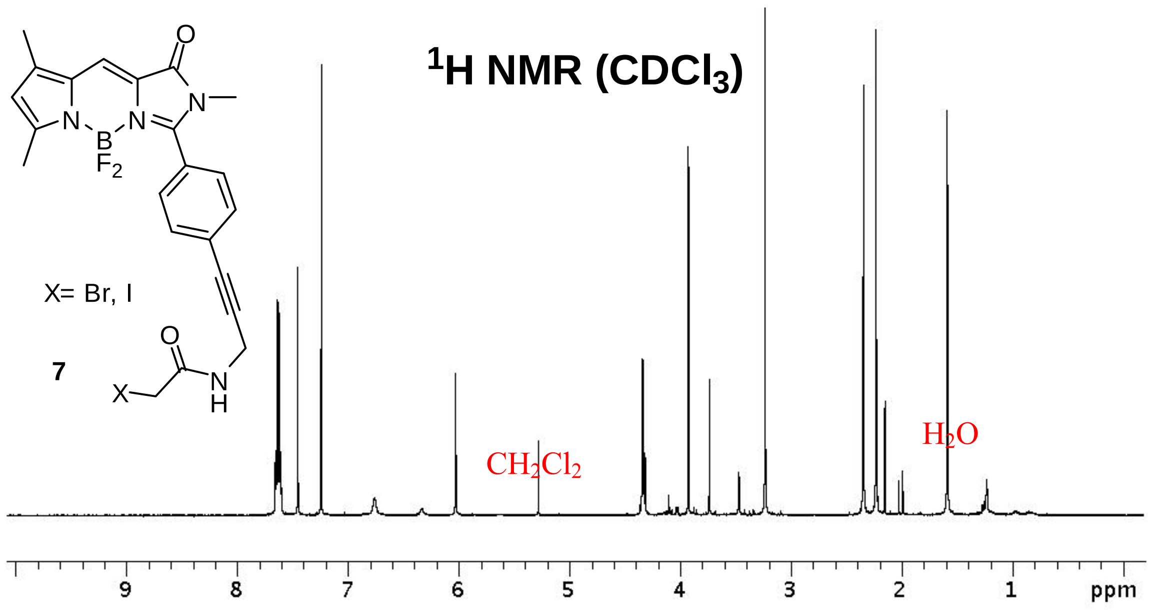
The solution of **1h** (1.9 mg, 3.7 μmol), NHS (2 mg, 18.5 μmol), DIC (3 μL, 18.5 μmol) in 0.2 mL DMF was stirred at 25 °C for 12 h. The above solution (21 μL) was added to a solution of 2 mg streptavidin in 0.2 mL 0.1M $NaHCO_3$ buffer (pH = 8.3). The resulting solution was stirred at 25 °C in the dark for 1 h. The mixture was purified by PD-10 desalting column to give the labeled streptavidin **1s**.

Synthesis of thiol-reactive GFP-chromophore analog 1i

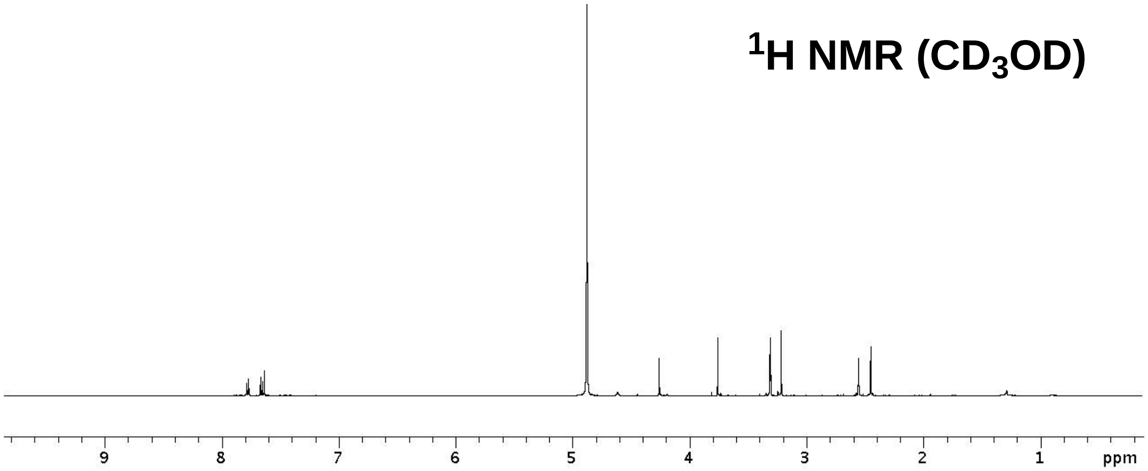


1c **7** **1i**

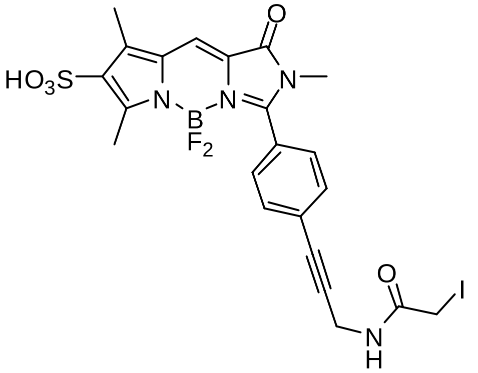
A solution of **1c** (91 mg, 0.2 mmol) , 2-bromoacetyl propargylamide (176 mg, 1.0 mmol), $PdCl_2(PPh_3)_2$ (7 mg, 0.01 mmol) and CuI (2 mg, 0.01 mmol) in 5 mL CH_2Cl_2 was degassed and filled with N_2 then DIEA (0.17 mL, 1.0 mmol) was added. The reaction mixture was stirred at 25 °C for 12 h then concentrated under reduced pressure. The residue was purified by flash chromatography (60 % EtOAc/Hexanes) to afford the product **7** (85 mg, 85 %) as a red solid. ^{1}H NMR and MS analysis shows **7** was a mixture of two compounds which corresponded to X = Br and I. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.65-7.60 (m, 4H), 7.45 (s, 1H), 6.76 (br, 1H) and 6.33 (br, 1H), 6.03 (s, 1H), 4.34 (d, 2H, *J* = 5.4 Hz) and 4.32 (d, 2H, *J* = 5.4 Hz), 3.92 (s, 2H) and 3.74 (s, 2H), 3.23 (s, 3H), 2.34 (s, 3H), 2.23 (s, 3H). MS (ESI) m/z calcd for $(M+H)^+$ $C_{22}H_{21}BBrF_2N_4O_2$ 501.09; found 501.10 (X = Br); calcd for $(M+H)^+$ $C_{22}H_{21}BF_2IN_4O_2$ 549.08; found 549.08 (X = I).



A solution of the above product **7** (70 mg, 0.14 mmol) in 10 mL CH_2Cl_2 was cooled to 0 °C and $ClSO_3H$ (38 µL, 0.56 mmol) was added dropwise over 1 min. The reaction mixture was stirred at 0 °C for 20 min then quenched with 1 mL H_2O. The mixture was concentrated under reduced pressure and the residue was purified by flash chromatography (20 % MeOH/CH_2Cl_2) to give a red solid. This intermediate was dissolved in 6 mL CH_3CN and 4 mL H_2O then NaI (63 mg, 0.42 mmol) was added in one portion. The reaction mixture was stirred at 25 °C for 24h then concentrated under reduced pressure. The residue was purified by flash chromatography (20 % MeOH/CH_2Cl_2) to afford the product **1i** (86 mg, 98 %) as a red solid. ^{1}H NMR (500 MHz, CD_3OD) δ 7.78 (d, 2H, *J* = 8.3 Hz), 7.66 (d, 2H, *J* = 8.3 Hz), 7.63 (s, 1H), 4.26 (s, 2H), 3.76 (s, 2H), 3.22 (s, 3H), 2.55 (s, 3H), 2.45 (s, 3H); ^{13}C NMR (125 MHz, CD_3OD) δ 171.1, 163.1, 157.0, 149.5, 136.0, 132.7 (2C), 131.1, 129.2, 128.2, 125.3, 120.0, 119.9, 88.8, 82.7, 31.0, 28.9, 13.6, 10.9, -2.5. MS (ESI) m/z calcd for $(M-H)^-$ $C_{22}H_{19}BF_2IN_4O_5S$ 627.02; found 626.93.



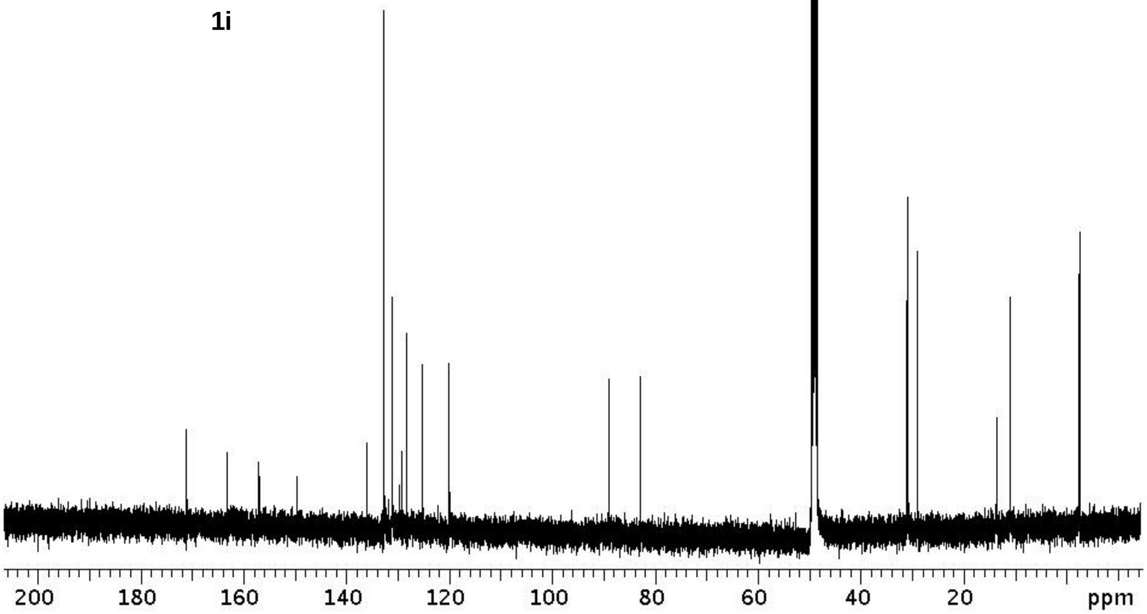
^{1}H NMR (CD$_3$OD)



HO$_3$S
N
B
F$_2$
N
N
O
O
I
N
H

1i

^{13}C NMR (CD$_3$OD)



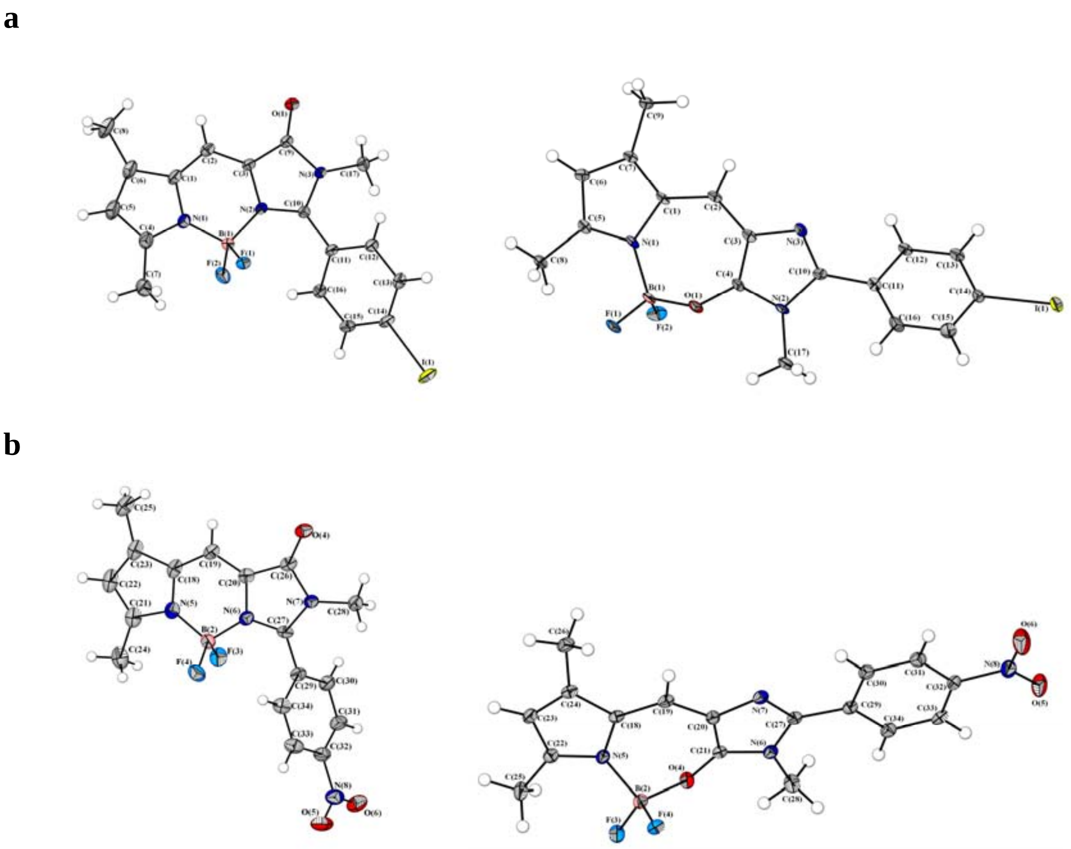
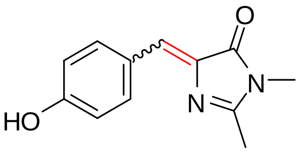
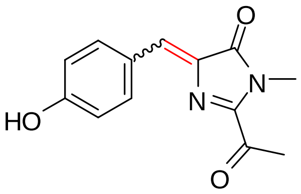

a
b
O(1)
N(1)
N(2)
N(3)
B(1)
F(1)
F(2)
I(1)
O(4)
O(5)
O(6)
N(5)
N(6)
N(7)
N(8)
B(2)
F(3)
F(4)

Figure B1. X-ray structures for GFP-chromophore analogs: **a** **1c** and **5c**; **b** **1f** and **5f**.

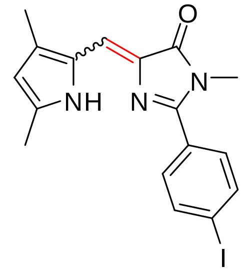
Table B1. Calculated free energies of ***E*** and ***Z*** isomers



A



G



4c

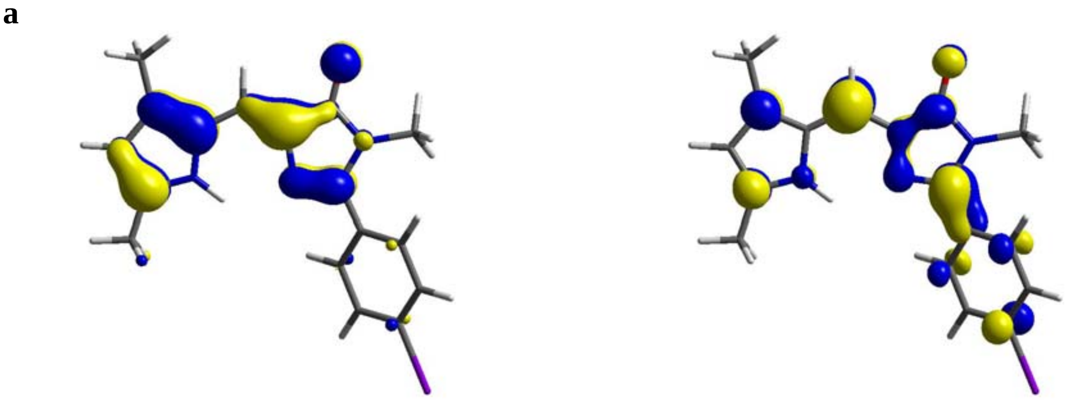
	A[a]	**G**[b,88]		**4c**[b]	
	Aqueous solution	Gas phase	Aqueous solution	Gas phase	DMSO
Z (Hartree)	–	- 837.84379	- 837.84976	- 908.160846	- 908.180318
E (Hartree)	–	- 837.84114	-837.84715	- 908.162969	- 908.181293
Energy of E vs Z (Kcal/mol)	2.1	1.66	1.64	- 1.33	- 0.61

[a] Experimental value;[91] [b] calculated with B3LYP/6-31+G** (SDD for I in **4c**)

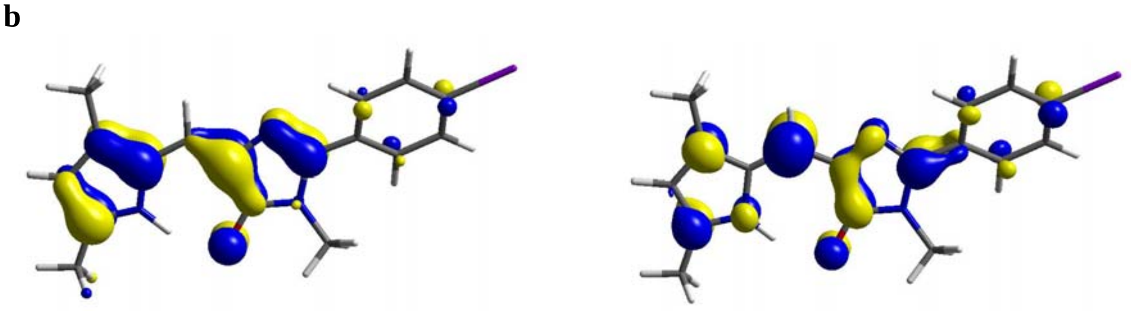
Table B2. Calculated spectroscopic properties of **4c**[a]

	ΔE, eV	λ_{max}, nm	Oscillator strength (*f*)
		Gas phase	
***Z* – 4c**	2.83	438	0.70
***E* – 4c**	2.89	430	0.99
		DMSO	
***Z* – 4c**	2.78	446 **(458)**[b]	0.89
***E* – 4c**	2.81	442 **(464)**[b]	1.09

[a] Calculated with B3LYP/6-31+G** (SDD for I); [b] data in bold are experimental values



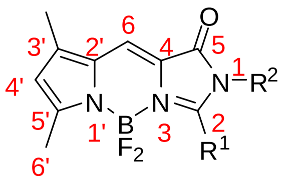
HOMO of ***Z*-4c** **LUMO** of ***Z*-4c**



HOMO of ***E*-4c** **LUMO** of ***E*-4c**

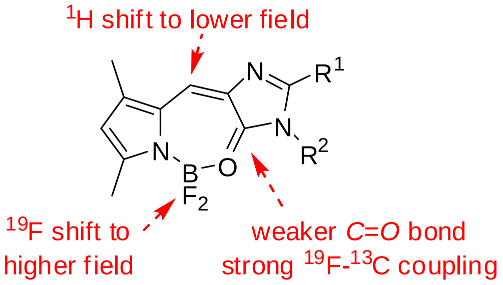
Figure B2. HOMO and LUMO pictures from B3LYP/6-31+G** (SDD for I) calculation: a ***Z*-4c**; and, b ***E*-4c**.

Table B3. Spectroscopic differences between compounds **1** and **5**



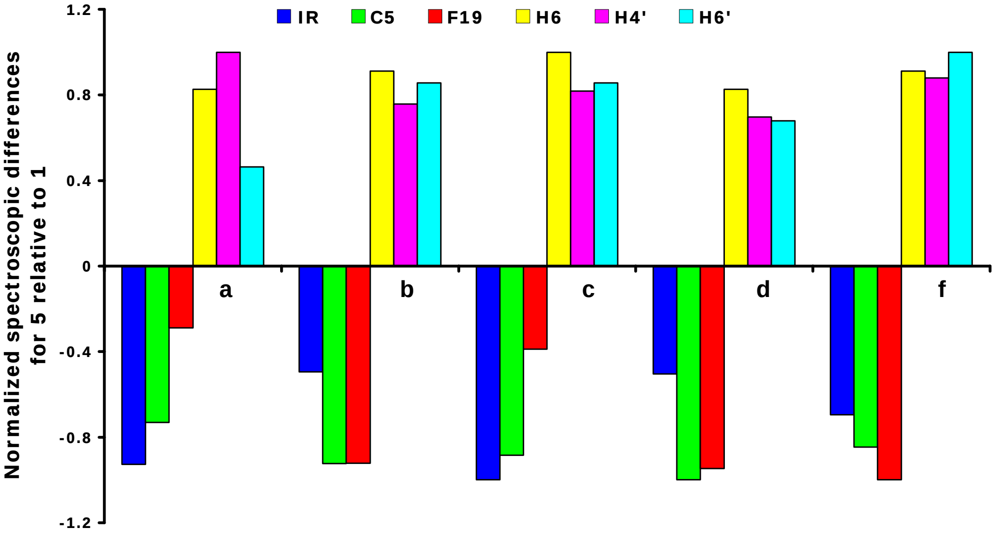
O
6
5
3'
2'
4
1
N
R2
4'
N
N
5'
B
3
2
1'
F2
R1
6'

1



1H shift to lower field
N
R1
N
N
R2
B
O
F2
19F shift to higher field
weaker C=O bond
strong 19F-13C coupling

5



Trends in chemical shift differences for compounds **5** relative to compounds **1** illustrate that these can be used to differentiate between them. Here the zero line is the chemical shift values for compounds **1**, and the relative chemical shift difference for the corresponding compound **5** is indicated. The chemical shift differences for the ^{19}F signals shown here are 0.1 x their actual values.

Table B3. Continued

	$\nu_{C=O}$ (cm^{-1})	δ_{C5} (ppm)	δ_F (ppm)	δ_{H6} (ppm)	$\delta_{H4'}$ (ppm)	$\delta_{H6'}$ (ppm)
1a	1710	160.6	39.43	7.49	6.04	2.31
5a	1622	158.7 (t, *J* = 5.3Hz)	37.33	7.68	6.37	2.44 (t, *J* = 3.0Hz)
1b	1687	161.0	40.92	7.45	6.01	2.34
5b	1640	158.6 (t, *J* = 5.3Hz)	34.22	7.66	6.26	2.58 (t, *J* = 3.0Hz)
1c	1705	160.7	37.02	7.41	6.01	2.34
5c	1610	158.4 (t, *J* = 6.1Hz)	34.20	7.64	6.28	2.58 (t, *J* = 3.0Hz)
1d	1664	161.3	41.17	7.41	5.99	2.35
5d	1616	158.7 (t, *J* = 5.5Hz)	34.29	7.60	6.22	2.54 (t, *J* = 3.0Hz)
1e	1670	163.6	41.35	7.41	6.05	2.40
5e	-	-	-	-	-	-
1f	1678	160.3	41.58	7.44	6.03	2.33
5f	1612	158.1 (t, *J* = 5.0Hz)	34.31	7.65	6.32	2.61 (t, *J* = 2.9Hz)

Table B4. Spectroscopic properties of GFP-chromophore analogs

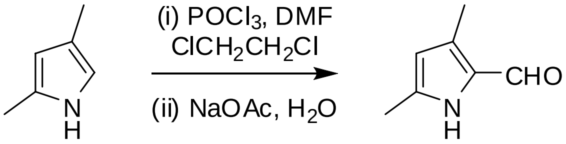
	$\lambda_{\text{max abs}}$ (nm)	ε ($M^{-1}cm^{-1}$)	$\lambda_{\text{max emiss}}$ (nm)	fwhm (nm)	Φ[a]	solvent
1a	477	57286	485	19	0.89 ± 0.01	MeOH
1a	480	52809	491	21	n.d.	CH_2Cl_2
5a	426	37009	485	25	0.0007	MeOH
1b	490	58684	521	46	0.87 ± 0.01	MeOH
1b	494	42641	530	48	n.d.	CH_2Cl_2
5b	463	50452	517	64	0.006	MeOH
1c	492	57303	532	54	0.86 ± 0.02	MeOH
1c	498	44990	546	53	n.d.	CH_2Cl_2
5c	463	42669	530	54	0.01	MeOH
1d	493	51269	521	38	0.85 ± 0.03	MeOH
1d	496	47448	530	43	n.d.	CH_2Cl_2
5d	455	38642	528	69	0.03	MeOH
1e	495	37867	531	48	0.80 ± 0.01	MeOH
1e	498	44990	546	46	n.d.	CH_2Cl_2
1f	492	44003	598	120	0.0004	MeOH
1f	498	n.d	587	53	0.53 ± 0.01	Hexanes
1f	499	38658	642	150	n.d	CH_2Cl_2
5f	473	35502	527	73	0.005	MeOH
5f	486	n.d.	545	97	0.05	Hexanes
1g	488	48132	531	53	0.88 ± 0.01	MeOH
1g	481	46461	518	50	0.87 ± 0.01	Phos 7.4[b]
1h	488	35024	538	58	0.84 ± 0.01	MeOH
1h	482	34766	526	56	0.82 ± 0.01	Phos 7.4[b]
1s	482	n.d.	529	56	n.d.	Phos 7.4[b]

[a] Fluorescein in 0.1M NaOH as standard (Φ = 0.92); [b] 0.1M lithium phosphate buffer, pH = 7.4

APPENDIX C

EXPERIMENTAL DATA FOR CHAPTER III

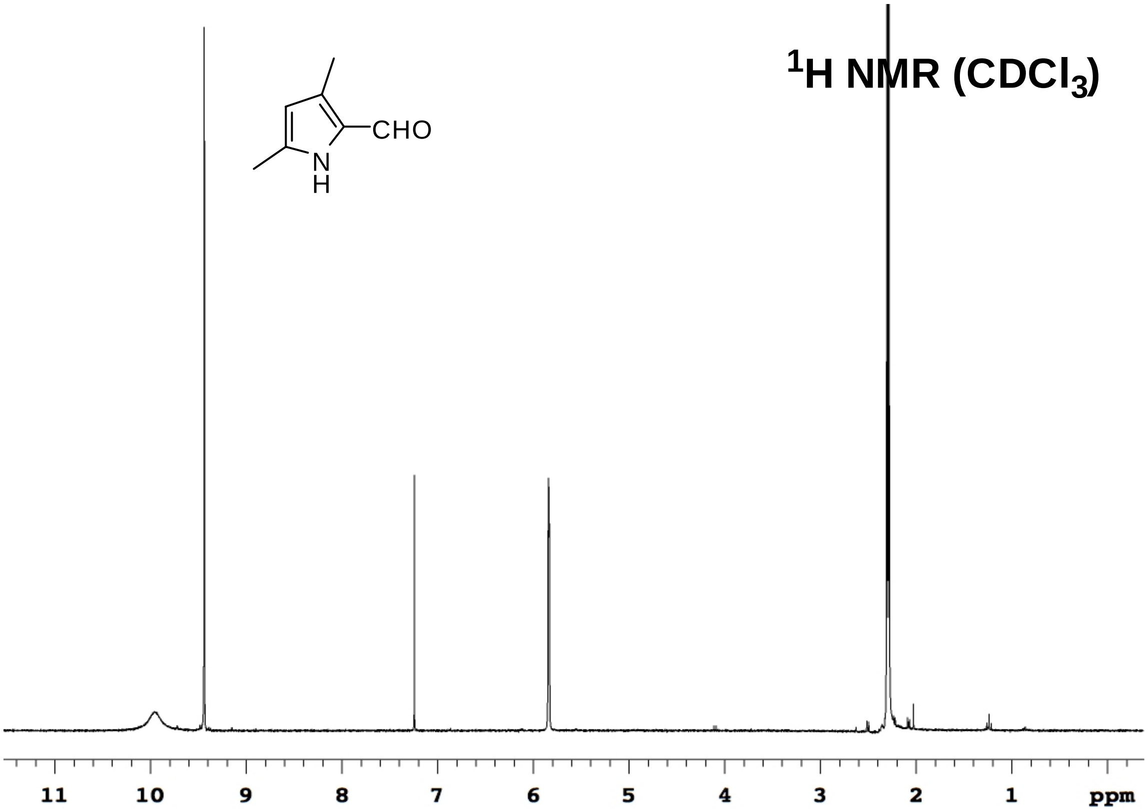
Typical procedure for the synthesis of pyrrole-2-carboxaldehyde[240]

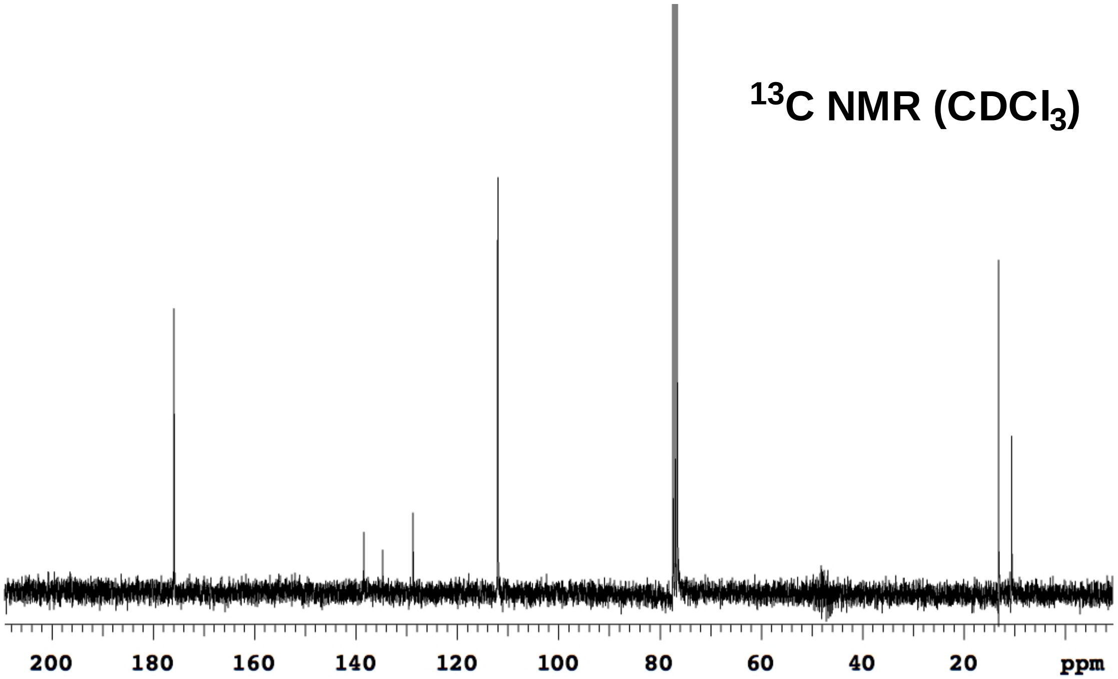


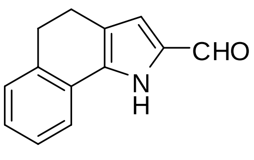
8a

$POCl_3$ (5.9 mL, 63.0 mmol) was added dropwise to DMF (4.9 mL, 63.0 mmol) at 0 °C. The mixture was warmed to room temperature and stirred for 15 min. The ice bath was replaced to cool the mixture back to 0 °C, then 30 mL of 1,2-dichloroethane was added to the mixture. A solution of 2,4-dimethyl pyrrole (5.0 g, 52.5 mmol) in 50 mL of 1,2-dichloroethane was added dropwise over 20 min at 0°C. After the addition was complete, the reaction mixture was refluxed for 30 min and then cooled to room temperature. A solution of NaOAc (23.7 g, 289 mmol) in 100 mL of water was added. The reaction mixture was again refluxed for 30 min. The cooled mixture was washed with water (1 x 100 mL), saturated Na_2CO_3 solution (2 x 50 mL) and brine (1 x 50 mL). The organic layer was dried over Na_2SO_4 and the solvents were removed under reduced pressure. The residue was purified by flash chromatography (SiO_2, 20 % EtOAc/Hexanes) to afford the pure product **8a** as a light yellow solid (5.8 g, 89 %). ^1H NMR (300 MHz, $CDCl_3$) δ 9.96 (br, 1H), 9.44 (s, 1H), 5.84 (d, 1H, *J* = 2.6 Hz), 2.30 (s, 3H), 2.28 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 175.9, 138.4, 134.7, 128.7, 112.0, 13.2, 10.6.

8b-**8h** were prepared using similar method as described for **8a**.


CHO
N
H
^{1}H NMR ($CDCl_3$)
11
10
9
8
7
6
5
4
3
2
1
ppm

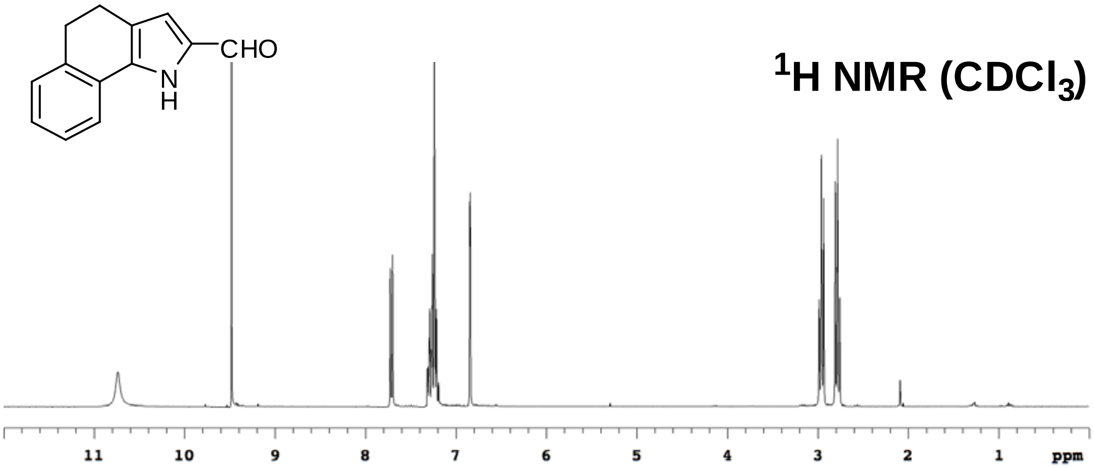

^{13}C NMR ($CDCl_3$)
200
180
160
140
120
100
80
60
40
20
ppm

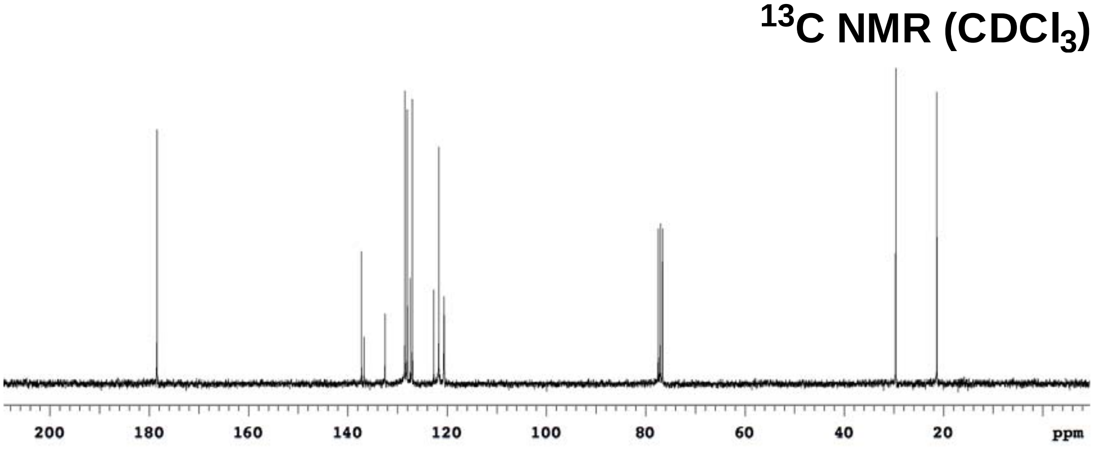


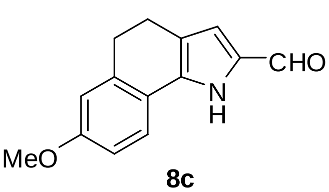
8b

4,5-Dihydro-1H-benzo[g]indole-2-carbaldehyde (8b)

Yellow solid (910 mg, 85 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 10.74 (br, 1H), 9.48 (s, 1H), 7.71 (d, 1H, *J* = 7.4 Hz), 7.32-7.19 (m, 3H), 6.84 (d, 1H, *J* = 2.0 Hz), 2.96 (t, 2H, *J* = 7.5 Hz), 2.78 (t, 2H, *J* = 7.5 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 178.4, 137.2, 136.7, 132.5, 128.5, 128.0, 127.4, 127.0, 122.7, 121.7, 120.6, 29.6, 21.3. MS (ESI) m/z calcd for $(M+H)^+$ $C_{13}H_{12}NO$ 198.09; found 198.09.



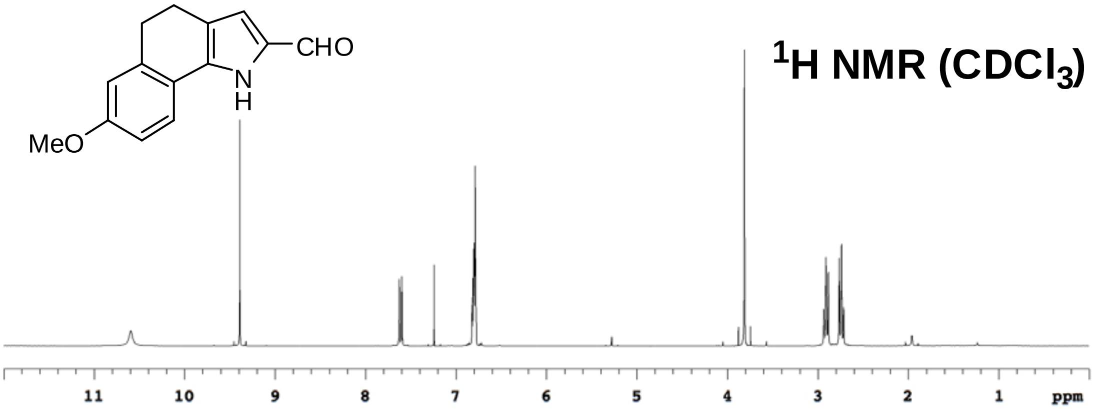


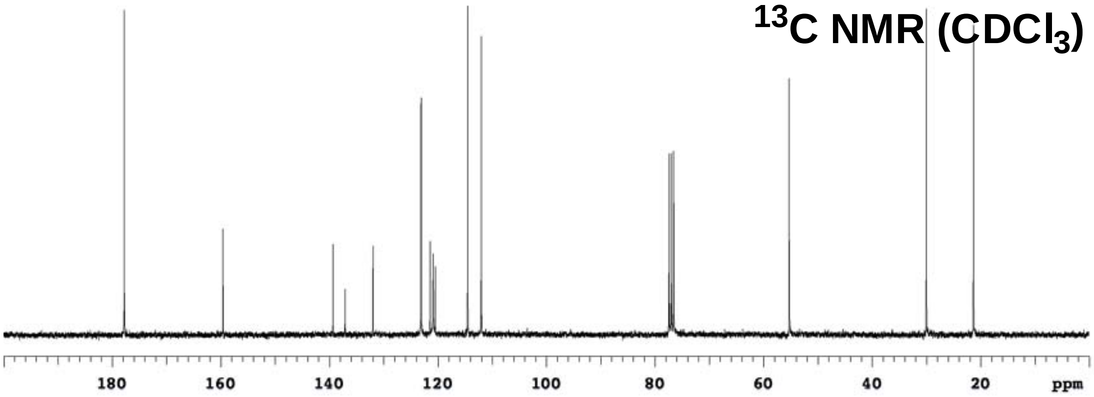


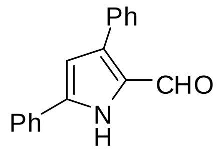
8c

7-Methoxy-4,5-dihydro-1H-benzo[g]indole-2-carbaldehyde (8c)

Light green solid (345 mg, 100 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 10.59 (br, 1H), 9.39 (s, 1H), 7.61 (d, 1H, *J* = 8.2 Hz), 6.82-6.78 (m, 3H), 3.81 (s, 3H), 2.91 (t, 2H, *J* = 7.5 Hz), 2.74 (t, 2H, *J* = 7.5 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 177.8, 159.6, 139.3, 137.1, 132.0, 123.1, 121.4, 120.9, 120.5, 114.5, 112.0, 55.3, 30.0, 21.3. MS (ESI) m/z calcd for (M+H)$^+$ C$_{14}$H$_{14}$NO$_2$ 228.10; found 228.10.



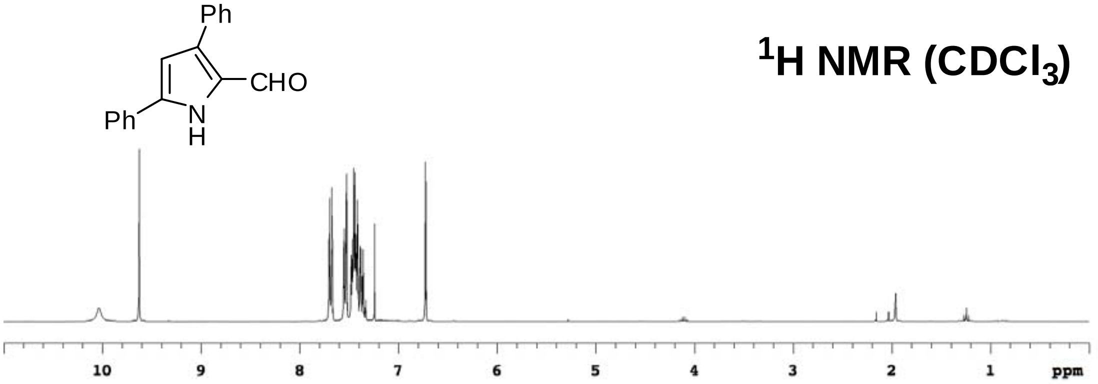


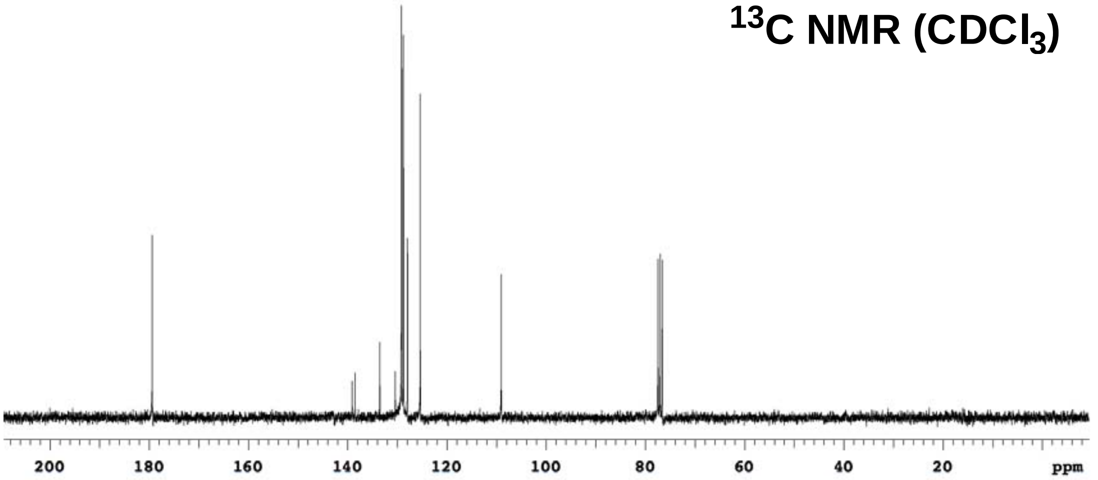


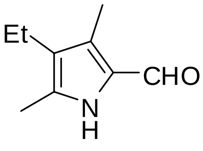
8d

3,5-Diphenyl-1H-pyrrole-2-carbaldehyde (8d)

White solid (290 mg, 86 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 10.04 (br, 1H), 9.63 (s, 1H), 7.70-7.67 (m, 2H), 7.55-7.52 (m, 2H), 7.48-7.33 (m, 6H), 6.72 (d, 1H, *J* = 2.8 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 179.4, 139.1, 138.5, 133.5, 130.4, 129.2, 129.1, 129.0, 128.8, 128.7, 127.9, 125.3, 109.0. MS (ESI) m/z calcd for (M+H)$^+$ $C_{17}H_{14}NO$ 248.11; found 248.11.



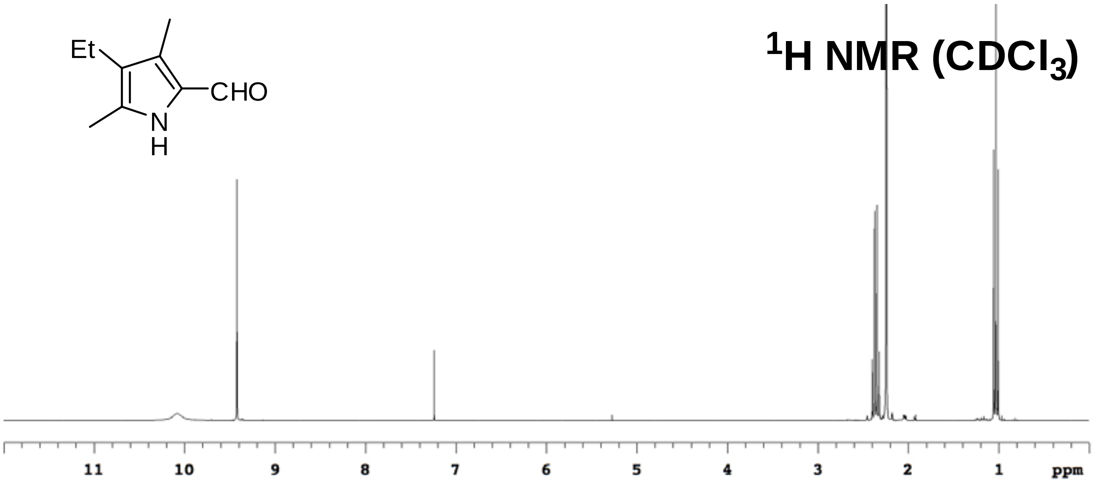


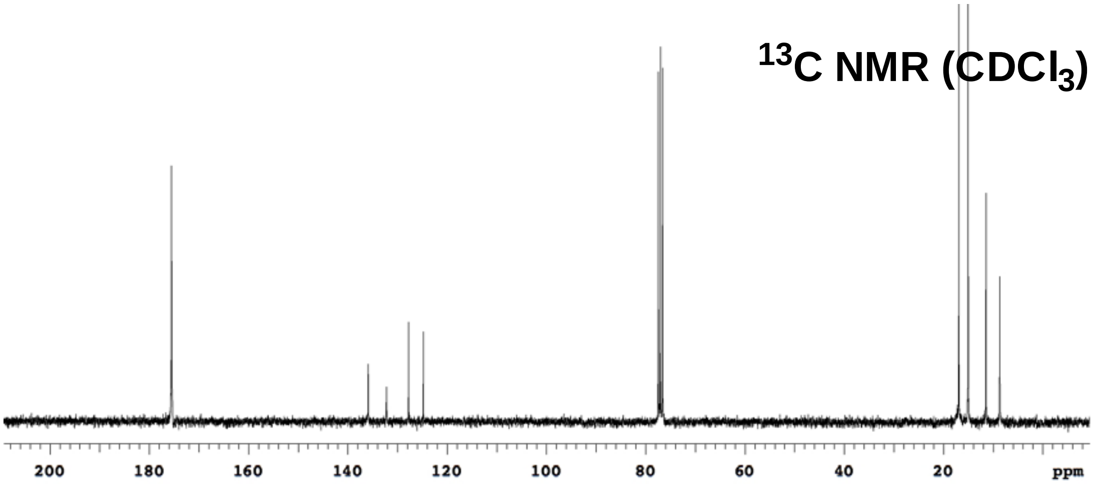


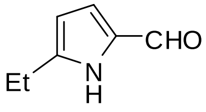
8e

4-Ethyl-3,5-dimethyl-1H-pyrrole-2-carbaldehyde (8e)

Brown solid (1.1 g, 85 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 10.08 (br, 1H), 9.42 (s, 1H), 2.36 (q, 2H, *J* = 7.5 Hz), 2.25 (s, 3H), 2.24 (s, 3H), 1.03 (t, 3H, *J* = 7.5 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 175.5, 135.9, 132.2, 127.8, 124.8, 16.9, 15.0, 11.4, 8.7. MS (ESI) m/z calcd for $(M+H)^+$ $C_9H_{14}NO$ 152.11; found 152.10.



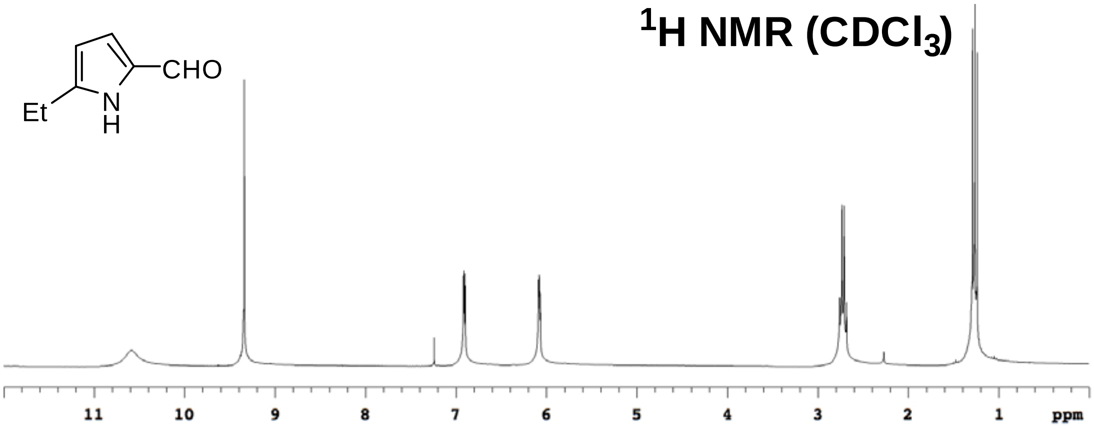


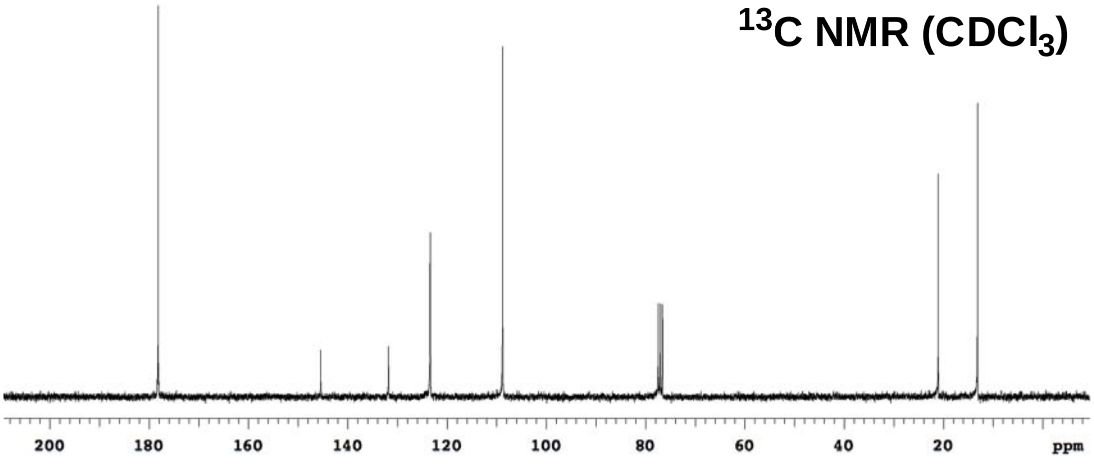


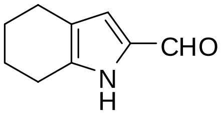
8f

5-Ethyl-1H-pyrrole-2-carbaldehyde (8f)

Red solid (1.0 g, 78 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.58 (br, 1H), 9.34 (s, 1H), 6.92-6.90 (m, 1H), 6.09-6.07 (m, 1H), 2.72 (q, 2H, *J* = 7.5 Hz), 1.26 (t, 3H, *J* =7.5 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 178.2, 145.4, 131.8, 123.4, 108.8, 21.0, 13.1. MS (ESI) m/z calcd for $(M+H)^+$ $C_7H_{10}NO$ 124.08; found 124.07.



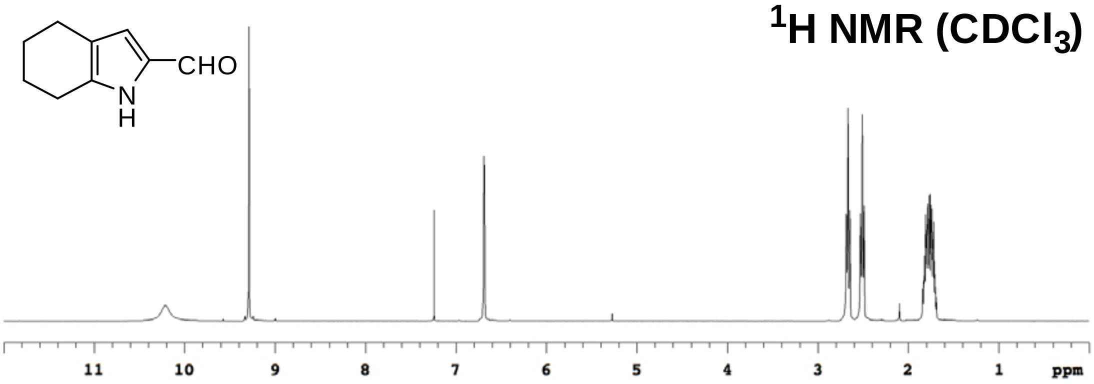


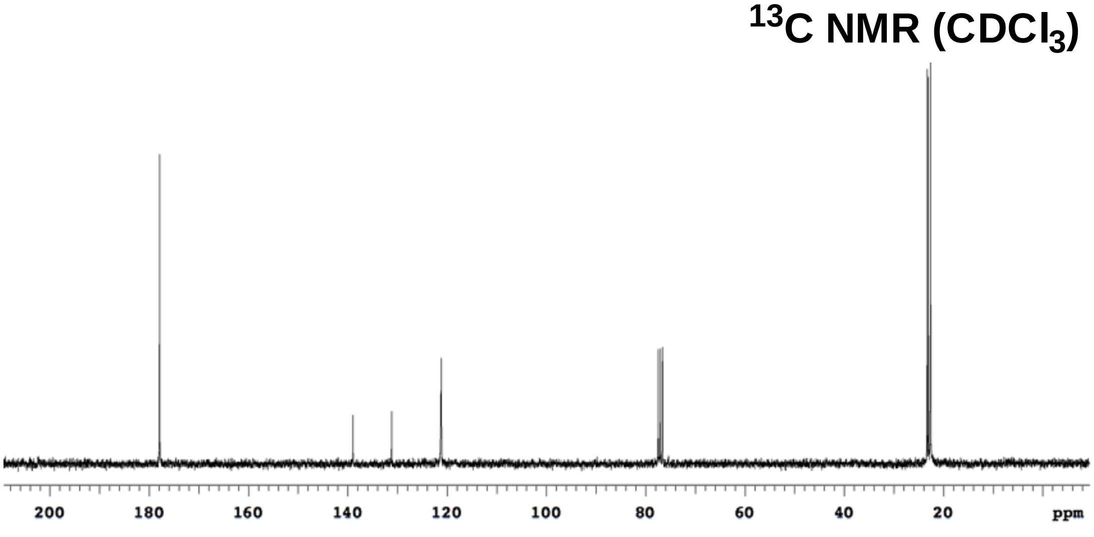


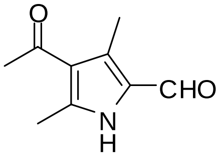
8g

4,5,6,7-Tetrahydro-1H-indole-2-carbaldehyde (8g)

Yellow solid (960 mg, 78 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 10.21 (br, 1H), 9.29 (s, 1H), 6.69 (d, 1H, *J* = 2.4 Hz), 2.67 (t, 2H, *J* = 6.1 Hz), 2.51 (t, 2H, *J* = 6.0 Hz), 1.84-1.69 (m, 4H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 177.9, 138.9, 131.2, 121.2, 121.1, 23.3, 23.0, 22.6 (2C: 22.61, 22.59). MS (ESI) m/z calcd for (M+H)$^+$ $C_9H_{12}NO$ 150.09; found 150.09.



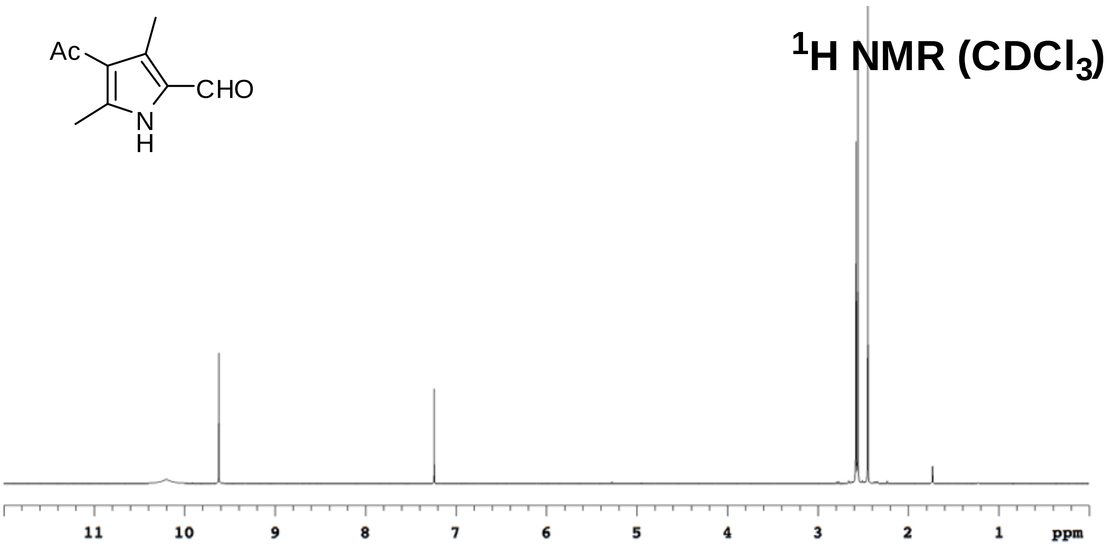


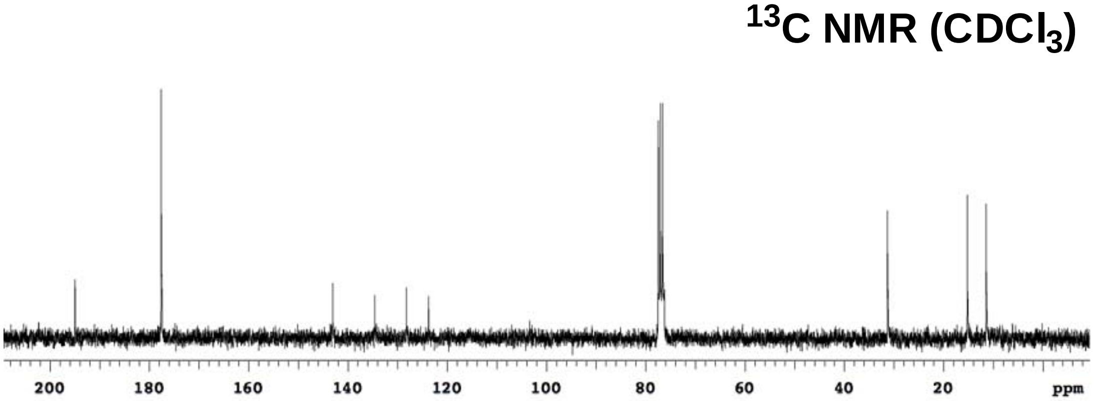


8h

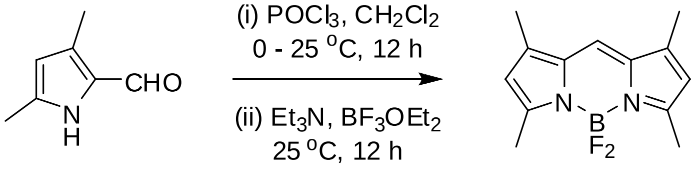
4-Acetyl-3,5-dimethyl-1H-pyrrole-2-carbaldehyde (8h)

Light yellow solid (481 mg, 40 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 10.20 (br, 1H), 9.62 (s, 1H), 2.58 (s, 3H), 2.56 (s, 3H), 2.45 (s, 3H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 194.9, 177.6, 143.1, 134.6, 128.2, 123.7, 31.3, 15.2, 11.4. MS (ESI) m/z calcd for (M+H)$^+$ $C_9H_{12}NO_2$ 166.09; found 166.09.





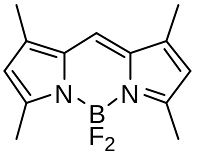
Typical procedure for the synthesis of symmetric BODIPYs



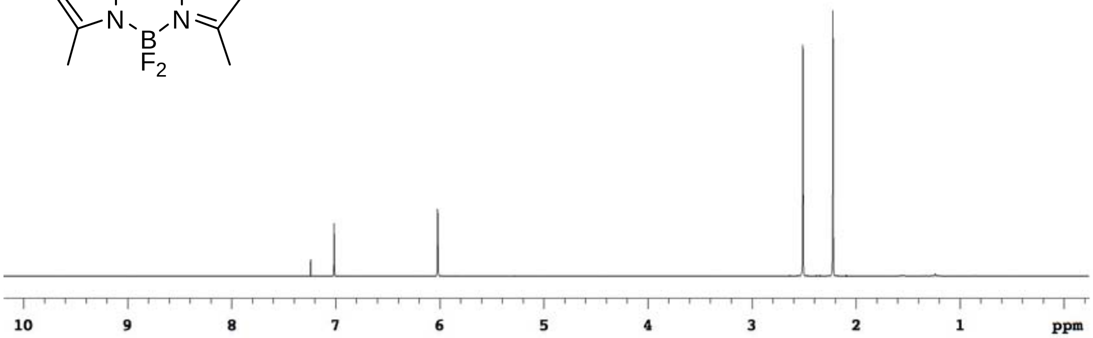
8a **9a**

3,5-Dimethyl-1H-pyrrole-2-carbaldehyde **8a** (246 mg, 2 mmol) was dissolved in 10 mL CH_2Cl_2 and $POCl_3$ (0.22 mL, 2.4 mmol) was added dropwise over 1 min at 0 ^{o}C. The solution was warmed to room temperature slowly and stirred for 12 h. The mixture was cooled to 0 ^{o}C and Et_3N (1.4 mL, 10 mmol) was added dropwise over 5 min. After stirring for 15 min, BF_3OEt_2 (2.0 mL, 16 mmol) was added dropwise to the solution over 5 min. The reaction mixture was warmed to room temperature and stirred for 12 h. The mixture was passed through a short pad of silica gel eluting with CH_2Cl_2 to remove the polar impurities. The solvents were removed under reduced pressure. The residue was dissolved in CH_2Cl_2 and water was added, and the mixture was stirred at room temperature overnight. (to decompose excess BF_3OEt_2 and other impurities). The organic layer was washed with water, brine and dried over Na_2SO_4. The solvent was removed under reduced pressure and the residue was purified by flash chromatography (5 % EtOAc/hexanes) to give the pure product **9a** (229 mg, 92 %) as a red solid. ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.01 (s, 1H), 6.02 (s, 2H), 2.51 (s, 6H), 2.22 (s, 6H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 156.7, 141.2, 133.4, 120.0, 119.0, 14.6, 11.2; ^{19}F NMR ($CDCl_3$, 282 MHz) δ 30.68 (q, *J* = 33.5 Hz). HRMS (ESI) m/z calcd for $(M+H)^+$ $C_{13}H_{16}BF_2N_2$ 249.1375; found 249.1373.

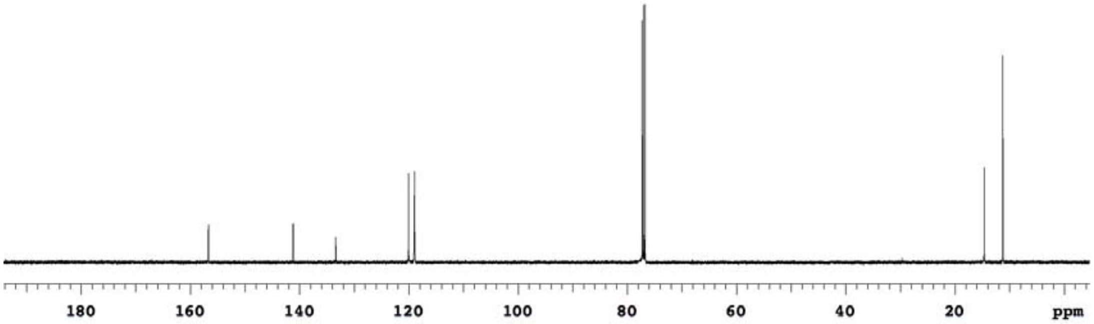
9b-**9h** were prepared using similar methods as described for **9a**.



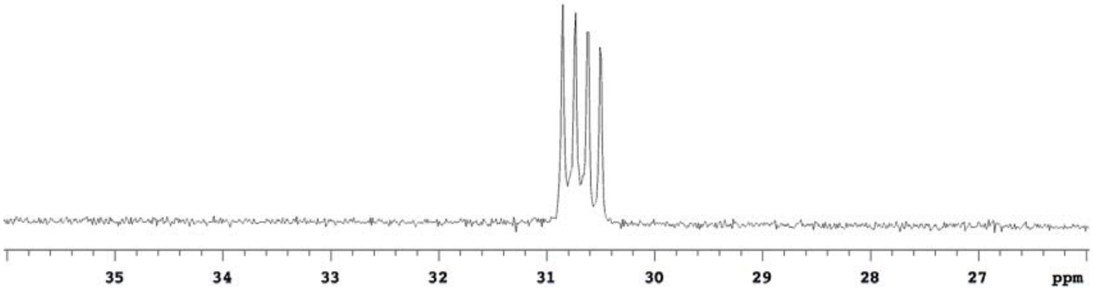
^{1}H NMR (CDCl$_3$)

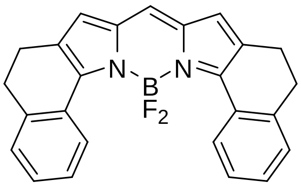


^{13}C NMR (CDCl$_3$)



^{19}F NMR (CDCl$_3$)

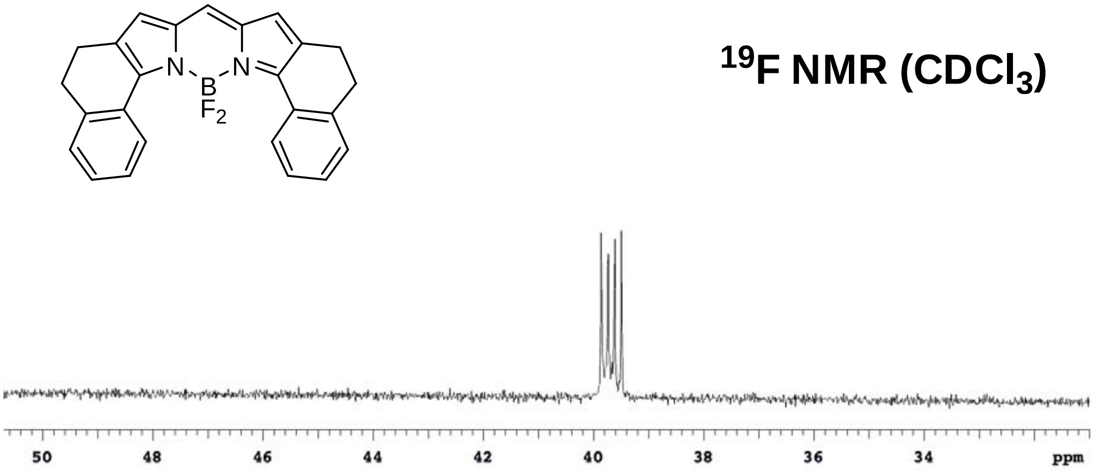


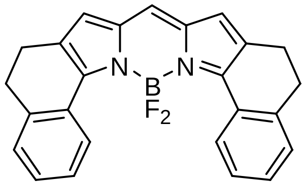


9b

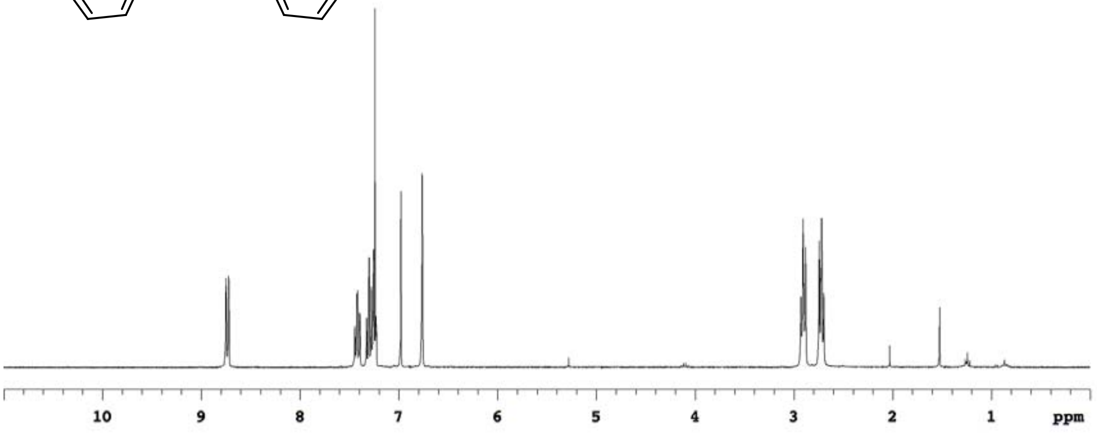
BODIPY 9b

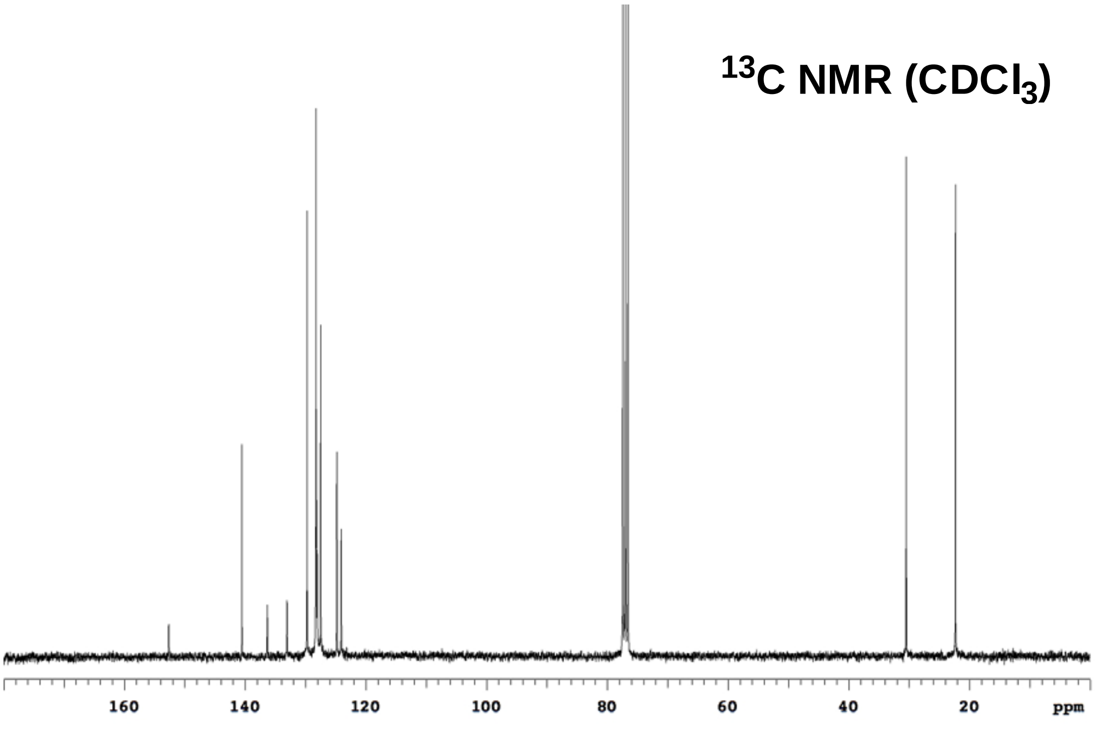
Green solid (360 mg, 91 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.73 (d, 2H, *J* = 8.0 Hz), 7.45-7.39 (m, 2H), 7.33-7.23 (m, 4H), 6.98 (s, 1H), 6.76 (s, 2H), 2.90 (t, 4H, *J* = 7.0 Hz), 2.72 (t, 4H, *J* = 7.0 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 152.6, 140.5, 136.3, 133.1, 129.7, 128.3, 128.2, 128.1, 127.5, 124.8, 124.1, 30.5, 22.3; ^{19}F NMR ($CDCl_3$, 282 MHz) δ 39.67 (q, *J* = 33.6 Hz). MS (MALDI) m/z calcd for M^+ $C_{25}H_{19}BF_2N_2$ 396.16; found 395.88.



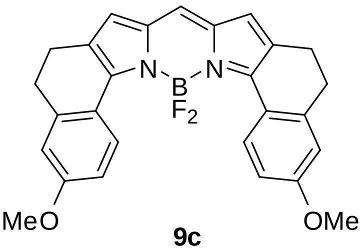


^{1}H NMR (CDCl$_3$)





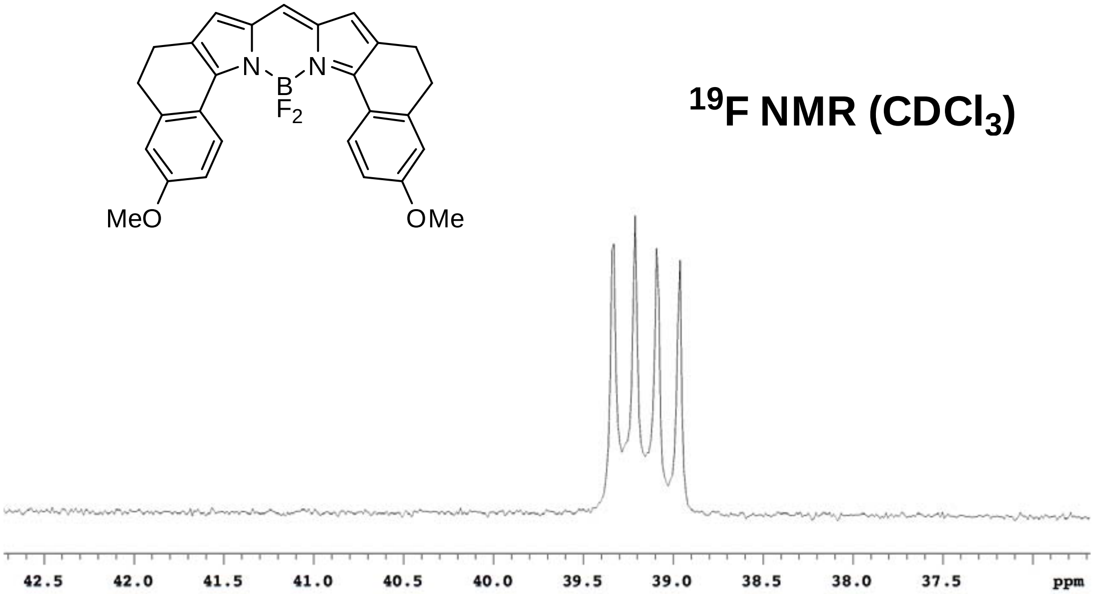
^{13}C NMR (CDCl$_3$)

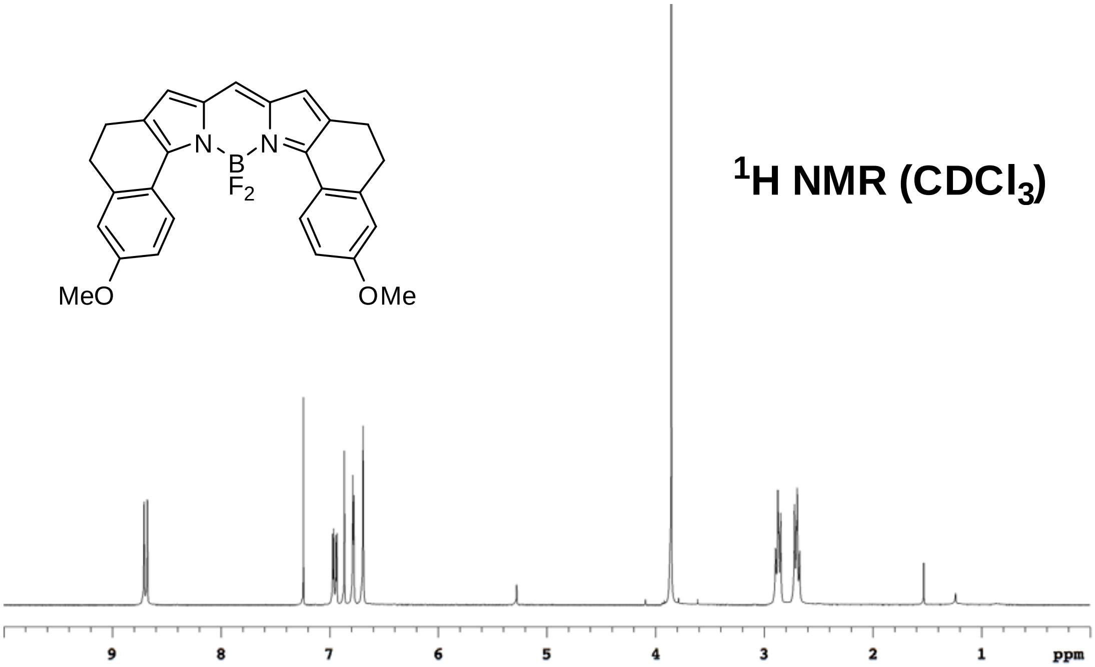


9c

BODIPY 9c

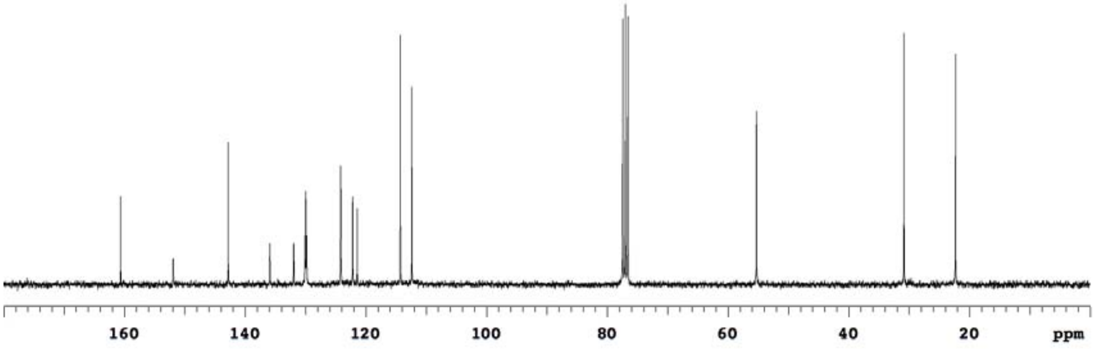
Blue solid (194 mg, 85 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.69 (d, 2H, *J* = 9.0 Hz), 6.95 (dd, 2H, *J* = 9.0, 2.8 Hz), 6.86 (s, 1H), 6.78 (d, 2H, *J* = 2.8 Hz), 6.69 (s, 2H), 3.85 (s, 6H), 2.87 (t, 4H, *J* = 6.9 Hz), 2.70 (t, 4H, *J* = 6.9 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 160.6, 151.9, 142.8, 135.9, 131.9, 129.9 (t, *J* = 11.1 Hz), 124.2, 122.2, 121.4, 114.3, 112.4, 55.3, 30.8, 22.3; ^{19}F NMR ($CDCl_3$, 282 MHz) δ 39.15 (q, *J* = 33.6 Hz). MS (MALDI) m/z calcd for M^+ $C_{27}H_{23}BF_2N_2O_2$ 456.18; found 455.88.

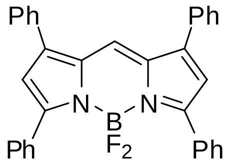




^{1}H NMR (CDCl$_3$)

^{13}C NMR (CDCl$_3$)

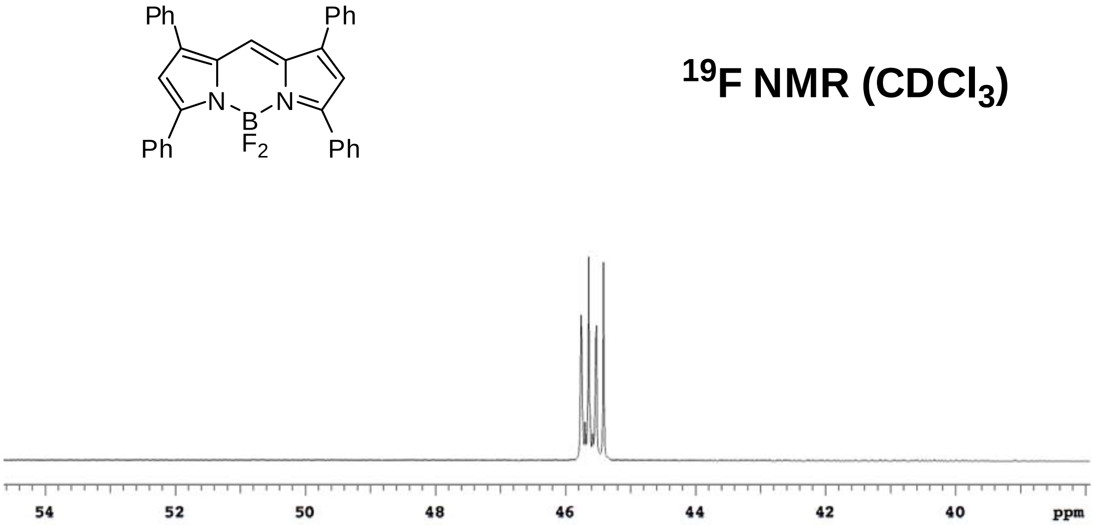


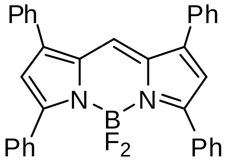


9d

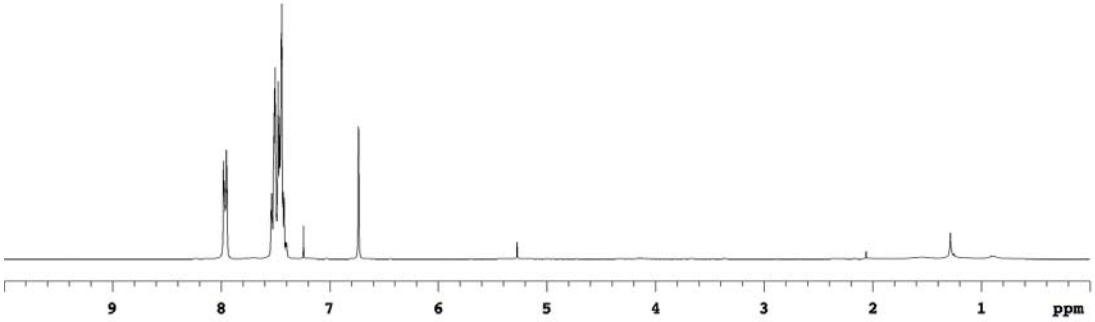
BODIPY 9d

Green solid (52 mg, 21 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.98-7.95 (m, 4H), 7.54-7.42 (m, 17H), 6.73 (s, 2H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 157.8, 145.7, 134.4, 133.3, 132.3, 129.7, 129.4 (t, *J* = 3.5 Hz), 129.1, 128.7 (2C), 128.2, 127.7, 119.0; ^{19}F NMR (CDCl$_3$, 282 MHz) δ 45.58 (q, *J* = 33.5 Hz). MS (MALDI) m/z calcd for M$^+$ $C_{33}H_{23}BF_2N_2$ 496.19; found 495.93.

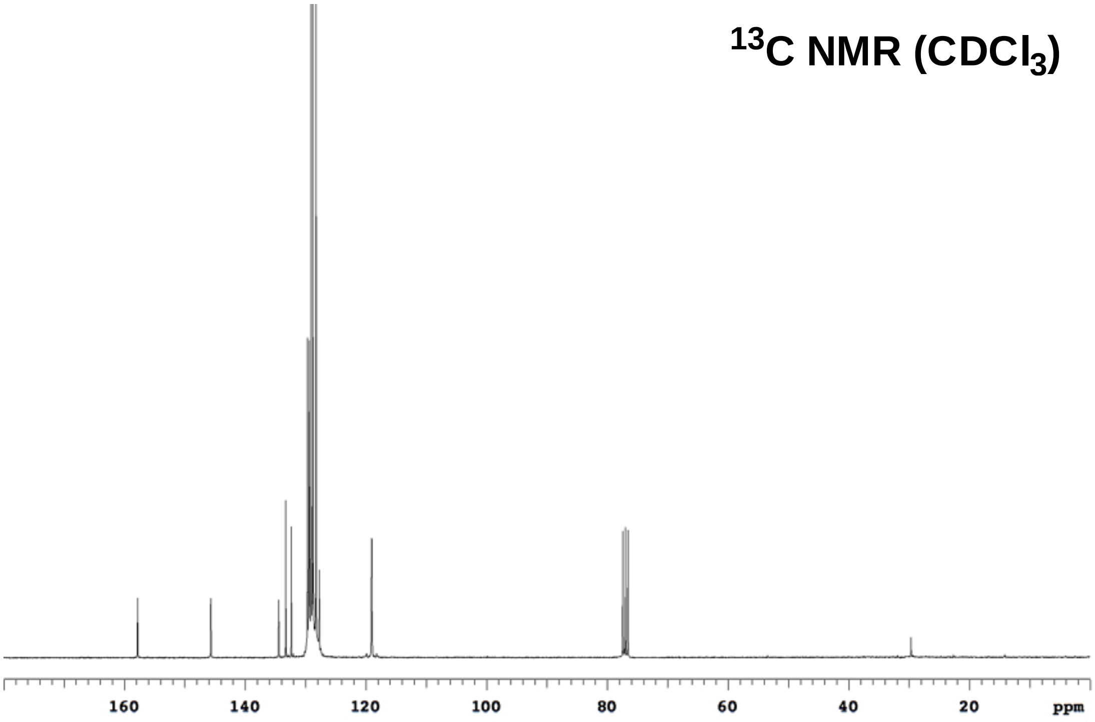


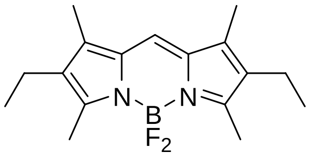

Ph
Ph
N
B
N
F2
Ph
Ph

^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

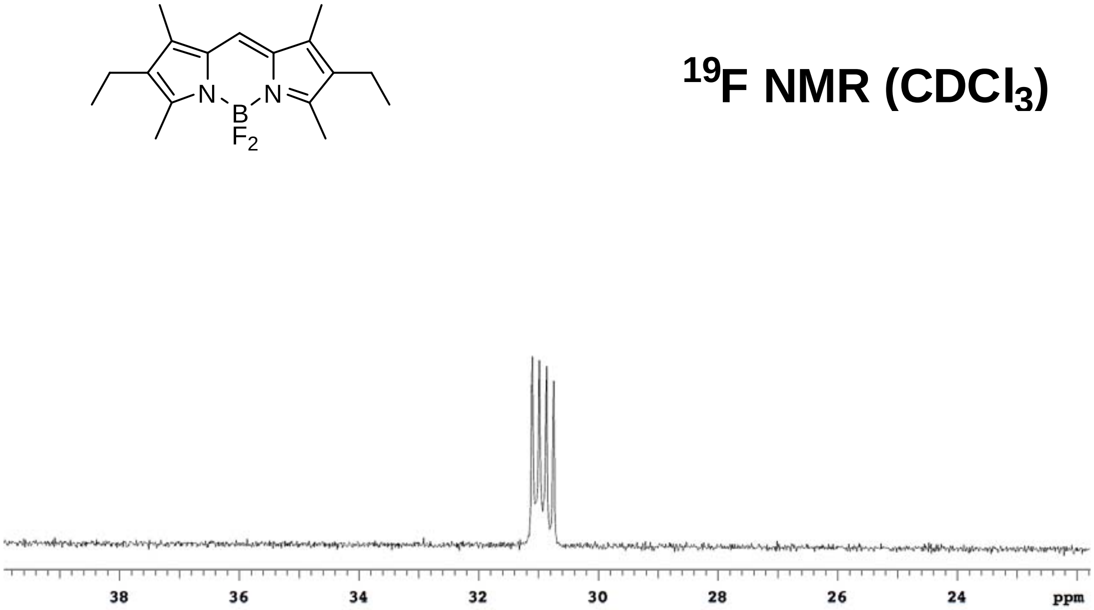


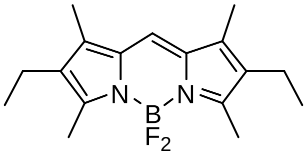


9e

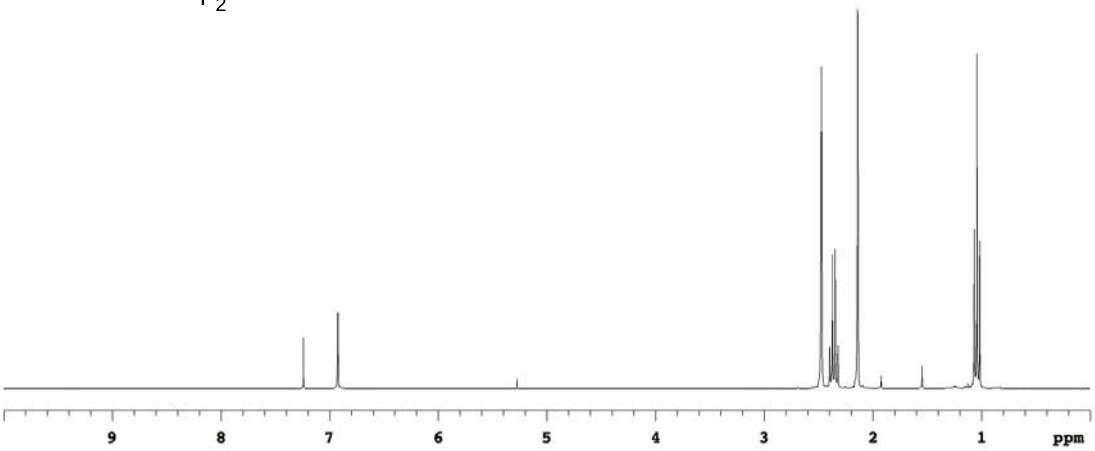
BODIPY 9e

Organge solid (227 mg, 75 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 6.92 (s, 1H), 2.47 (s, 6H), 2.36 (q, 4H, *J* = 7.5 Hz), 2.14 (s, 6H), 1.04 (t, 6H, *J* = 7.5 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 154.6, 136.6, 132.4, 131.6, 118.5, 17.3, 14.6, 12.5, 9.4; ^{19}F NMR ($CDCl_3$, 282 MHz) δ 30.93 (q, *J* = 33.6 Hz). MS (MALDI) m/z calcd for M^+ $C_{17}H_{23}BF_2N_2$ 304.19; found 303.96.

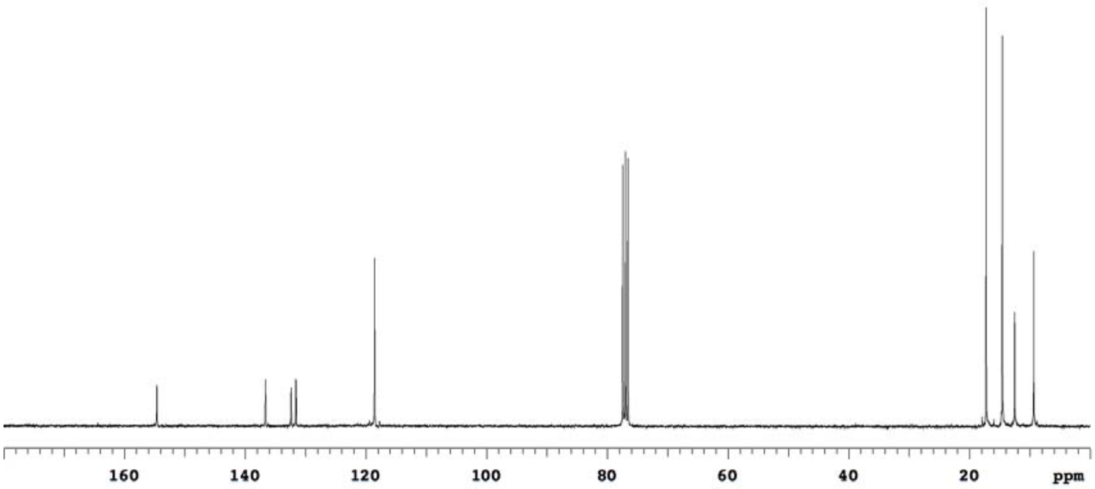


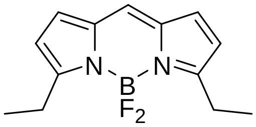


^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

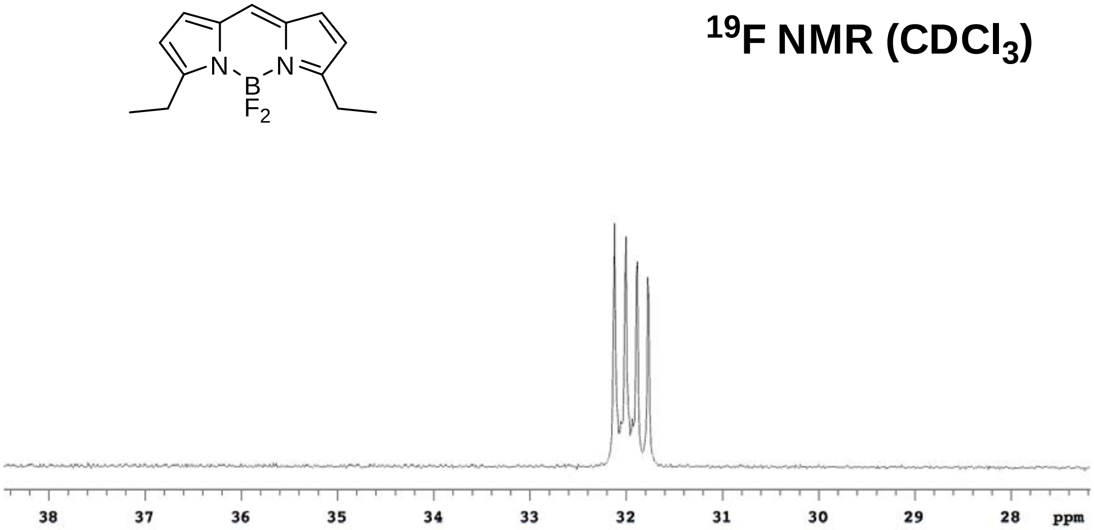


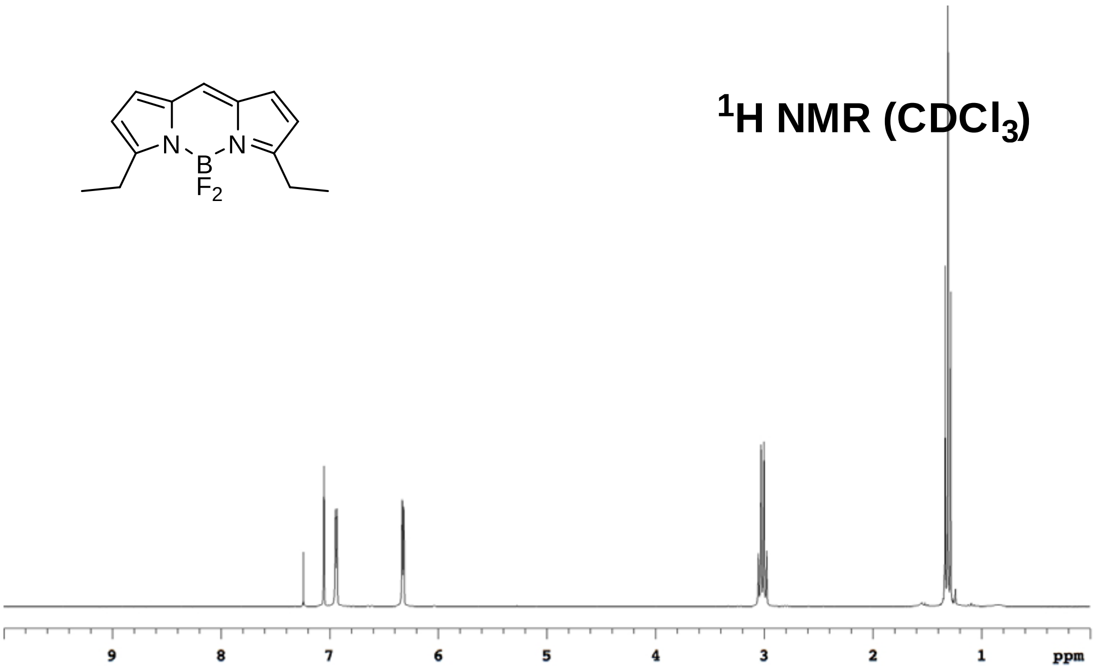


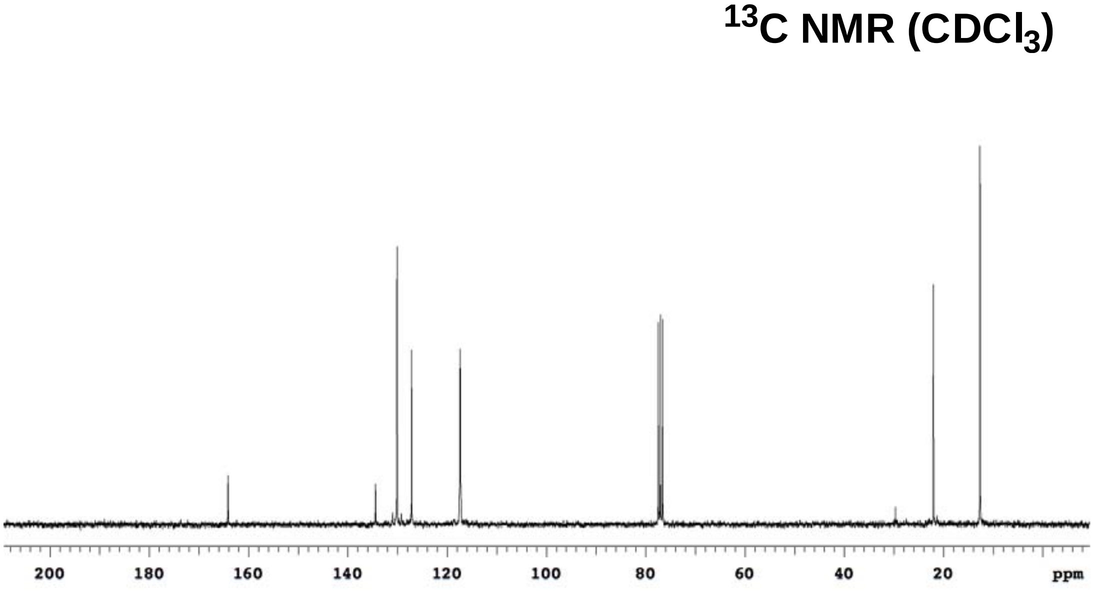
9f

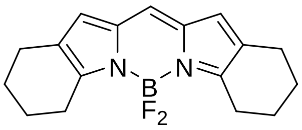
BODIPY 9f

Orange solid (70 mg, 28 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.05 (s, 1H), 6.94 (d, 2H, *J* = 4.2 Hz), 6.32 (d, 2H, *J* = 4.2 Hz), 3.02 (q, 4H, *J* = 7.5 Hz), 1.31 (t, 6H, *J* = 7.5 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 164.1, 134.4, 130.1, 127.1, 117.4, 22.0, 12.6; ^{19}F NMR (CDCl$_3$, 282 MHz) δ 31.94 (q, *J* = 33.4 Hz). MS (MALDI) m/z calcd for M$^+$ $C_{13}H_{15}BF_2N_2$ 248.13; found 247.85.




N
B
N
F2
1H NMR (CDCl3)
9
8
7
6
5
4
3
2
1
ppm

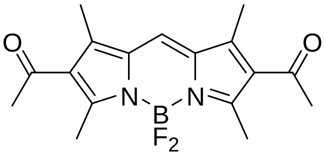

13C NMR (CDCl3)
200
180
160
140
120
100
80
60
40
20
ppm



9g

BODIPY 9g

Red solid (65 mg, 22 %). ^{1}H NMR (300 MHz, CDCl$_3$) δ 6.89 (s, 1H), 6.59 (s, 2H), 3.01 (t, 4H, *J* = 6.2 Hz), 2.52 (t, 4H, *J* = 6.2 Hz), 1.86-1.78 (m, 4H), 1.76-1.68 (m, 4H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 158.0, 134.0, 129.3, 125.9, 125.3, 24.7, 23.1, 22.8, 22.3; ^{19}F NMR (CDCl$_3$, 282 MHz) δ 27.29 (q, *J* = 33.5 Hz). MS (MALDI) m/z calcd for M$^+$ $C_{17}H_{19}BF_2N_2$ 300.16; found 299.91.

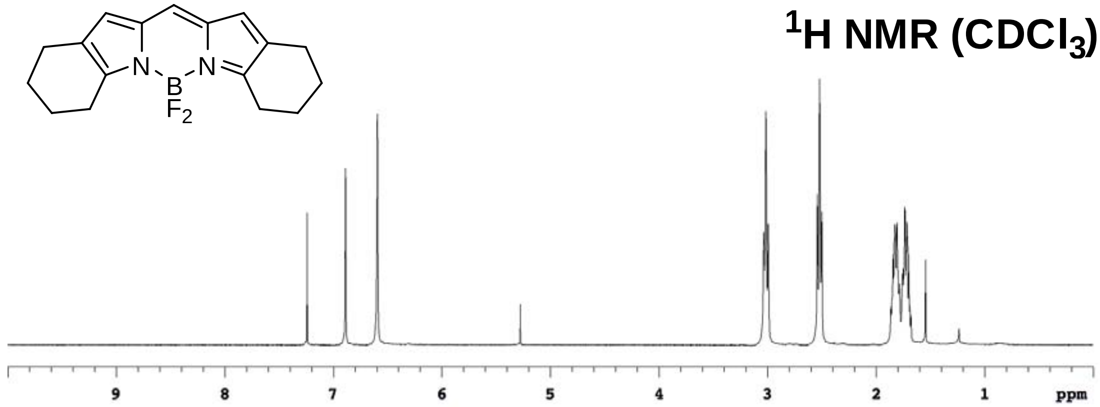


9h

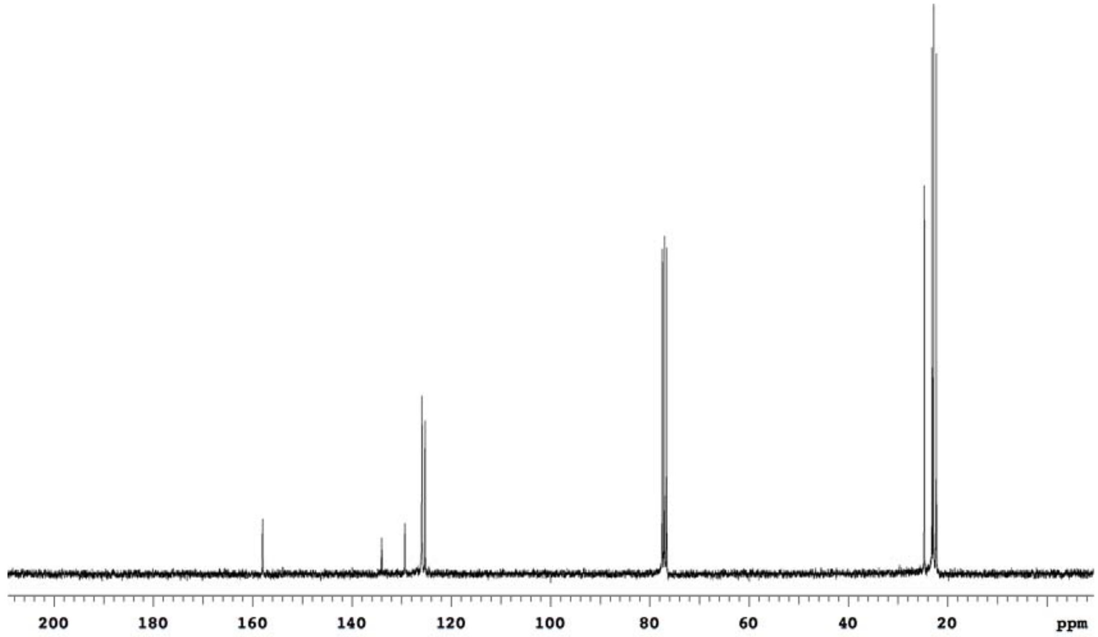
BODIPY 9h

Messy reaction, obtained complex mixtures.

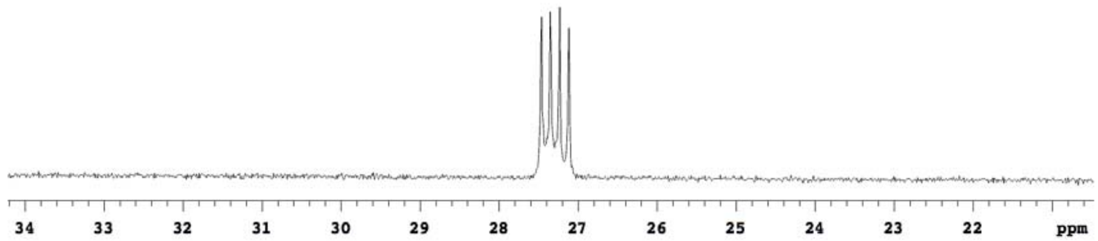
^{1}H NMR ($CDCl_3$)

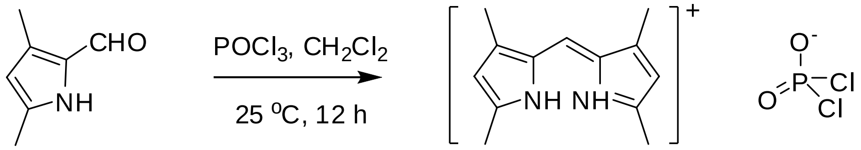


^{13}C NMR ($CDCl_3$)



^{19}F NMR ($CDCl_3$)

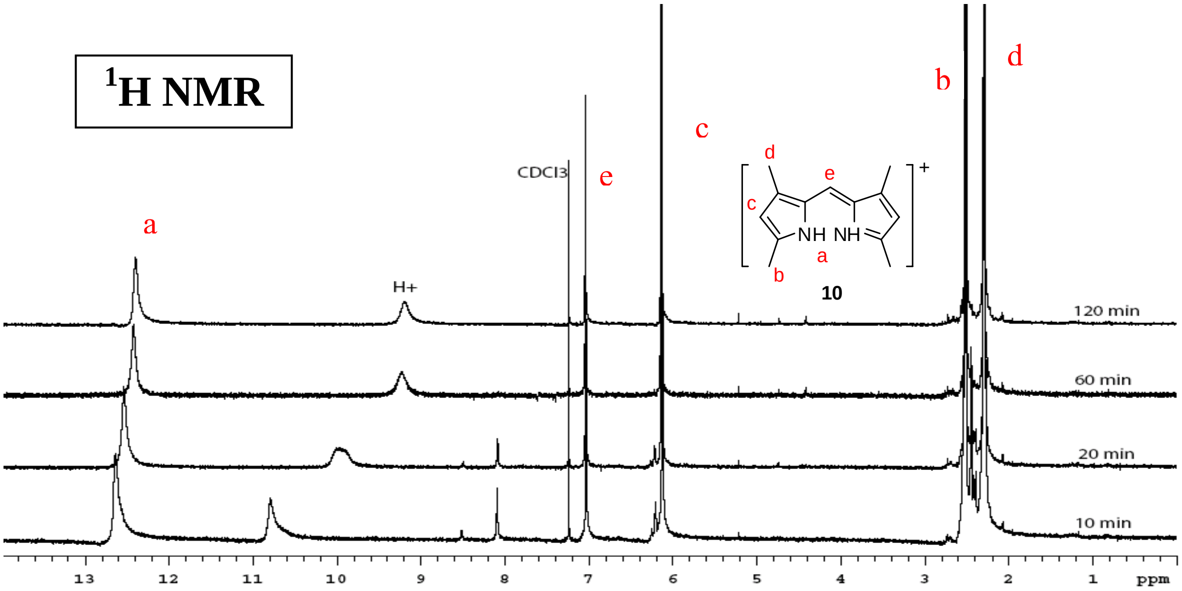


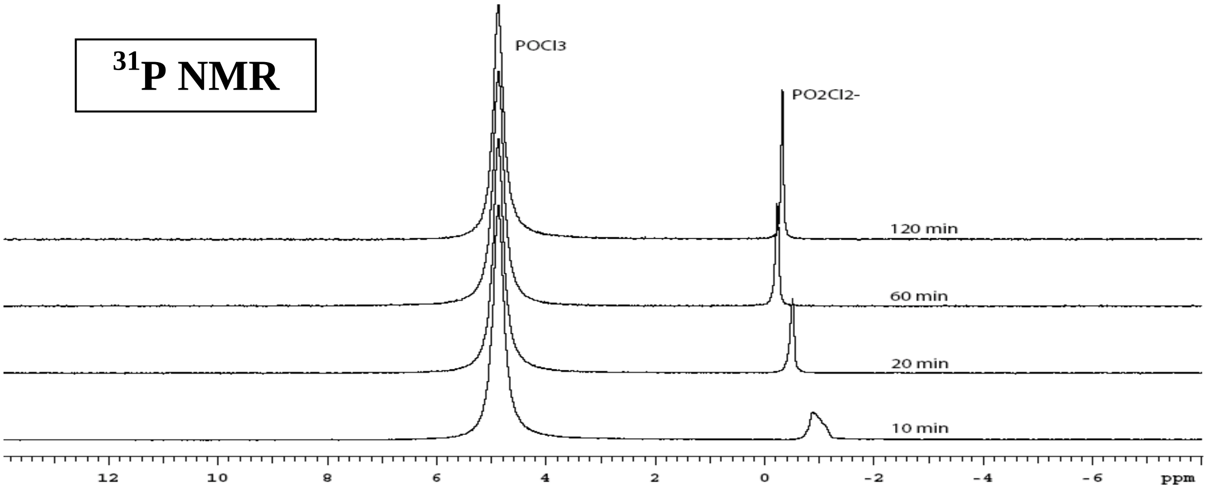

CHO
NH
POCl3, CH2Cl2
25 °C, 12 h
NH NH
+
O-
P
Cl
Cl
O

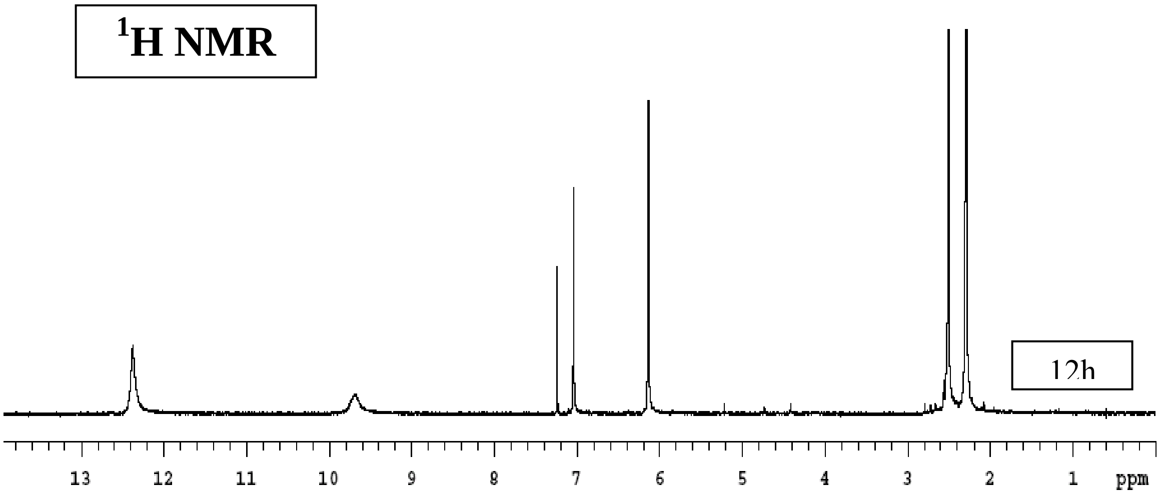
8a **10**

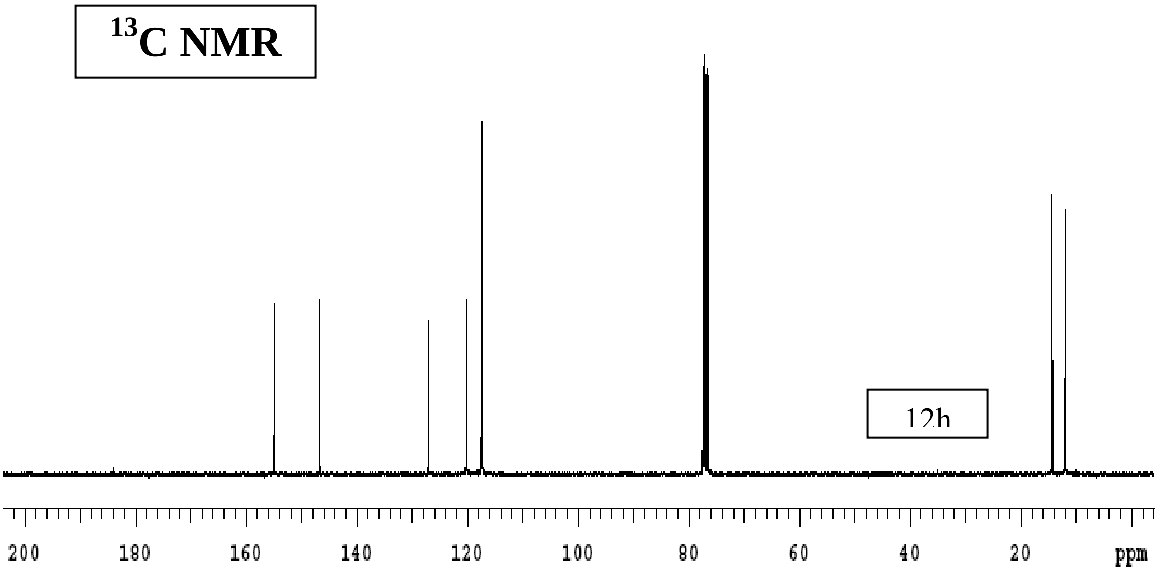
NMR and MS study of the self-condensation reaction:

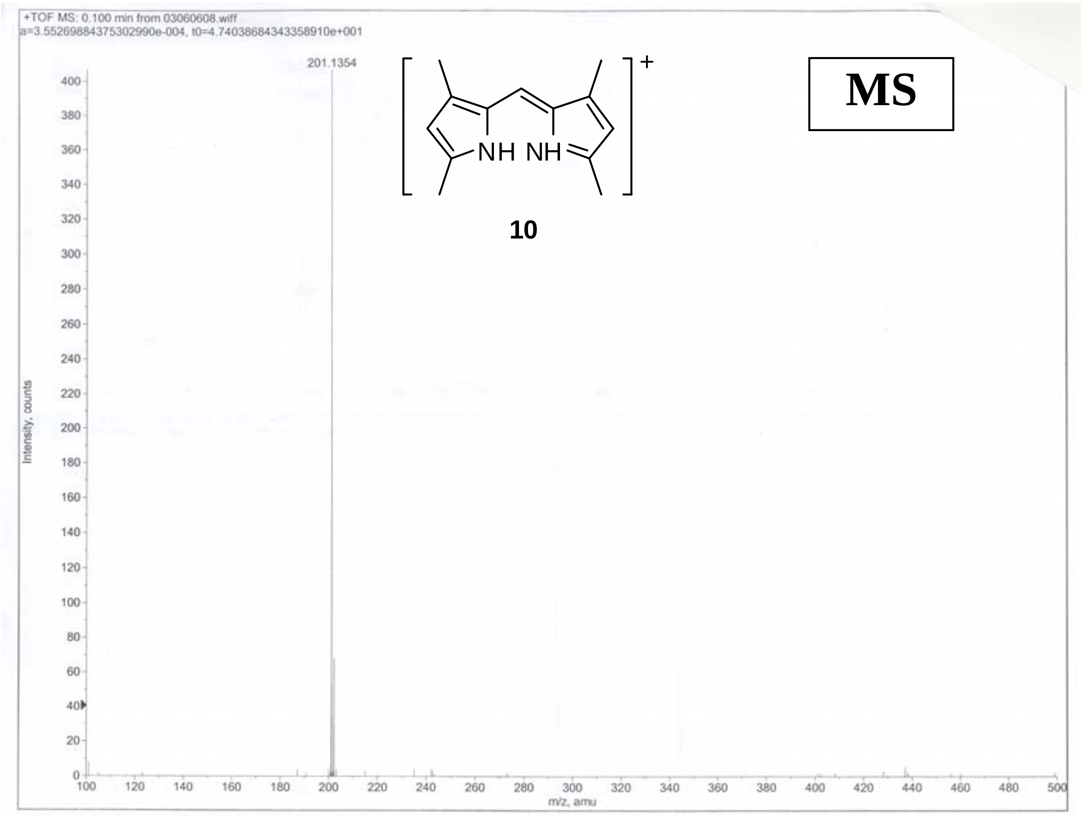
Compound **8a** (25 mg, 0.2 mmol) was dissolved in 1 mL of $CDCl_3$ and then $POCl_3$ (50 μL) was added. The reaction was monitored by ^1H, ^{13}C and ^{31}P over 10 min intervals at room temperature.



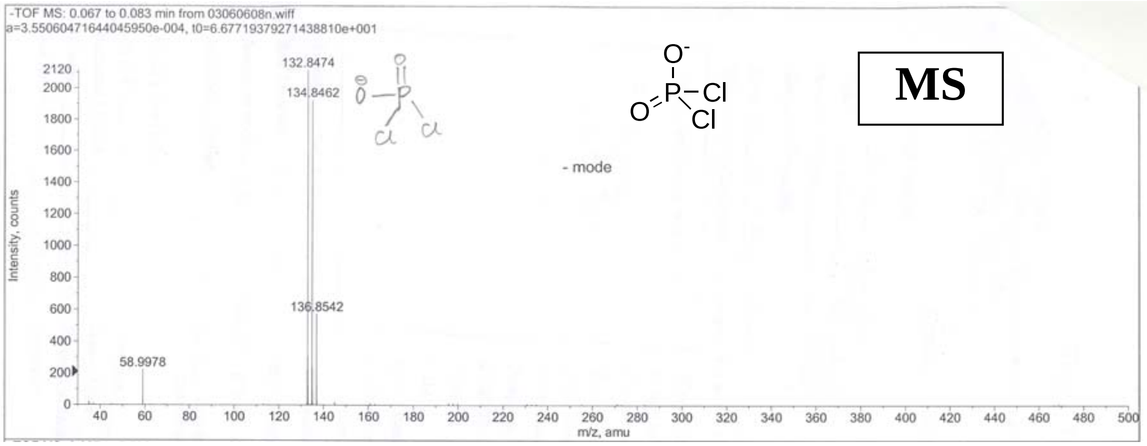



1H NMR
12h
13
12
11
10
9
8
7
6
5
4
3
2
1
ppm


13C NMR
12h
200
180
160
140
120
100
80
60
40
20
ppm


+TOF MS: 0.100 min from 03060608.wiff
a=3.55269884375302990e-004, t0=4.74038684343358910e+001
201.1354
NH
NH
+
MS
Intensity, counts
m/z, amu
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
400
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
400
420
440
460
480
500

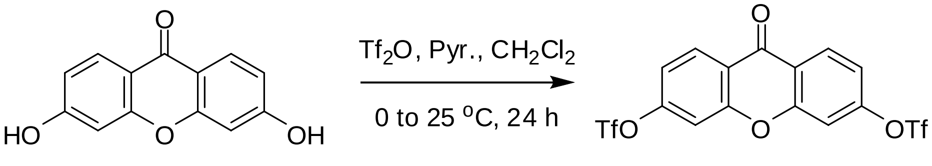
10


-TOF MS: 0.067 to 0.083 min from 03060608n.wiff
a=3.55060471644045950e-004, t0=6.67719379271438810e+001
132.8474
134.8462
136.8542
58.9978
O
P
Cl
Cl
- mode
MS
Intensity, counts
m/z, amu
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2120
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
400
420
440
460
480
500

APPENDIX D

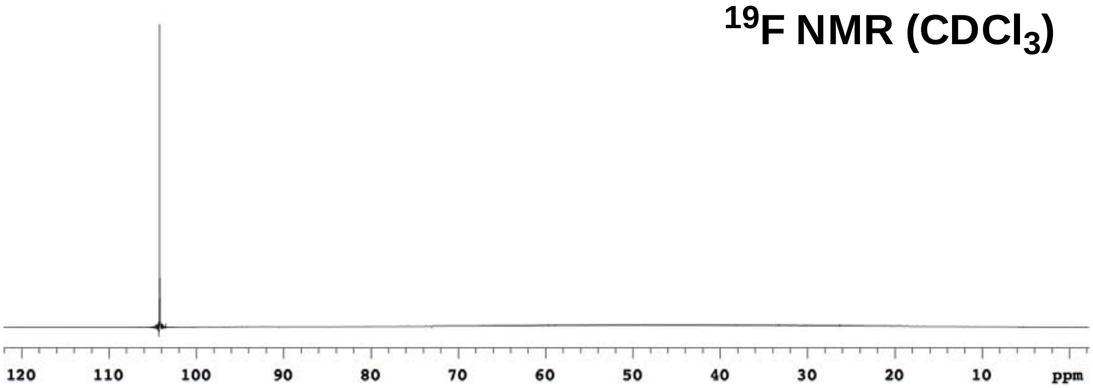
EXPERIMENTAL DATA FOR CHAPTER IV

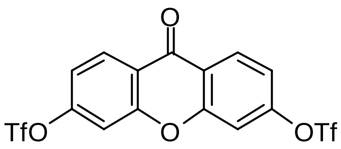
Synthesis of 3,6-ditrifylxanthone 16



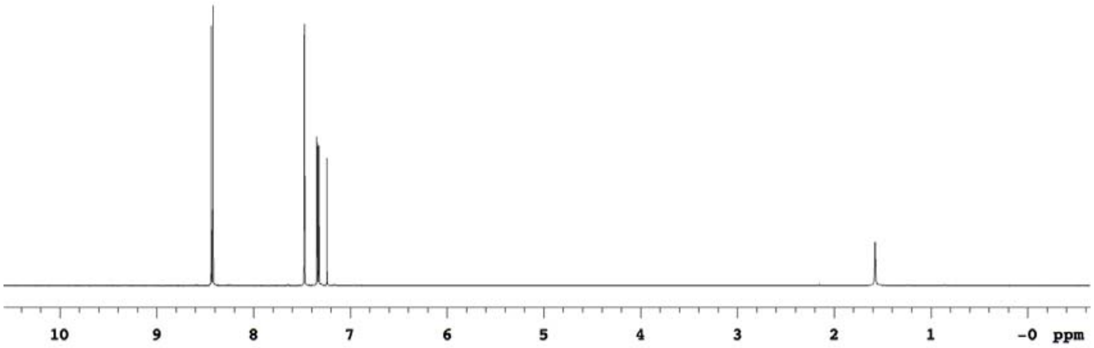
16

3,6-Dihydroxy-xanthen-9-one (6.85 g, 30 mmol) was dissolved in 150 mL CH_2Cl_2 and pyridine (24.5 mL, 300 mmol) was added slowly over 5 min at 0 ^oC. The mixture was stirred at 0 ^oC for 10 min then Tf_2O (15 mL, 90 mmol) was added dropwise over 10 min. The reaction mixture was warmed to room temperature slowly and stirred for 24 h. The reaction was quenched with water and the organic layer was washed with water (1 x 30 mL), 1N HCl (3 x 30 mL), brine (1 x 30 mL) and dried over Na_2SO_4. The solvents were removed under reduced pressure and the residue was recrystallized from CH_2Cl_2/hexanes to afford the pure product as a white crystal (13.2 g, 89 %). ^1H NMR (500 MHz, $CDCl_3$) δ 8.43 (d, 2H, *J* = 8.9 Hz), 7.47 (d, 2H, *J* = 2.3 Hz), 7.33 (dd, 2H, *J* = 8.9, 2.3 Hz); ^{13}C NMR (125 MHz, $CDCl_3$) δ 174.5, 156.6, 153.3, 129.6, 121.4, 118.7 (q, *J* = 319.1 Hz), 118.2, 111.4; ^{19}F NMR (282 MHz, $CDCl_3$) δ 104.2 (s).

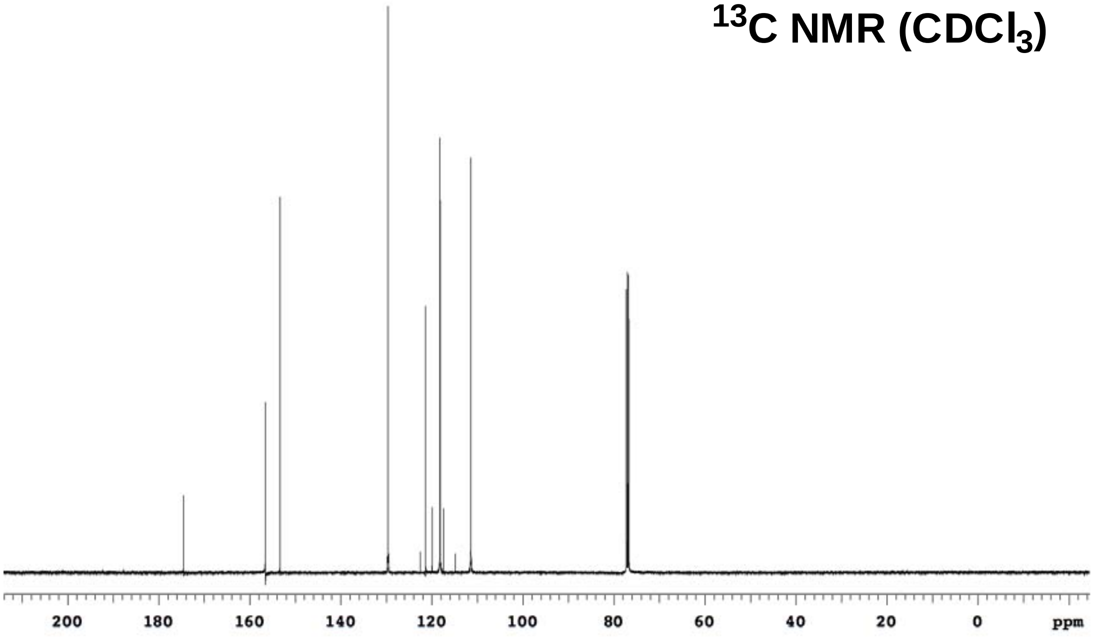




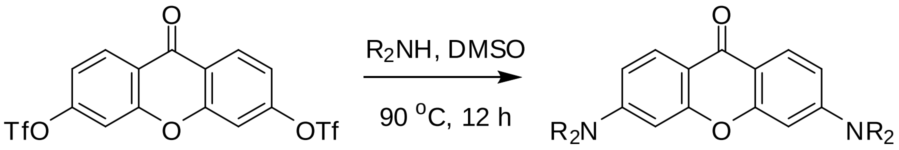
^{1}H NMR ($CDCl_3$)



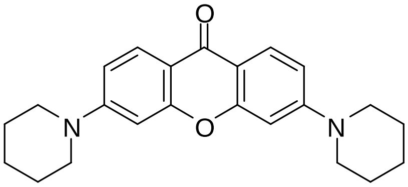
^{13}C NMR ($CDCl_3$)



General procedure for the synthesis of 3,6-diamino-xanthen-9-one



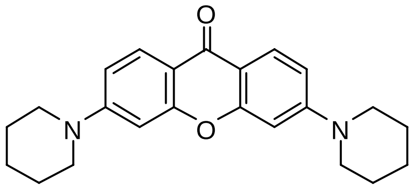
3,6-Di-triflate-xanthone **16** (1.0 eq.) was dissolved in DMSO (0.2 M) and the corresponding amine (10 eq.) was added. The reaction mixture was heated to 90 °C and stirred for 12 h. After cooling to room temperature, the reaction was quenched with water. The precipitate was collected, washed with saturated Na_2CO_3 $_{(aq.)}$ and water to give the crude product, which was recrystallized from EtOAc/hexanes to afford the pure product.



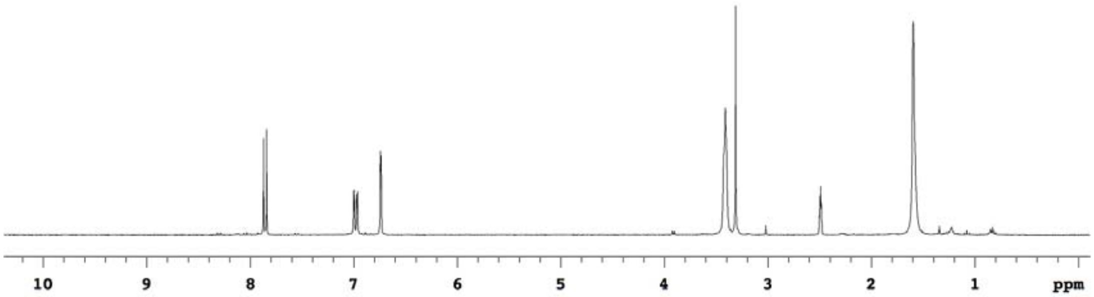
17

3,6-Di-piperidin-1-yl-xanthen-9-one (17)

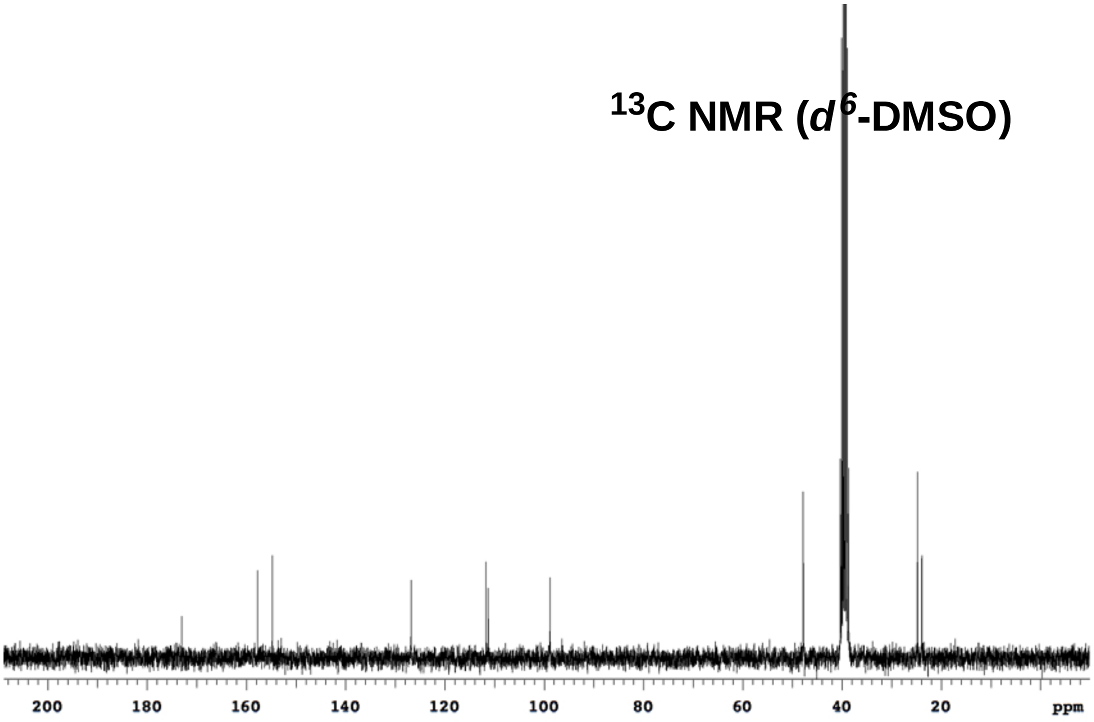
Yellow solid (3.0 g, 89 %). ^{1}H NMR (300 MHz, d^6-DMSO) δ 7.86 (d, 2H, J = 9.0 Hz), 6.98 (dd, 2H, J = 9.0, 2.3 Hz), 6.74 (d, 2H, J = 2.3 Hz), 3.41 (br, 8H), 1.60 (br, 12H); ^{13}C NMR (75 MHz, d^6-DMSO) δ 173.0, 157.7, 154.8, 126.8, 111.7, 111.3, 98.8, 47.8, 24.8, 23.9. MS (ESI) m/z calcd for $(M+H)^+$ $C_{23}H_{26}N_2O_2$ 363.21; found 363.21.

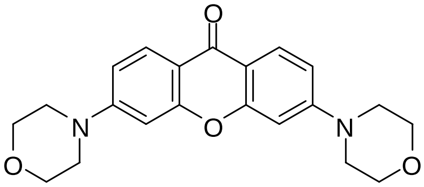


^{1}H NMR (d^6-DMSO)



^{13}C NMR (d^6-DMSO)

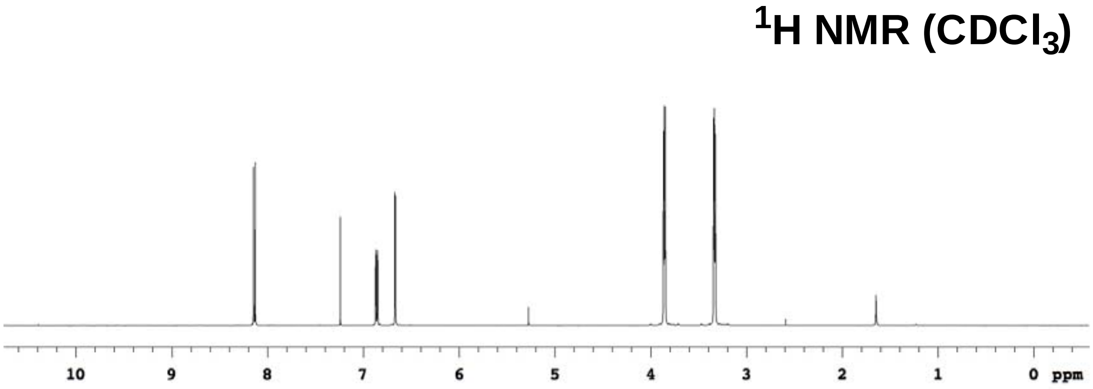


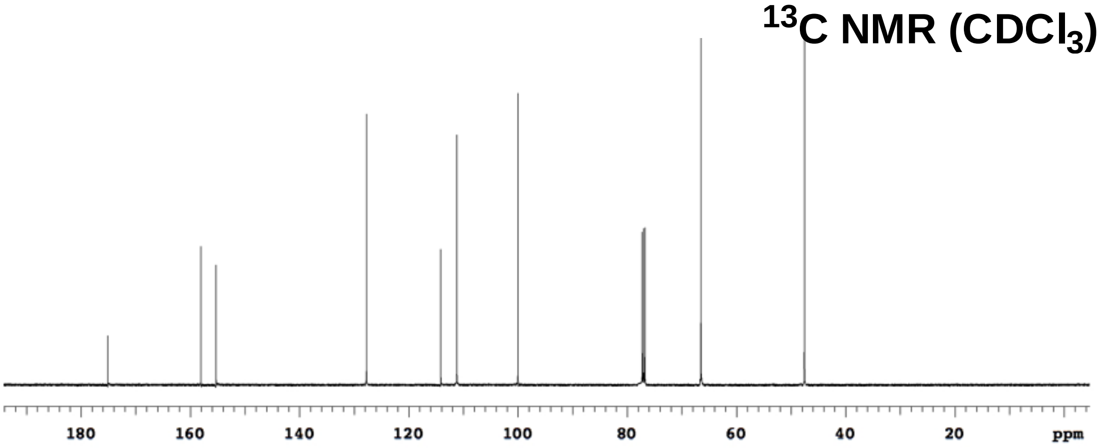


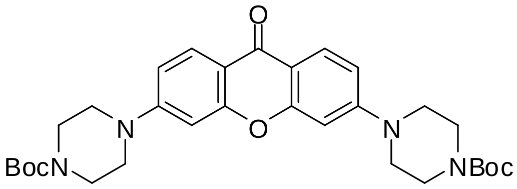
18

3,6-Di-morpholin-4-yl-xanthen-9-one (18)

White solid (308 mg, 93 %). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.14 (d, 2H, *J* = 9.0 Hz), 6.86 (dd, 2H, *J* = 9.0, 2.4 Hz), 6.67 (d, 2H, *J* = 2.4 Hz), 3.86 (t, 8H, *J* = 4.9 Hz), 3.33 (t, 8H, *J* = 4.9 Hz); ^{13}C NMR (125 MHz, CDCl$_3$) δ 175.1, 158.0, 155.3, 127.7, 114.1, 111.2, 100.0, 66.5, 47.5. MS (ESI) m/z calcd for (M+H)$^+$ $C_{21}H_{23}N_2O_4$ 367.17; found 367.17.



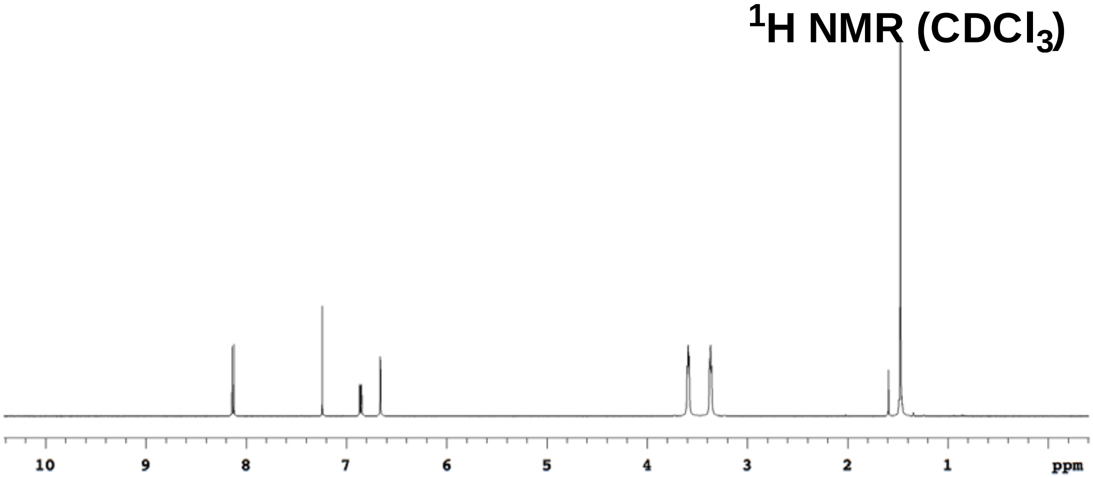


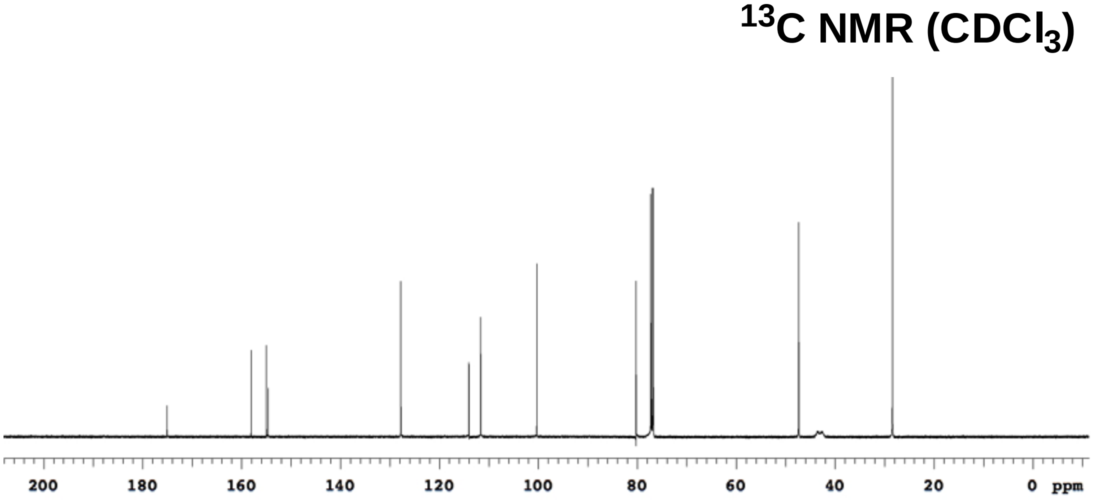


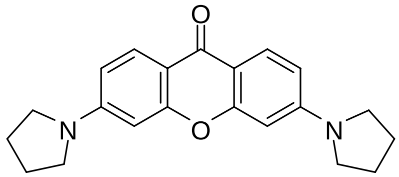
19

3,6-Bis-(4-Boc-piperazin-1-yl)-xanthen-9-one (19)

White solid (680 mg, 60 %). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.13 (d, 2H, *J* = 9.0 Hz), 6.86 (dd, 2H, *J* = 9.0, 2.3 Hz), 6.66 (d, 2H, *J* = 2.3 Hz), 3.59 (t, 8H, *J* = 5.2 Hz), 3.37 (t, 8H, *J* = 5.2 Hz), 1.47 (s, 18H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 175.0, 158.1, 155.0, 154.6, 127.8, 114.0, 111.7, 100.3, 80.2, 47.3(2C), 28.4. MS (ESI) m/z calcd for (M+H)$^+$ $C_{31}H_{41}N_4O_6$ 565.30; found 565.29.


^{1}H NMR (CDCl$_3$)


^{13}C NMR (CDCl$_3$)

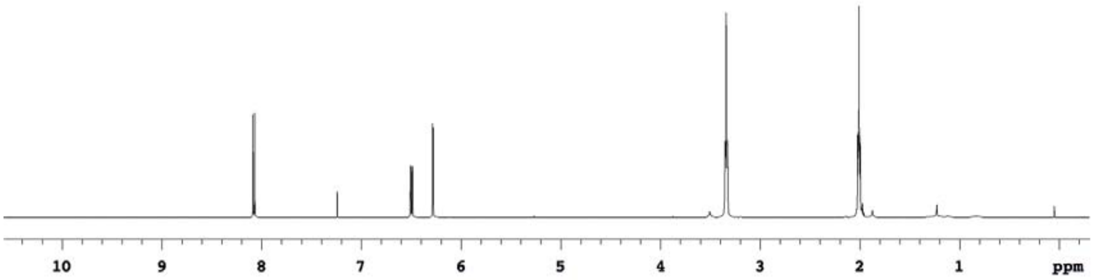


20

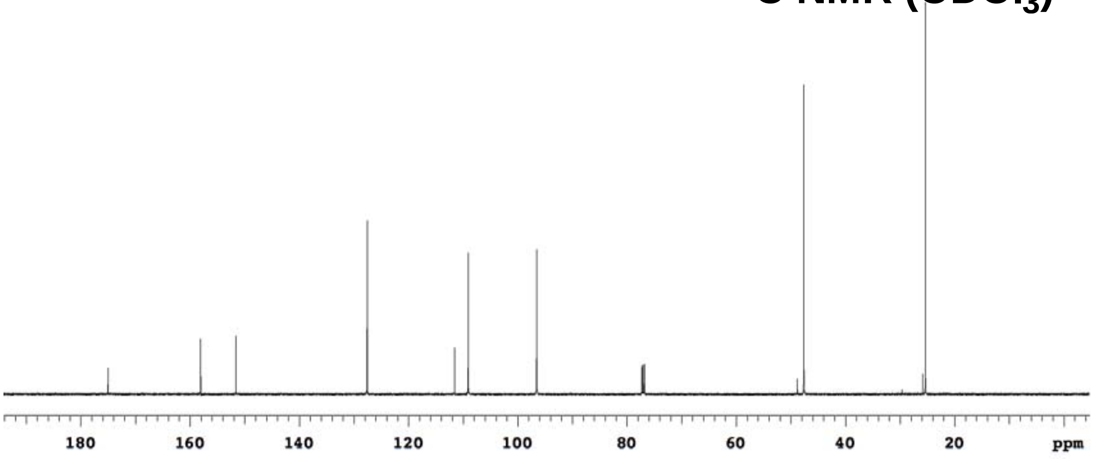
3,6-Di-pyrrolidin-1-yl-xanthen-9-one (20)

Yellow solid (90 mg, 18 %). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.08 (d, 2H, *J* = 8.9 Hz), 6.50 (dd, 2H, *J* = 8.9, 2.2 Hz), 6.28 (d, 2H, *J* = 2.2 Hz), 3.34 (t, 8H, *J* = 6.6 Hz), 2.01 (t, 8H, *J* = 6.6 Hz); ^{13}C NMR (125 MHz, CDCl$_3$) δ 175.0, 158.1, 151.6, 127.6, 111.5, 109.1, 96.5, 47.6, 25.4. MS (ESI) m/z calcd for $(M+H)^+$ $C_{21}H_{23}N_2O_2$ 335.18; found 335.18.

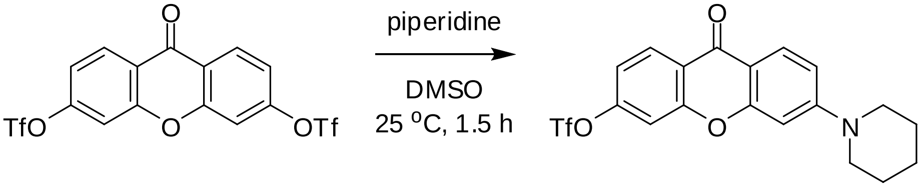
^{1}H NMR (CDCl$_3$)



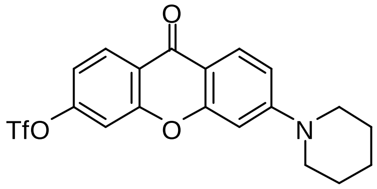
^{13}C NMR (CDCl$_3$)



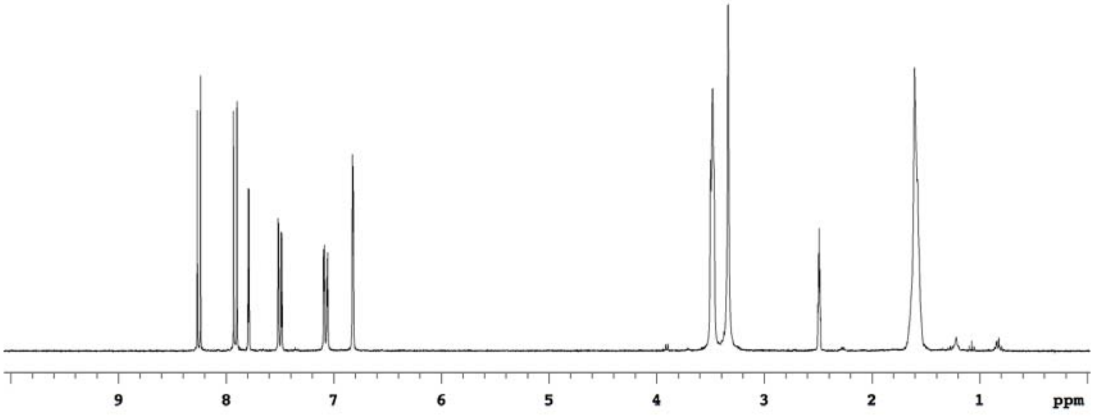
Synthesis of 3-trifyl-6-piperidine-xanthone



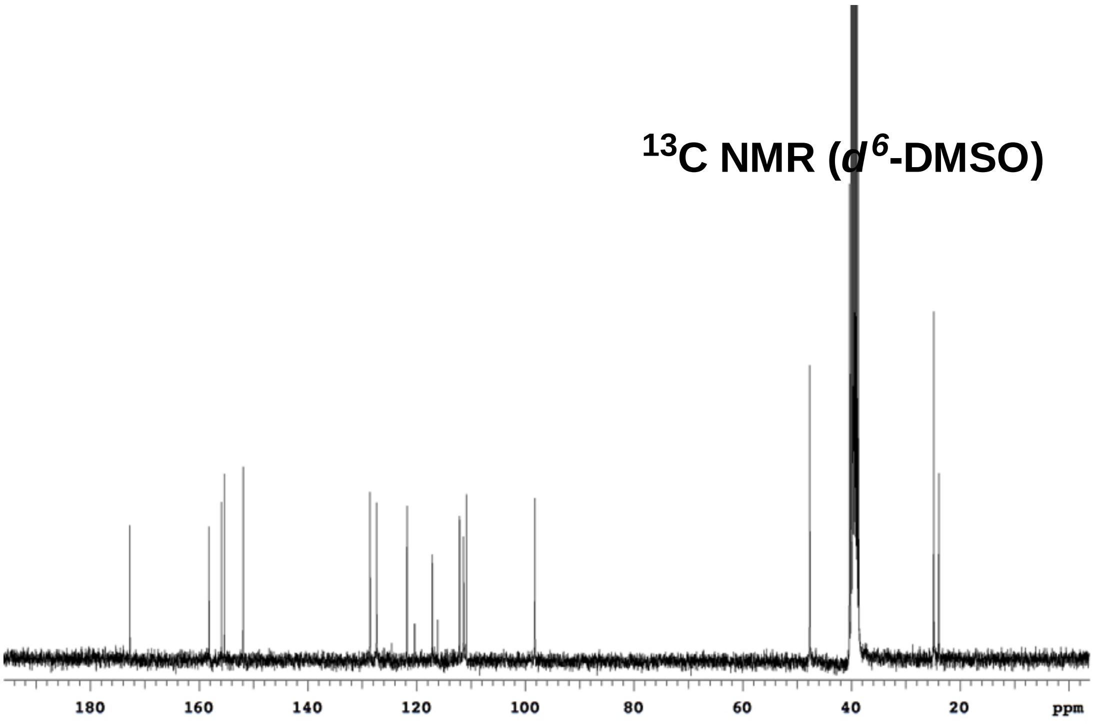
3,6-Di-OTf-xanthone (492 mg, 1.0 mmol) was dissolved in 10 mL DMSO and piperidine (0.49 mL, 5.0 mmol) was added. After stirring at 25 °C for 1.5 h, the solution was diluted with 30 mL CH_2Cl_2 and washed with water (3 x 30 mL). The solvents were removed under reduced pressure and the residue was purified by flash chromatography (10 % to 20 % EtOAc/Hexanes) to afford the pure product (372 mg, 87 %) as a light yellow solid. $R_f = 0.66$ (40 % EtOAc/hexanes). ^{1}H NMR (300 MHz, d^6-DMSO) δ 8.25 (d, 1H, $J = 8.8$ Hz), 7.91 (d, 1H, $J = 9.1$ Hz), 7.79 (d, 1H, $J = 2.4$ Hz), 7.50 (dd, 1H, $J = 8.8, 2.4$ Hz), 7.07 (dd, 1H, $J = 9.1, 2.4$Hz), 6.82 (d, 1H, $J = 2.4$ Hz), 3.50-3.47 (m, 4H), 1.60 (br, 6H); ^{13}C NMR (75 MHz, d^6-DMSO) δ 172.7, 158.2, 155.9, 155.3, 151.9, 128.5, 127.3, 121.8, 118.4 (q, $J = 320.5$ Hz), 117.1, 112.1, 111.3, 110.8, 98.2, 47.7, 24.9, 23.9. MS (ESI) m/z calcd for $(M+H)^+$ $C_{19}H_{17}F_3NO_5S$ 428.08; found 428.08.



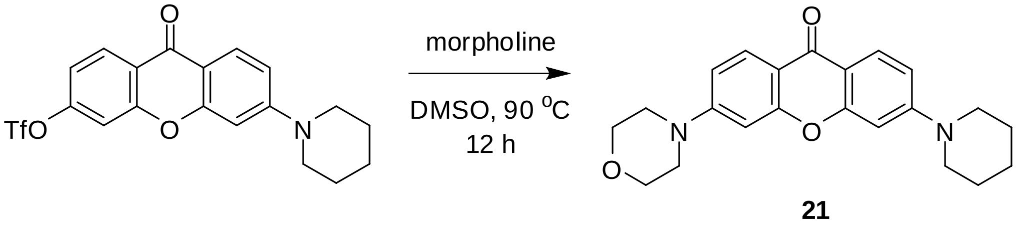
^{1}H NMR (d^6-DMSO)



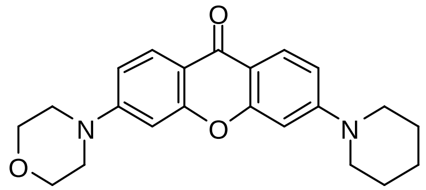
^{13}C NMR (d^6-DMSO)



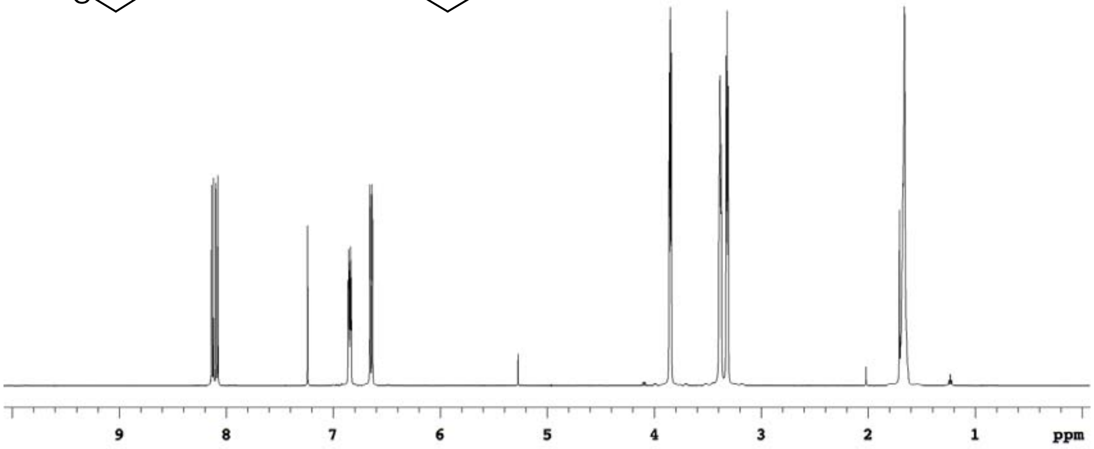
Synthesis of unsymmetrical 3,6-diamino-xanthen-9-one (21)



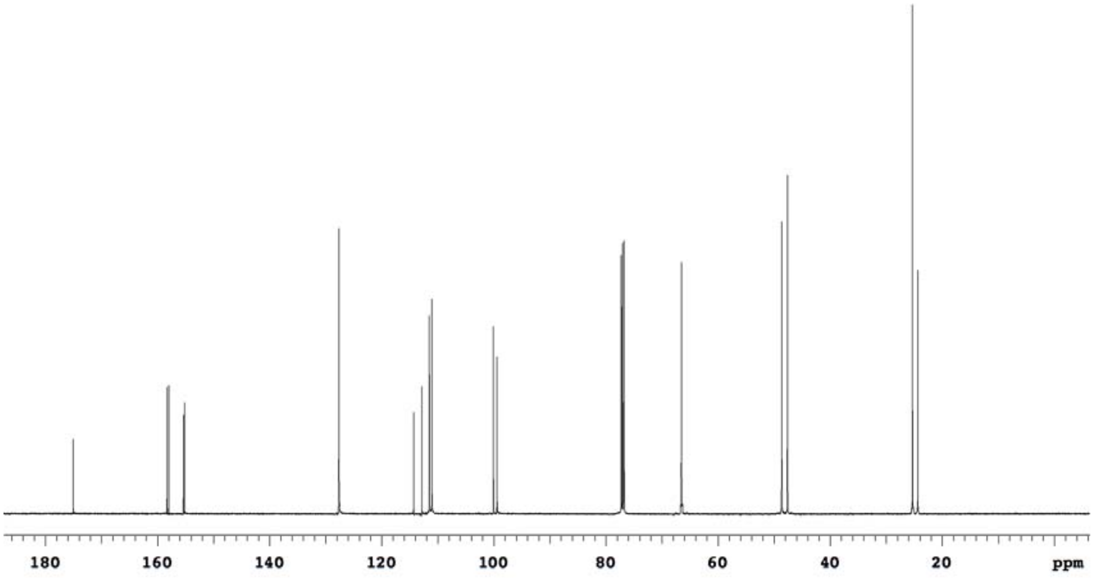
Morpholine (0.66 mL, 7.5 mmol) was added to the solution of 3-OTf-6-piperidine-xanthone (321 mg, 0.75 mmol) in 10 mL DMSO. The reaction mixture was heated to 90 ^{o}C and stirred for 12 h. After cooling to room temperature, the solution was diluted with 30 mL CH_2Cl_2 and washed with saturated Na_2CO_3 $_{(aq.)}$ (1 x 30 mL), water (1 x 30 mL) and dried over Na_2SO_4. The solvents were removed under reduced pressure and the residue was purified by flash chromatography (40 % EtOAc/Hexanes) to afford the pure product **21** (261 mg, 96 %) as a light yellow solid. R_f = 0.20 (40 % EtOAc/hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.13 (d, 1H, J = 9.1 Hz), 8.09 (d, 1H, J = 9.1 Hz), 6.86-6.83 (m, 2H), 6.66 (d, 1H, J = 2.2 Hz), 6.64 (d, 1H, J = 2.2 Hz), 3.85 (t, 4H, J = 4.9 Hz), 3.39-3.37 (m, 4H), 3.32 (t, 4H, J = 4.9 Hz), 1.66 (br, 6H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 175.0, 158.3, 158.0, 155.4, 155.1, 127.7, 127.6, 114.3, 112.8, 111.5, 111.1, 100.1, 99.4, 66.5, 48.6, 47.6, 25.3, 24.3. MS (ESI) m/z calcd for $(M+H)^+$ $C_{22}H_{25}N_2O_3$ 365.19; found 365.17.



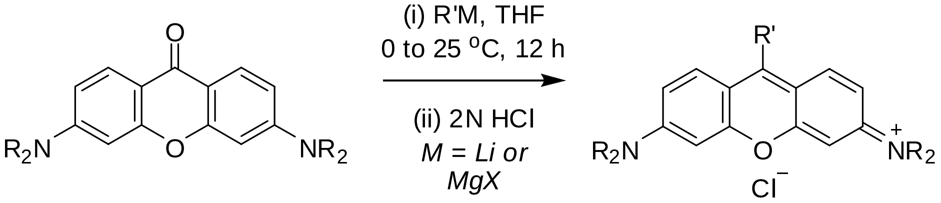
^{1}H NMR (CDCl$_3$)



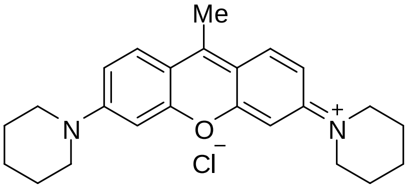
^{13}C NMR (CDCl$_3$)



Synthesis of rosamines



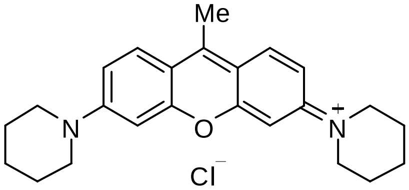
A Grignard reagent or Lithium reagent (1.0 mmol) was added dropwise over 1 min to the solution of 3,6-diamino-xanthen-9-one (0.2 mmol) in 5 mL THF at 0 ^{o}C. After stirring for 12 h at room temperature, the reaction mixture was quenched by 2 mL 2N HCl $_{(aq.)}$ and stirred for 10 min then diluted with 20 mL CH_2Cl_2. The organic layer was washed with water, brine, dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatography (5% to 10% MeOH/CH_2Cl_2) to give the pure product.



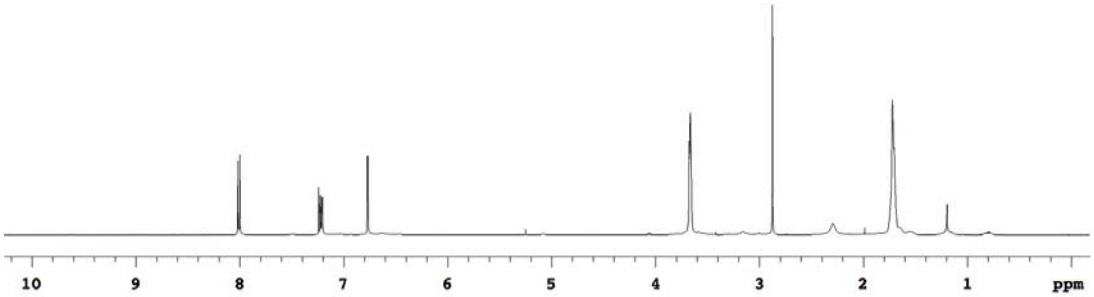
11a

Rosamine 11a

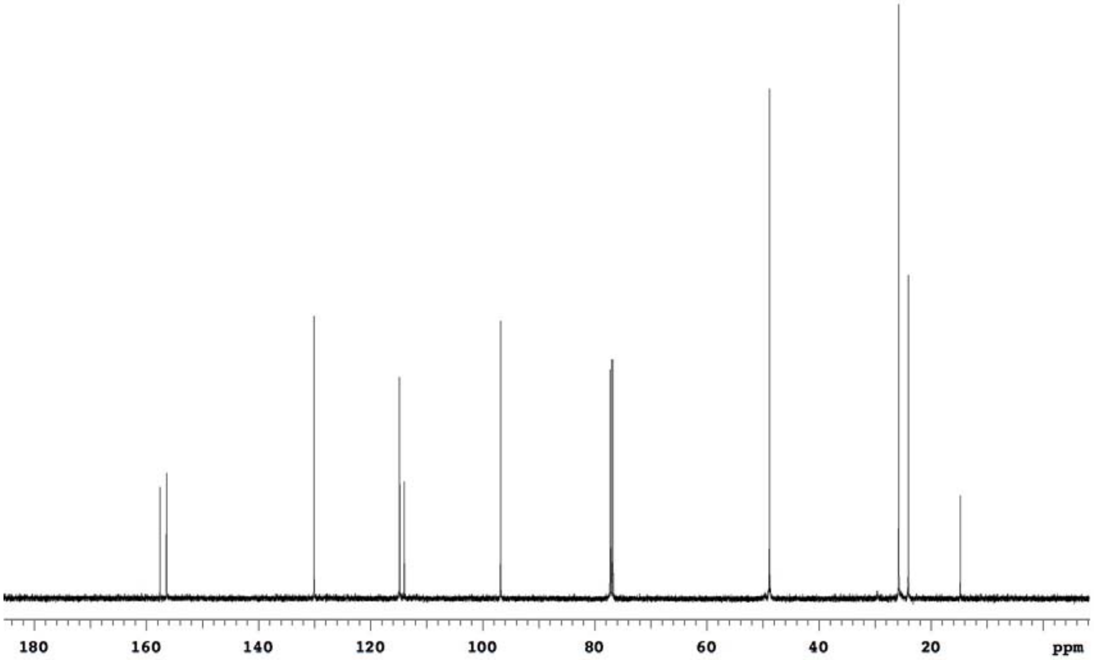
Purple solid (71 mg, 89 %). R_f = 0.28 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.01 (d, 2H, *J* = 9.6 Hz), 7.22 (dd, 2H, *J* = 9.6, 2.5 Hz), 6.77 (d, 2H, *J* = 2.5 Hz), 3.66 (br, 8H), 2.87 (s, 3H), 1.72 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 157.6, 156.4, 156.3, 130.1, 114.8, 114.0, 96.8, 48.8, 25.8, 24.1, 14.8; IR (thin film) 1643, 1597, 1486, 1401, 1235, 1200 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{24}H_{29}N_2O$ 361.2280; found 361.2287.

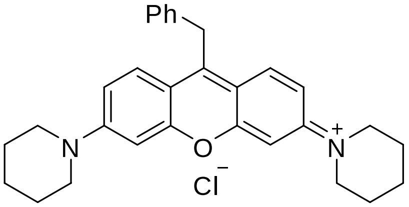


^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

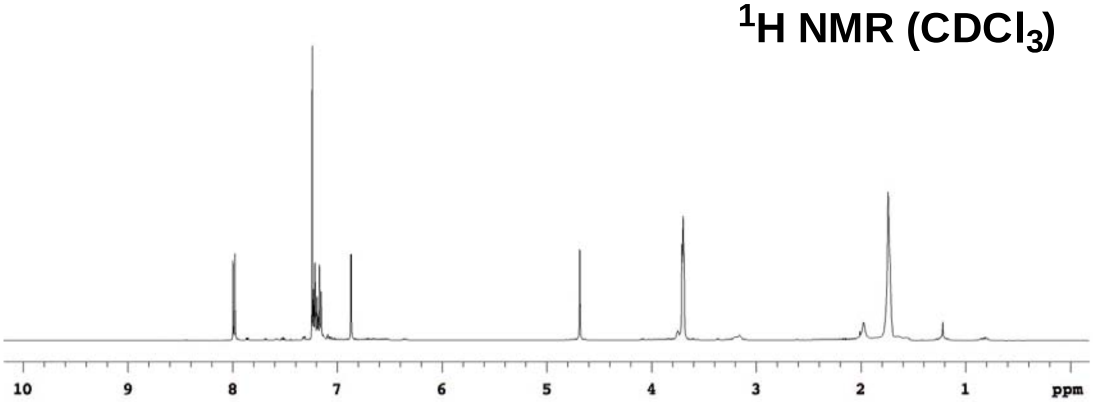


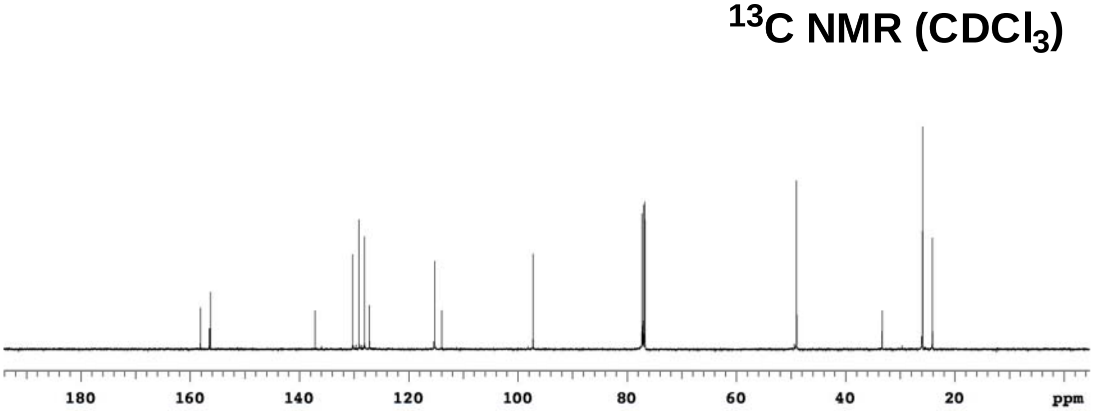


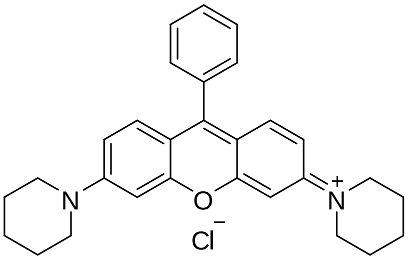
11b

Rosamine 11b

Purple solid (66 mg, 70 %). R_f = 0.34 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 (d, 2H, *J* = 9.7 Hz), 7.23-7.15 (m, 7H), 6.87 (d, 2H, *J* = 2.5 Hz), 4.68 (s, 2H), 3.70 (br, 8H), 1.74 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.1, 156.5, 156.3, 137.1, 130.3, 129.1, 128.1, 127.2, 115.2, 113.9, 97.2, 49.0, 33.3, 25.9, 24.1; IR (thin film) 1643, 1594, 1480, 1424, 1397, 1237, 1169 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{30}H_{33}N_2O$ 437.2593; found 437.2595.





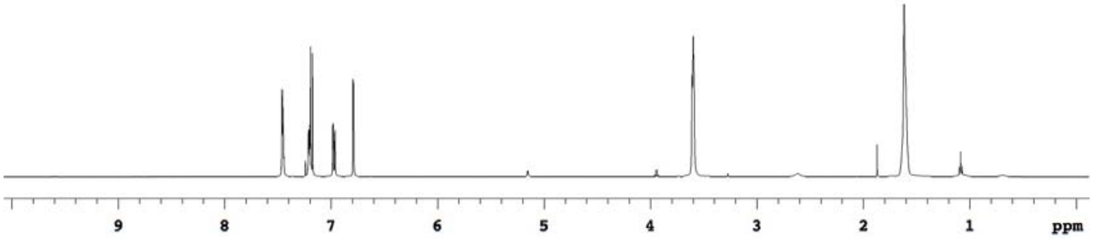


11c

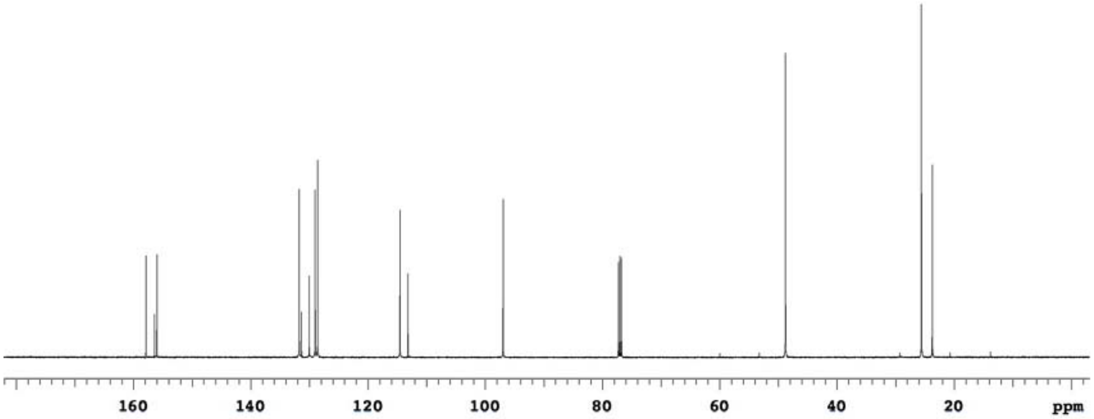
Rosamine 11c

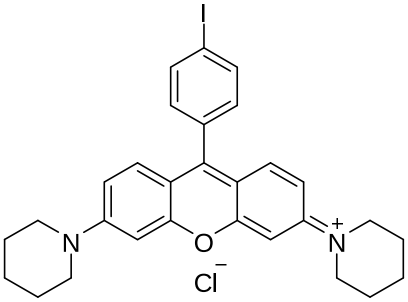
Green solid (92 mg, 100 %). R_f = 0.35 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.46-7.45 (m, 3H), 7.21-7.18 (m, 4H), 6.97 (dd, 2H, *J* = 9.6, 2.5 Hz), 6.79 (d, 2H, *J* = 2.5 Hz), 3.60 (br, 8H), 1.62 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 157.9, 156.4, 156.0, 131.7, 131.3, 130.0, 129.0, 128.6, 114.5, 113.2, 96.9, 48.8, 25.6, 23.7; IR (thin film) 1646, 1592, 1483, 1414, 1391, 1235, 1190 cm^{-1}; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{29}H_{31}N_2O$ 423.2436; found 423.2437.

^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

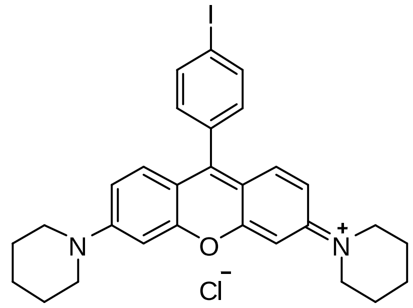




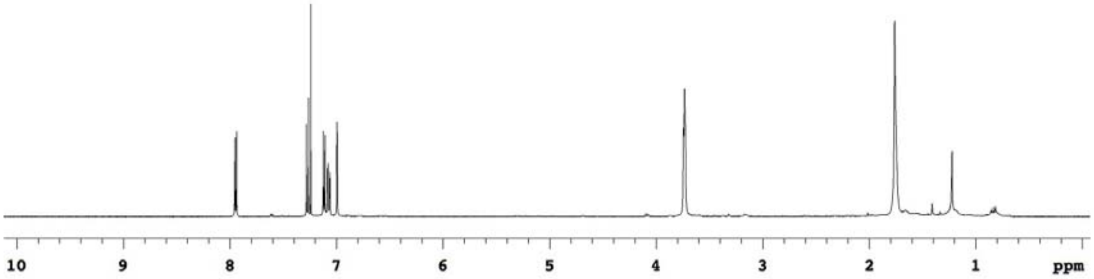
11d

Rosamine 11d

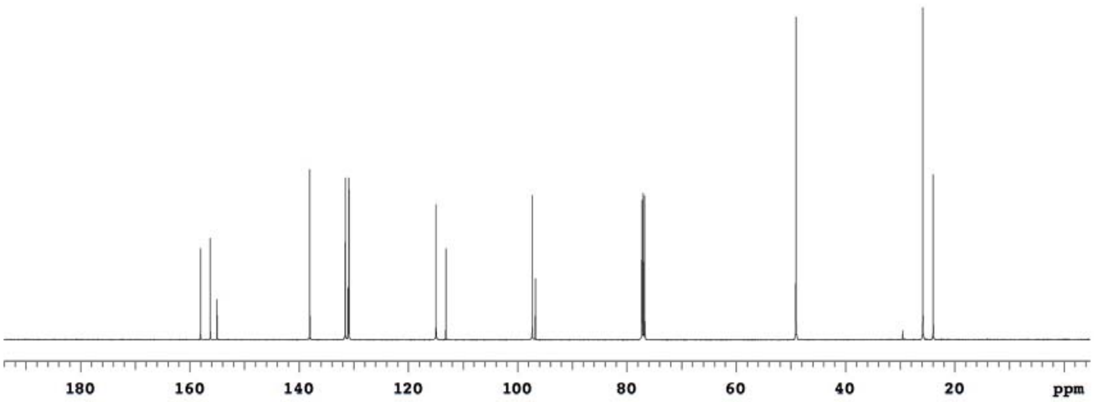
Green solid (105 mg, 90 %) [Made from 4-Iodophenyl magnesium chloride, which was prepared from 1,4-di-iodobenzene with ^{i}PrMgCl]. $R_f = 0.36$ (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.95 (d, 2H, *J* = 8.4 Hz), 7.27 (d, 2H, *J* = 9.7 Hz), 7.12 (d, 2H, *J* = 8.4 Hz), 7.07 (dd, 2H, *J* = 9.7, 2.5 Hz), 6.99 (d, 2H, *J* = 2.5 Hz), 3.73 (br, 8H), 1.76 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.0, 156.2, 155.0, 138.0, 131.5, 131.0, 130.9, 114.9, 113.1, 97.3, 96.7, 49.1, 25.8, 24.0; IR (thin film) 1644, 1592, 1483, 1420, 1391, 1235, 1193 cm^{-1}; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{29}H_{30}IN_2O$ 549.1403; found 549.1399. Anal. Calcd for $C_{29}H_{30}ClIN_2O$: C, 59.55; H, 5.17; N, 4.79. Found: C, 58.00; H, 5.17; N, 4.71. The elemental analysis data are consistent with the presence of one molecule of water per molecules of product.

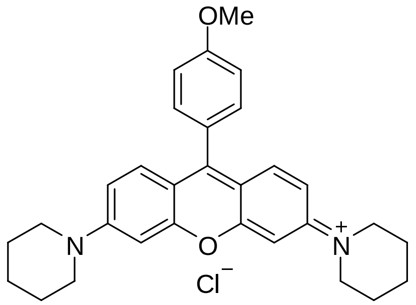


^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

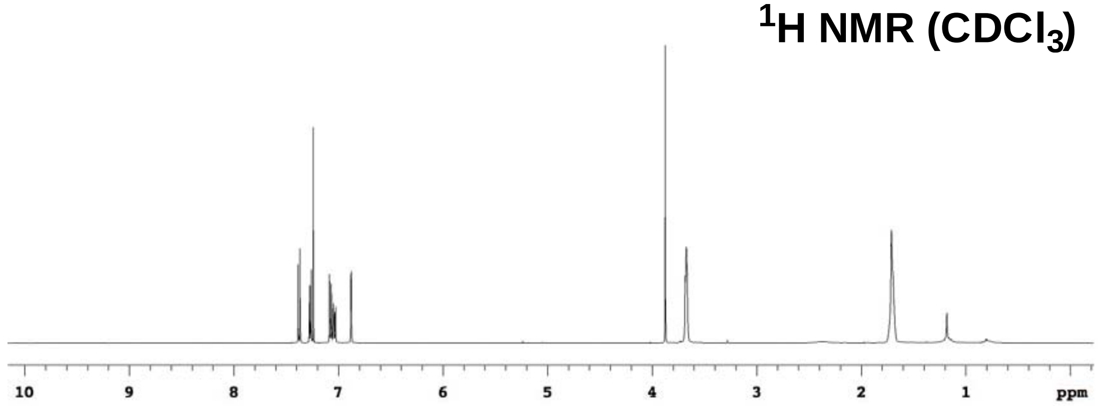


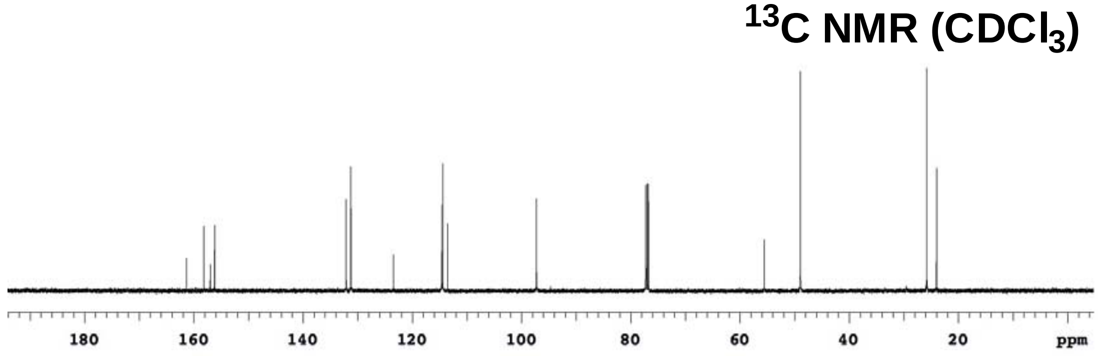


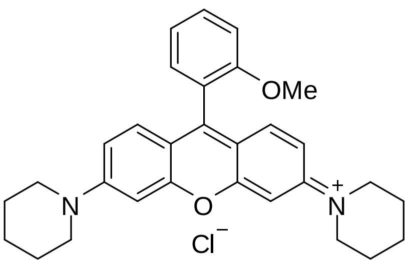
11e

Rosamine 11e

Green solid (98 mg, 100 %). $R_f = 0.33$ (10 % MeOH/CH_2Cl_2). ^1H NMR (500 MHz, $CDCl_3$) δ 7.38 (d, 2H, $J = 9.6$ Hz), 7.27 (d, 2H, $J = 8.8$ Hz), 7.08 (d, 2H, $J = 8.8$ Hz), 7.04 (dd, 2H, $J = 9.6, 2.5$ Hz), 6.88 (d, 2H, $J = 2.5$ Hz), 3.87 (s, 3H), 3.67 (br, 8H), 1.71 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 161.3, 158.2, 157.0, 156.2, 132.1, 131.2, 123.5, 114.6, 114.4, 113.5, 97.3, 55.5, 48.9, 25.8, 24.0; IR (thin film) 1644, 1592, 1480, 1391, 1235, 1191 cm^{-1}; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{30}H_{33}N_2O_2$ 453.2542; found 453.2546.





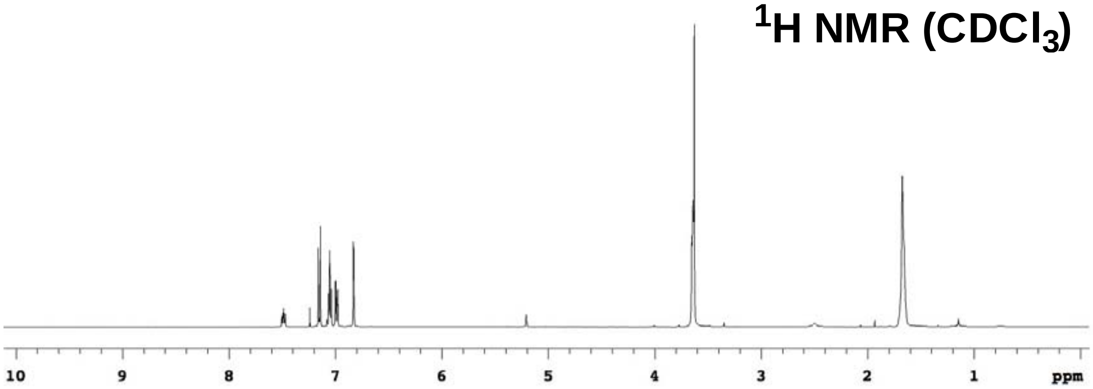


11f

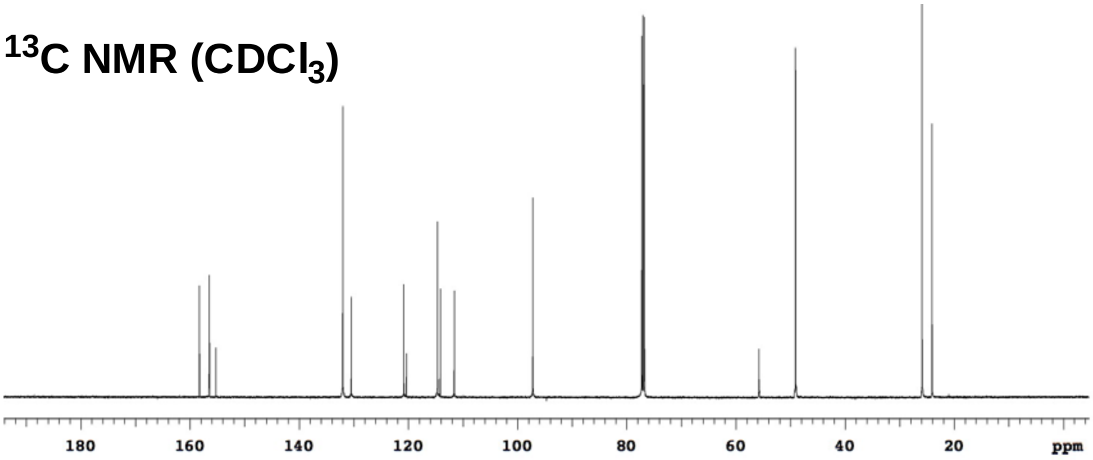
Rosamine 11f

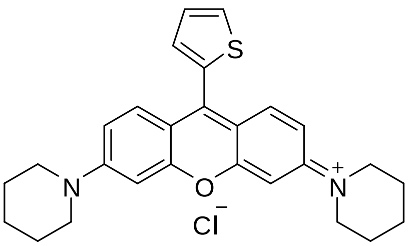
Green solid (97 mg, 100 %). R_f = 0.36 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.50-7.47 (m, 1H), 7.15 (d, 2H, *J* = 9.5 Hz), 7.08-7.04 (m, 3H), 6.99 (dd, 2H, *J* = 9.5, 2.5 Hz), 6.83 (d, 2H, *J* =2.5 Hz), 3.65-3.63 (m, 11H), 1.67 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.3, 156.5, 156.4, 155.2, 132.0 (2C), 130.5, 120.8, 120.4, 114.6, 114.1, 111.6, 97.2, 55.8, 49.1, 25.9, 24.1; IR (thin film) 1646, 1590, 1480, 1415, 1390, 1233, 1187 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{30}H_{33}N_2O_2$ 453.2542; found 453.2540.

^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)



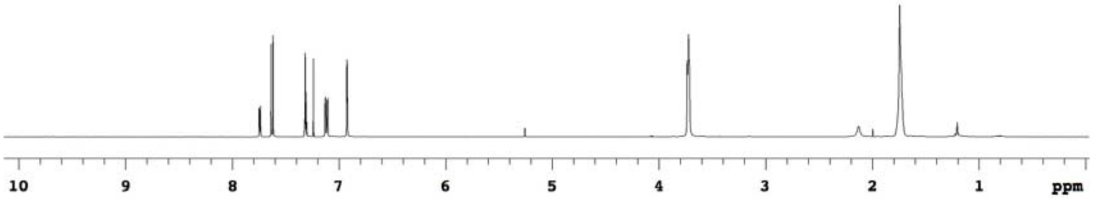


11g

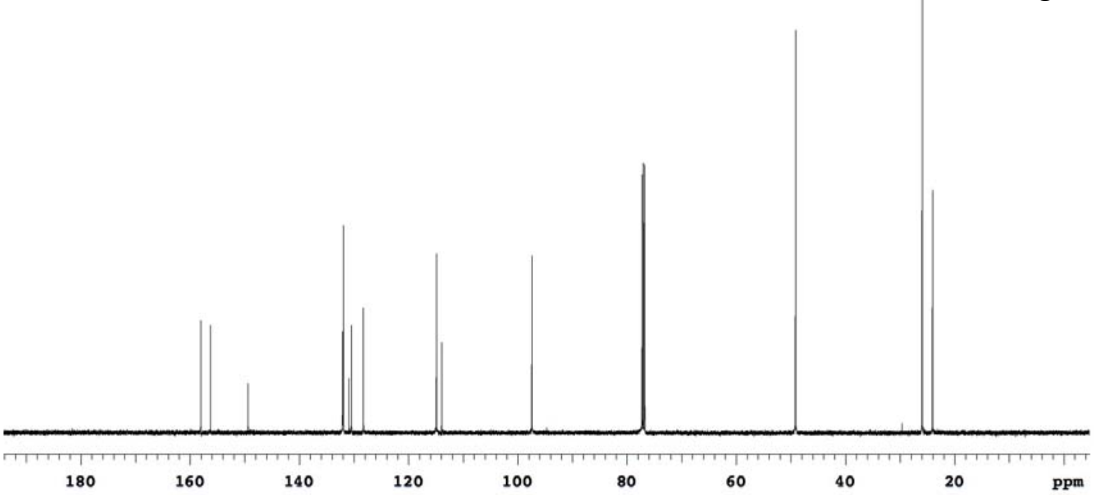
Rosamine 11g

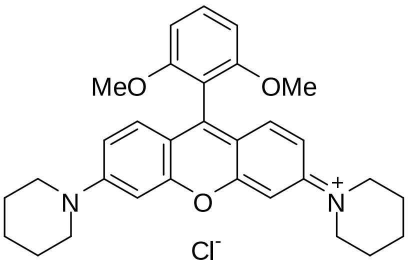
Green solid (69 mg, 74 %). R_f = 0.25 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.74 (dd, 1H, *J* = 4.5, 1.7 Hz), 7.63 (d, 2H, *J* = 9.5 Hz), 7.32-7.30 (m, 2H), 7.12 (dd, 2H, *J* = 9.5, 2.5 Hz), 6.92 (d, 2H, *J* = 2.5 Hz), 3.72 (br, 8H), 1.74 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.0, 156.3, 149.4, 132.1, 131.9, 130.9, 130.4, 128.3, 114.9, 113.9, 97.4, 49.1, 25.9, 24.1; IR (thin film) 1642, 1592, 1482, 1415, 1391, 1237, 1192 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{27}H_{29}N_2OS$ 429.2001; found 429.1995.

^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

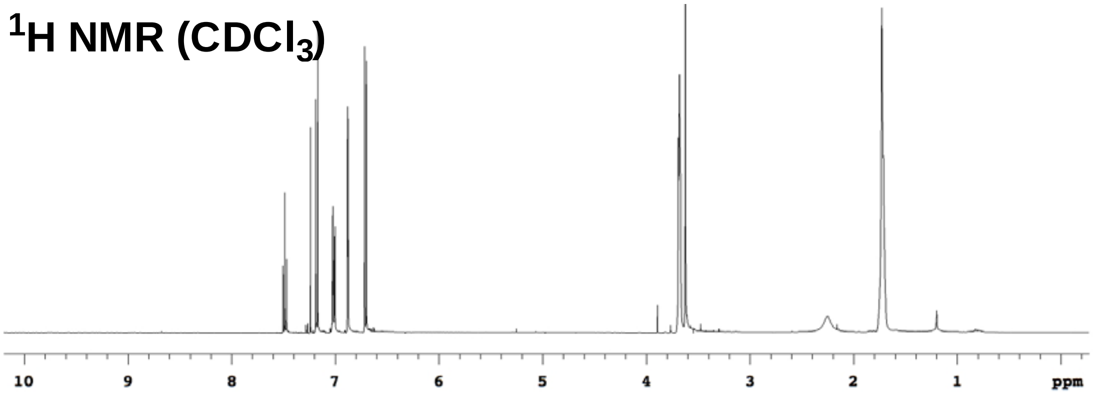


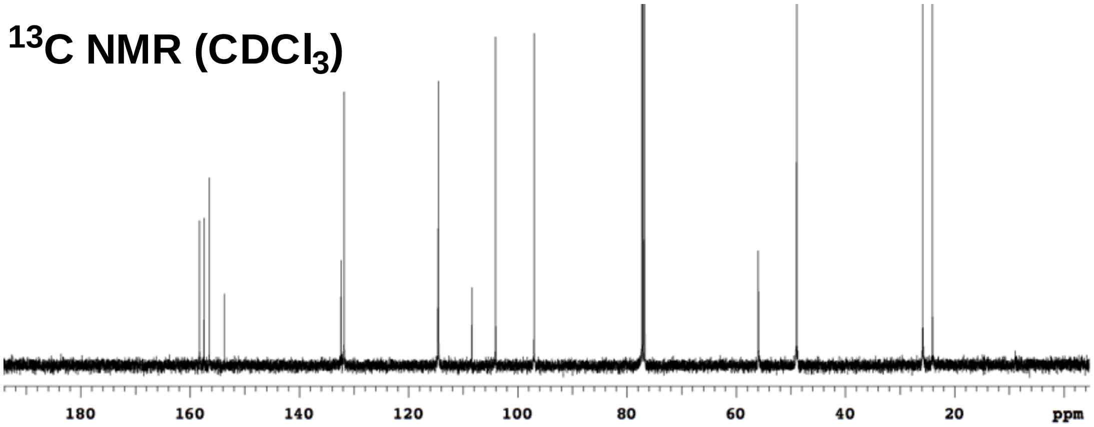


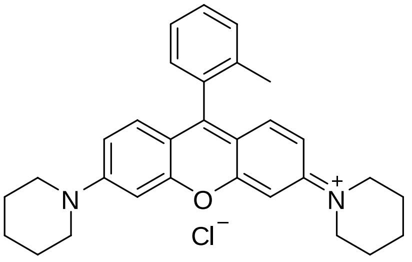
11h

Rosamine 11h

Green solid (100 mg, 96 %). R_f = 0.32 (10 % MeOH/CH_2Cl_2). [Made from 2,6-dimethoxyphenyl lithium, which was prepared from 1,3-dimethoxybenzene with ^{n}BuLi]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.49 (t, 1H, *J* = 8.5 Hz), 7.18 (d, 2H, *J* = 9.5 Hz), 7.01 (dd, 2H, *J* = 9.5, 2.5 Hz), 6.88 (d, 2H, *J* = 2.5 Hz), 6.71 (d, 2H, *J* = 8.5 Hz), 3.69-3.67 (m, 8H), 3.62 (s, 6H), 1.73 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.3, 157.4, 156.5, 153.7, 132.3, 131.7, 114.6, 114.5, 108.4, 104.0, 97.0, 55.9, 48.9, 25.8, 24.1; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{31}H_{35}N_2O_3$ 483.2648; found 483.2656.


^{1}H NMR ($CDCl_3$)

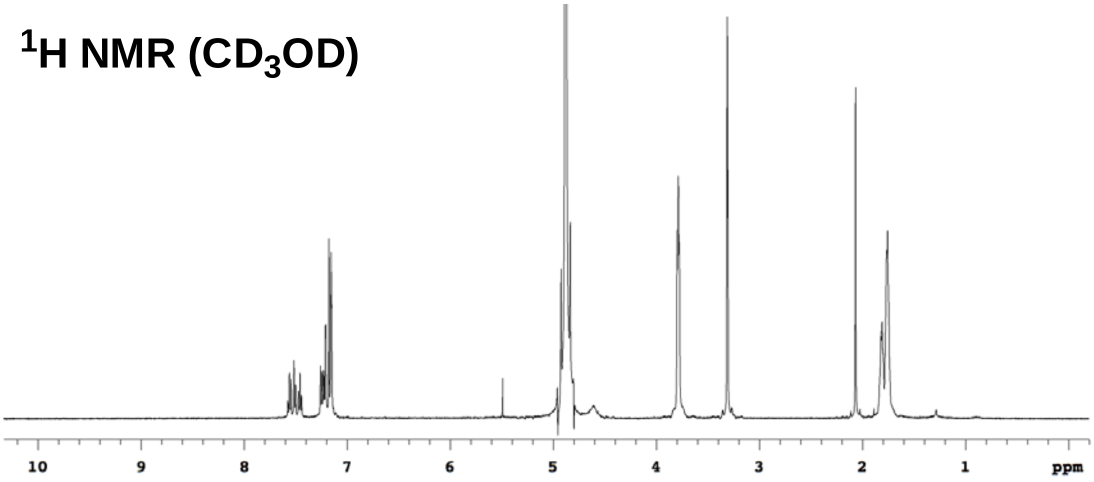

^{13}C NMR ($CDCl_3$)

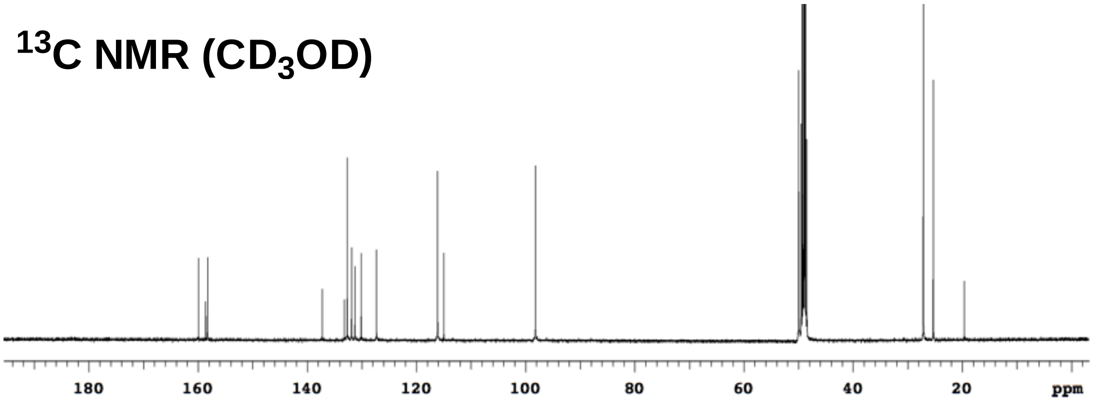


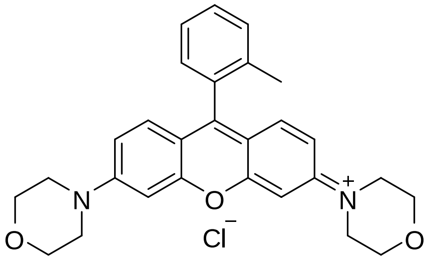
11i

Rosamine 11i

Green solid (95 mg, 99 %). *R_f* = 0.32 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, CD_3OD) δ 7.58-7.44 (m, 3H), 7.26-7.16 (m, 7H), 3.79 (br, 8H), 2.07 (s, 3H), 1.82-1.76 (m, 12H); ^{13}C NMR (125 MHz, CD_3OD) δ 159.9, 158.6, 158.2, 137.2, 133.2, 132.7, 131.9, 131.3, 130.1, 127.3, 116.1, 115.0, 98.2, 50.0, 27.1, 25.3, 19.6; IR (thin film) 1646, 1588, 1482, 1414, 1389, 1233, 1187 cm^{-1}; HRMS (ESI) m/z calcd for (M-Cl)$^+$ $C_{30}H_{33}N_2O$ 437.2593; found 437.2598.



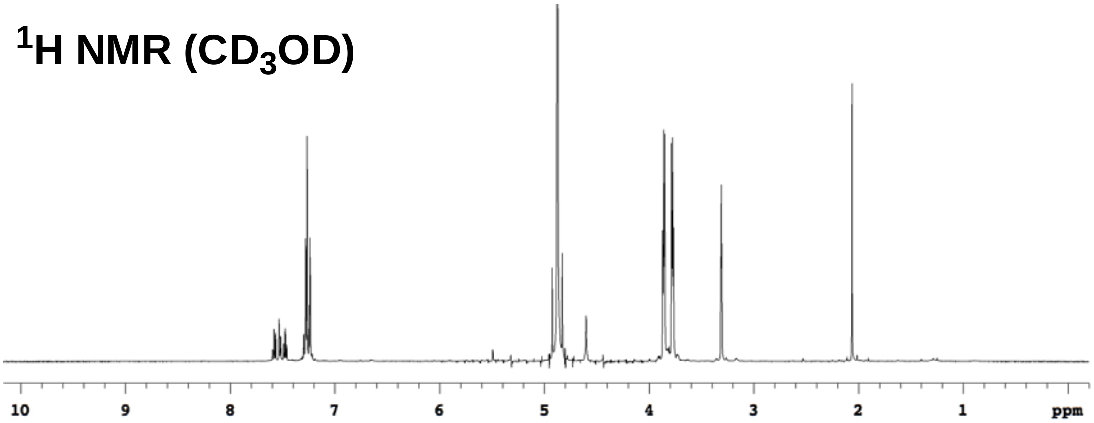


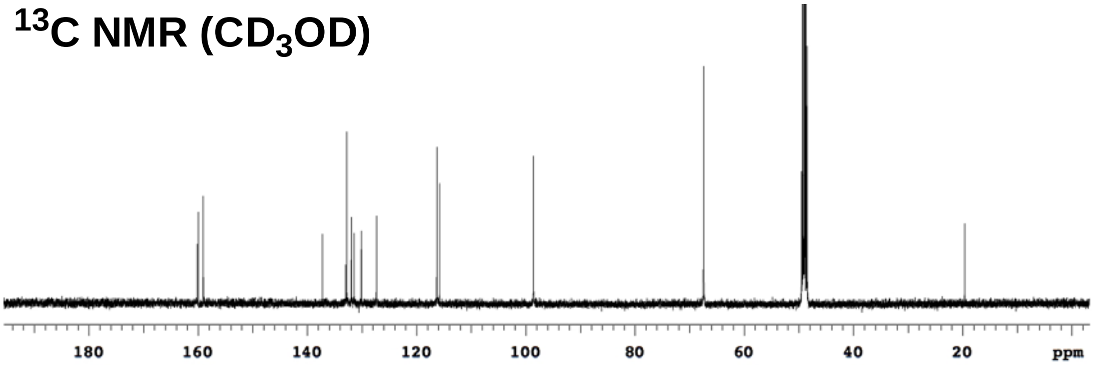


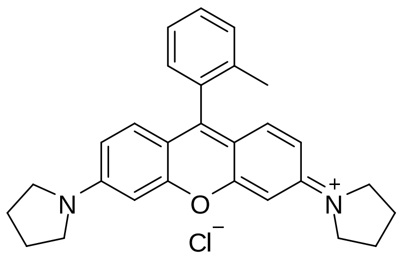
12i

Rosamine 12i

Green solid (94 mg, 98 %). R_f = 0.26 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, CD_3OD) δ 7.59-7.46 (m, 3H), 7.30-7.23 (m, 7H), 3.87-3.85 (m, 8H), 3.79-3.77 (m, 8H), 2.06 (s, 3H); ^{13}C NMR (125 MHz, CD_3OD) δ 160.2, 159.9, 159.1, 137.2, 133.0, 132.8, 131.9, 131.4, 130.1, 127.3, 116.3, 115.8, 98.6, 67.4, 48.5, 19.6; IR (thin film) 1646, 1590, 1481, 1415, 1383, 1235, 1190 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{28}H_{29}N_2O_3$ 441.2178; found 441.2184.



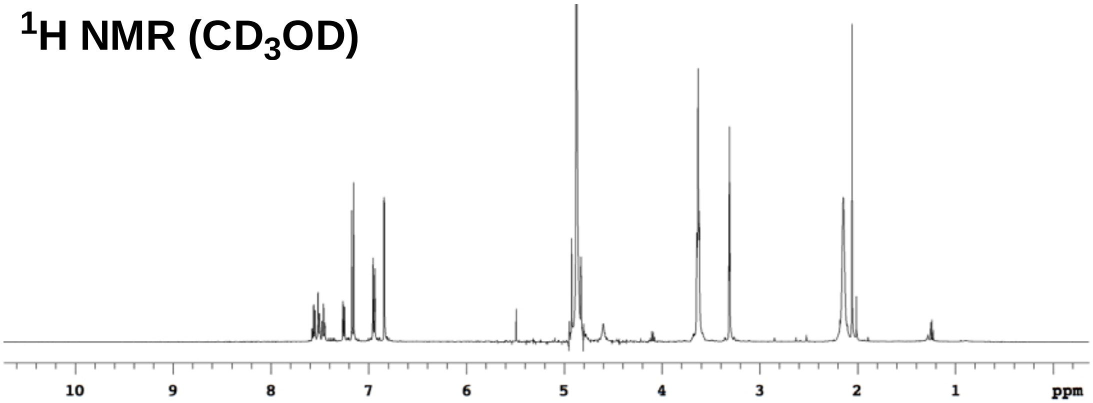


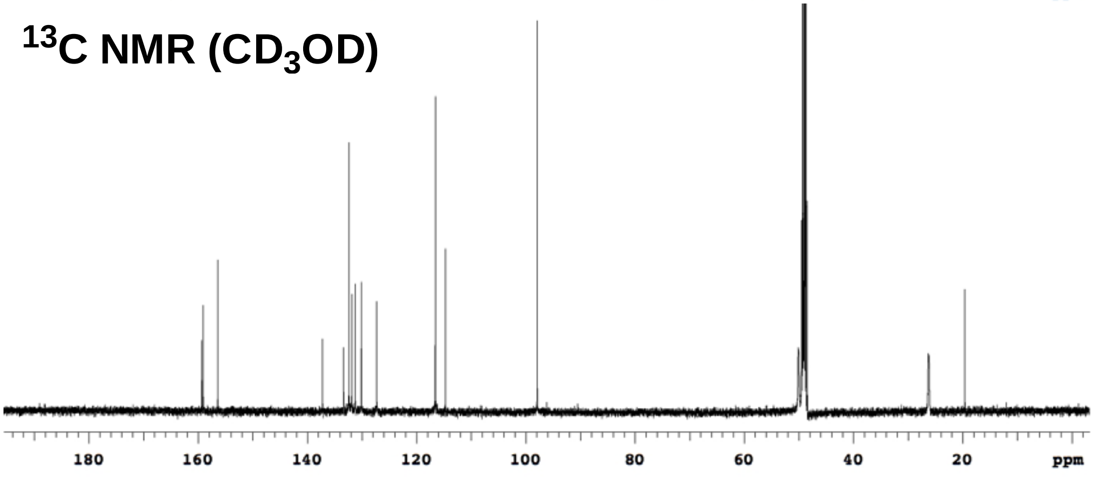


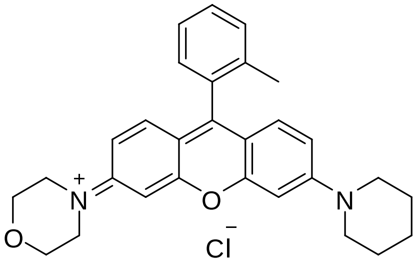
13i

Rosamine 13i

Green solid (47 mg, 53 %). R_f = 0.28 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, CD_3OD) δ 7.58-7.45 (m, 3H), 7.26 (d, 1H, *J* = 7.5 Hz), 7.16 (d, 2H, *J* = 9.5 Hz), 6.95 (dd, 2H, *J* = 9.5, 2.2 Hz), 6.84 (d, 2H, *J* = 2.2 Hz), 3.63 (br, 8H), 2.15 (br, 8H), 2.06 (s, 3H); ^{13}C NMR (125 MHz, CD_3OD) δ 159.3, 159.1, 156.4, 137.2, 133.4, 132.4, 131.9, 131.2, 130.1, 127.3, 116.6, 114.7, 97.9, 50.1, 26.3, 19.6; IR (thin film) 1648, 1596, 1413, 1378, 1344, 1189 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{28}H_{29}N_2O$ 409.2280; found 409.2277.



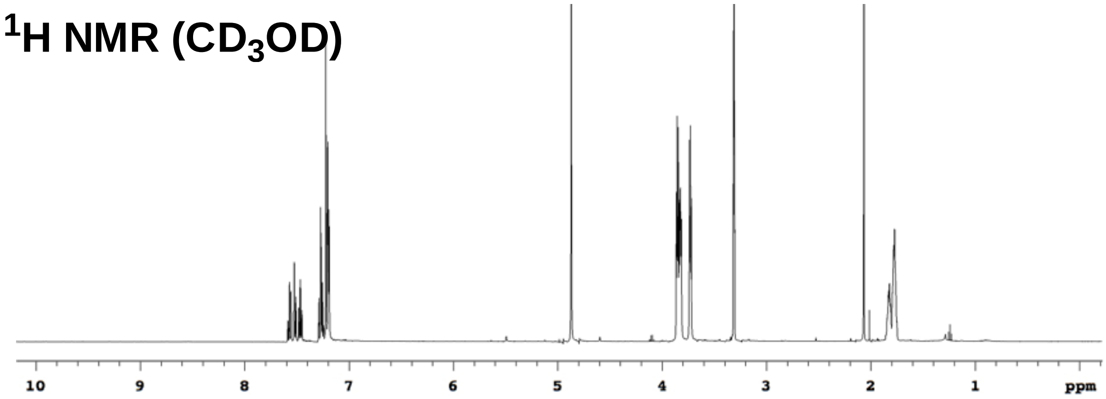


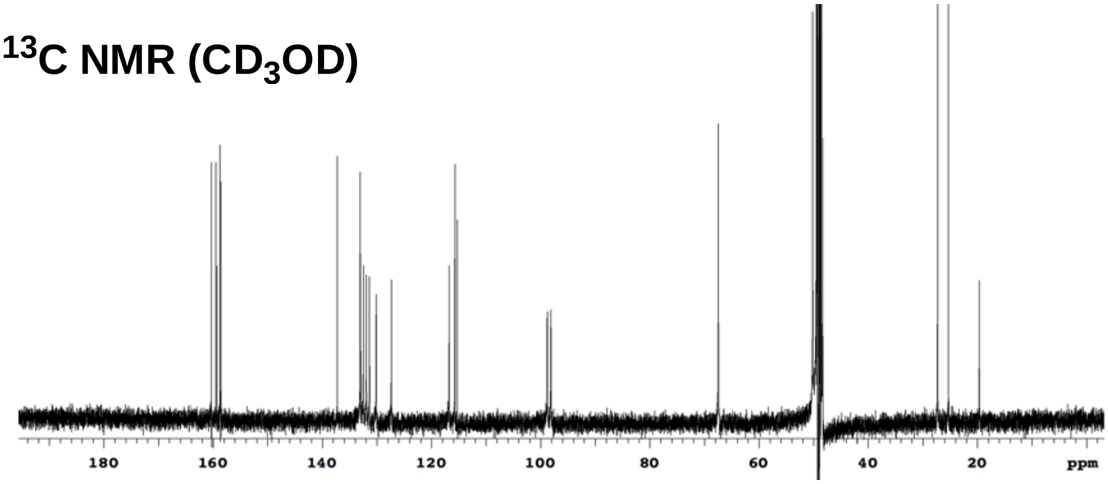


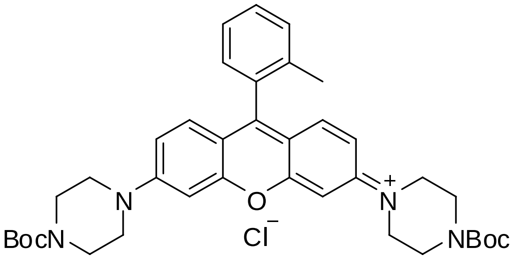
22i

Rosamine 22i

Green solid (93 mg, 98 %). R_f = 0.30 (10% MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, CD_3OD) δ 7.58-7.45 (m, 3H), 7.29-7.19 (m, 7H), 3.86-3.81 (m, 8H), 3.74-3.72 (m, 4H), 2.07 (s, 3H), 1.85-1.75 (m, 6H); ^{13}C NMR (125 MHz, CD_3OD) δ 160.3, 159.5, 159.3, 158.7, 158.6, 137.3, 133.1, 133.0, 132.5, 131.9, 131.3, 130.1, 127.3, 116.8, 115.7, 115.6, 115.3, 98.7, 98.1, 67.4, 50.2, 48.4, 27.3, 25.3, 19.6; IR (thin film) 1646, 1590, 1480, 1415, 1387, 1236, 1189 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{29}H_{31}N_2O_2$ 439.2386; found 439.2384.


^{1}H NMR (CD_3OD)

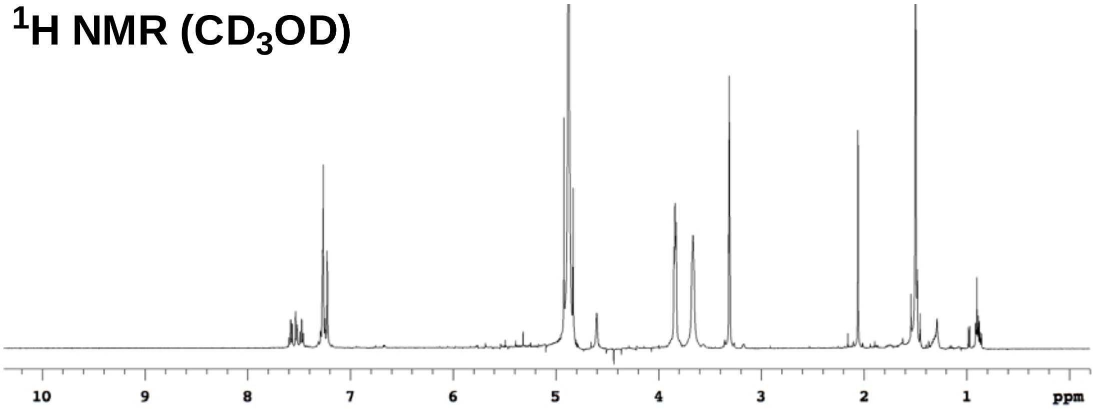

^{13}C NMR (CD_3OD)

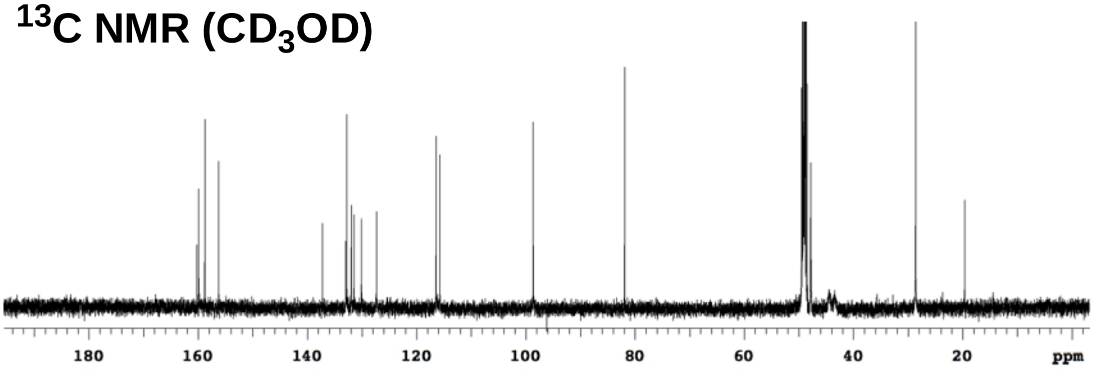


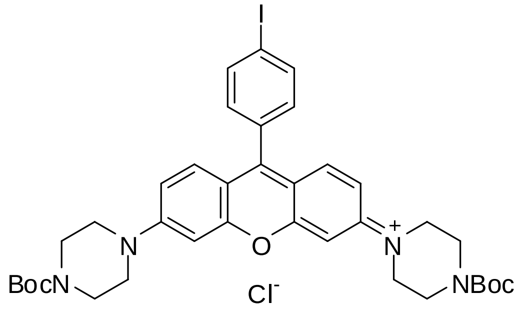
14i

Rosamine 14i

Purple solid (121 mg, 90 %). R_f = 0.37 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, CD_3OD) δ 7.60-7.46 (m, 3H), 7.29-7.22 (m, 7H), 3.84 (br, 8H), 3.67 (br, 8H), 2.06 (s, 3H), 1.50 (s, 18H); ^{13}C NMR (125 MHz, CD_3OD) δ 160.2, 159.9, 158.8, 156.2, 137.2, 133.0, 132.8, 131.9, 131.4, 130.1, 127.3, 116.4, 115.7, 98.6, 81.9, 47.8 (2C), 28.6, 19.6; IR (thin film) 1693, 1646, 1591, 1480, 1413, 1388, 1227, 1161 cm^{-1}; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{38}H_{47}N_4O_5$ 639.3546; found 639.3553.


^{1}H NMR (CD_3OD)

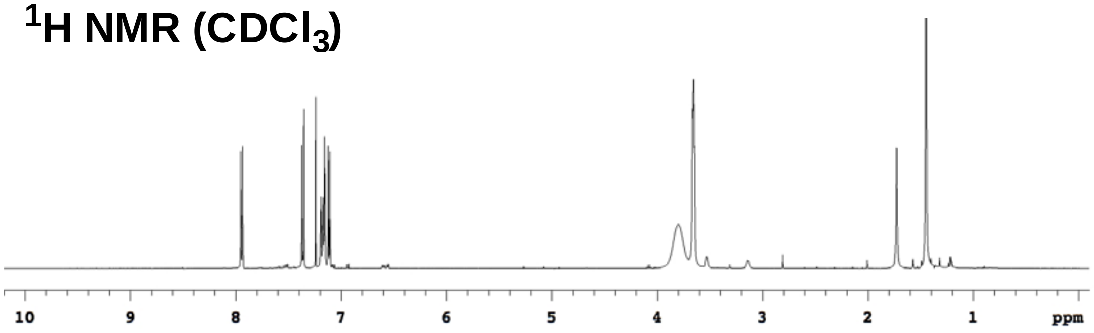

^{13}C NMR (CD_3OD)

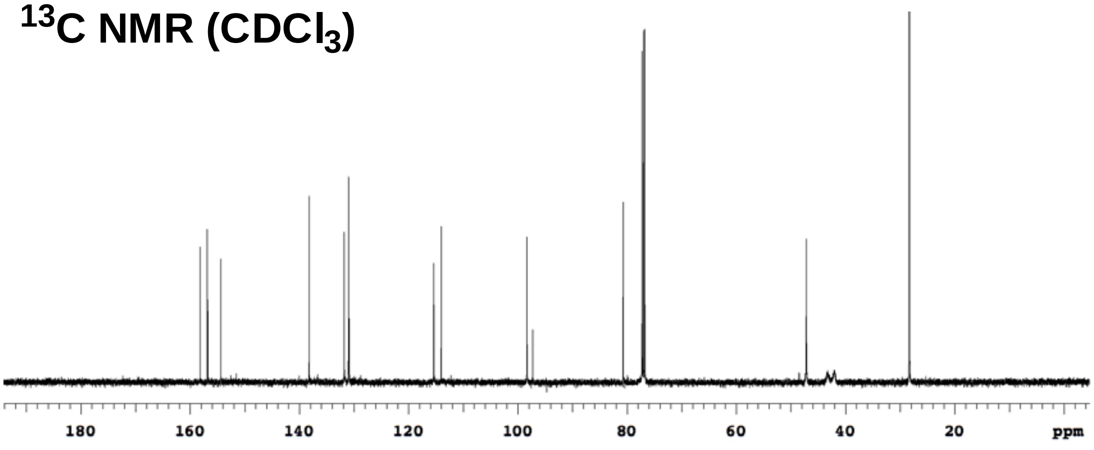


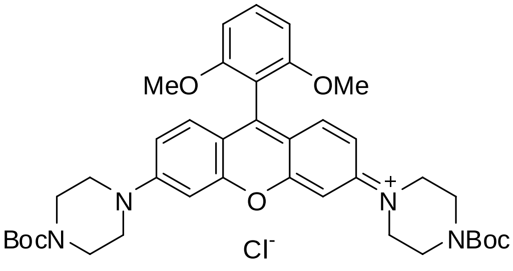
14d

Rosamine 14d

Green solid (107 mg, 68 %). R_f = 0.31 (10% MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 (d, 2H, *J* = 8.3 Hz), 7.37 (d, 2H, *J* = 9.5 Hz), 7.19-7.15 (m, 4H), 7.11 (d, 2H, *J* = 8.3 Hz), 3.80 (br, 8H), 3.67-3.65 (m, 8H), 1.45 (s, 18H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.2, 156.9, 156.8, 154.4, 138.2, 131.8, 131.0, 130.8, 115.4, 114.0, 98.3, 97.3, 80.7, 47.2(2C), 28.3; IR (thin film) 1694, 1644, 1592, 1479, 1415, 1387, 1226, 1161 cm^{-1}; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{37}H_{44}N_4O_5$ 751.2356; found 751.2342.


^{1}H NMR ($CDCl_3$)

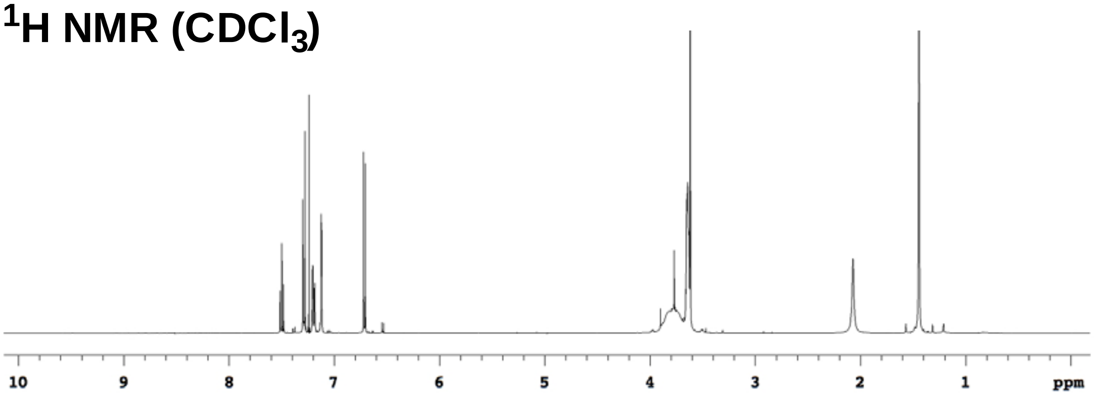

^{13}C NMR ($CDCl_3$)

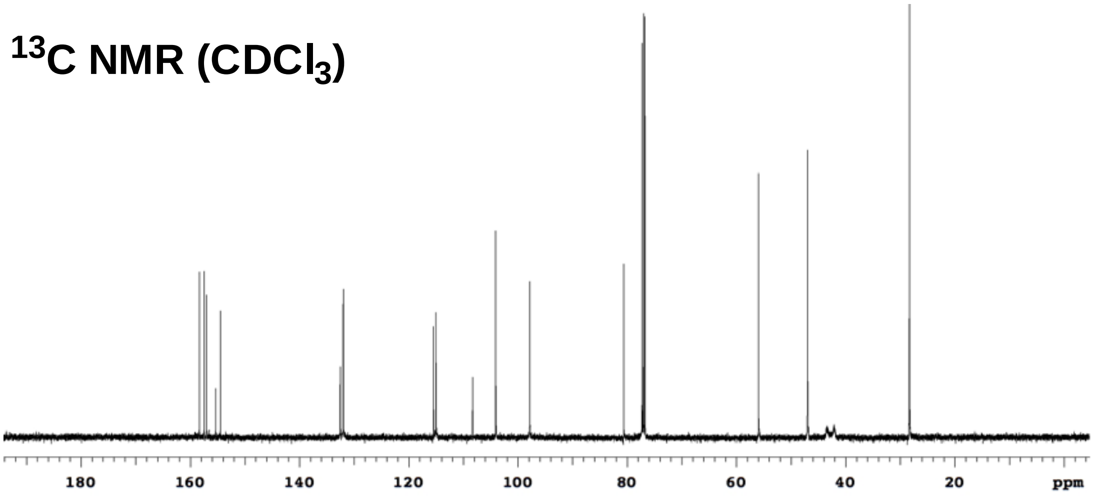


14h

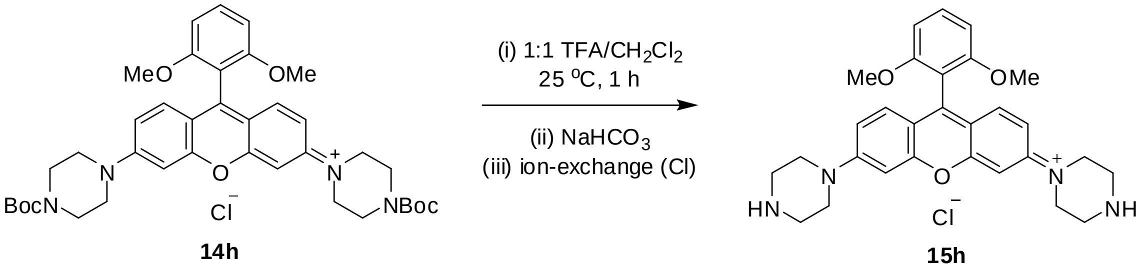
Rosamine 14h

Purple solid (135 mg, 94 %). R_f = 0.31 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.50 (t, 1H, *J* = 8.5 Hz), 7.29 (d, 2H, *J* = 9.4 Hz), 7.20 (dd, 2H, *J* = 9.4, 2.5 Hz), 7.12 (d, 2H, *J* = 2.5 Hz), 6.71 (d, 2H, *J* = 8.5 Hz), 3.77 (br, 8H), 3.65-3.63 (m, 8H), 3.62 (s, 6H), 1.44 (s, 18H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 158.3, 157.4, 157.0, 155.4, 154.5, 132.5, 132.0, 115.5, 115.0, 108.3, 104.1, 97.8, 80.6, 55.9, 47.0 (2C), 28.3. HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{39}H_{49}N_4O_7$ 685.3601; found 685.3598.

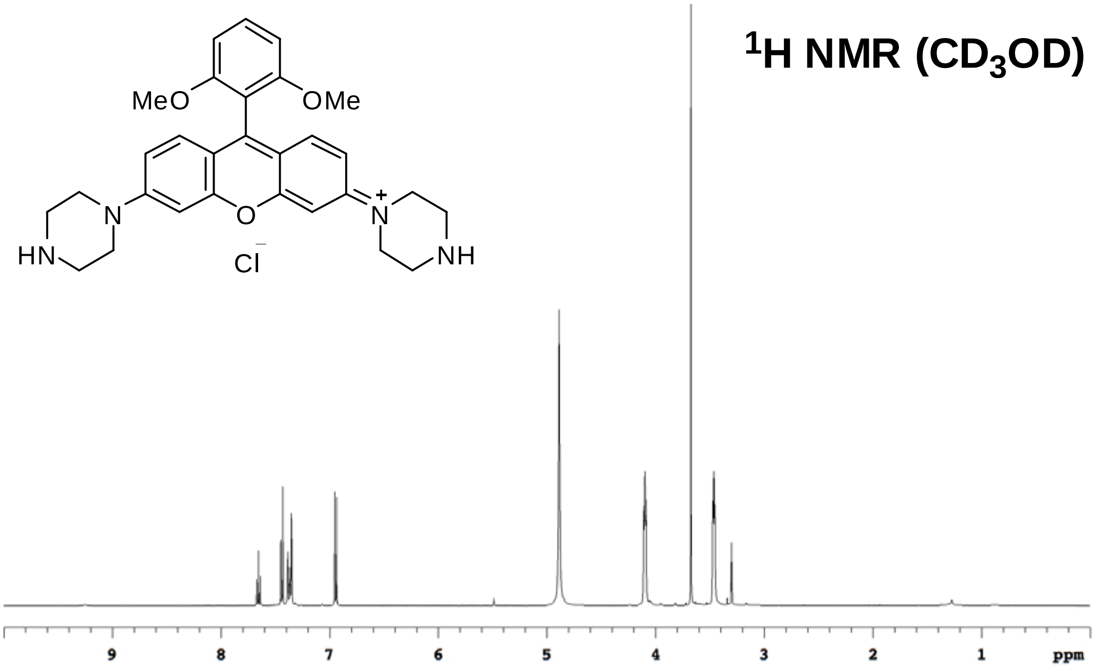


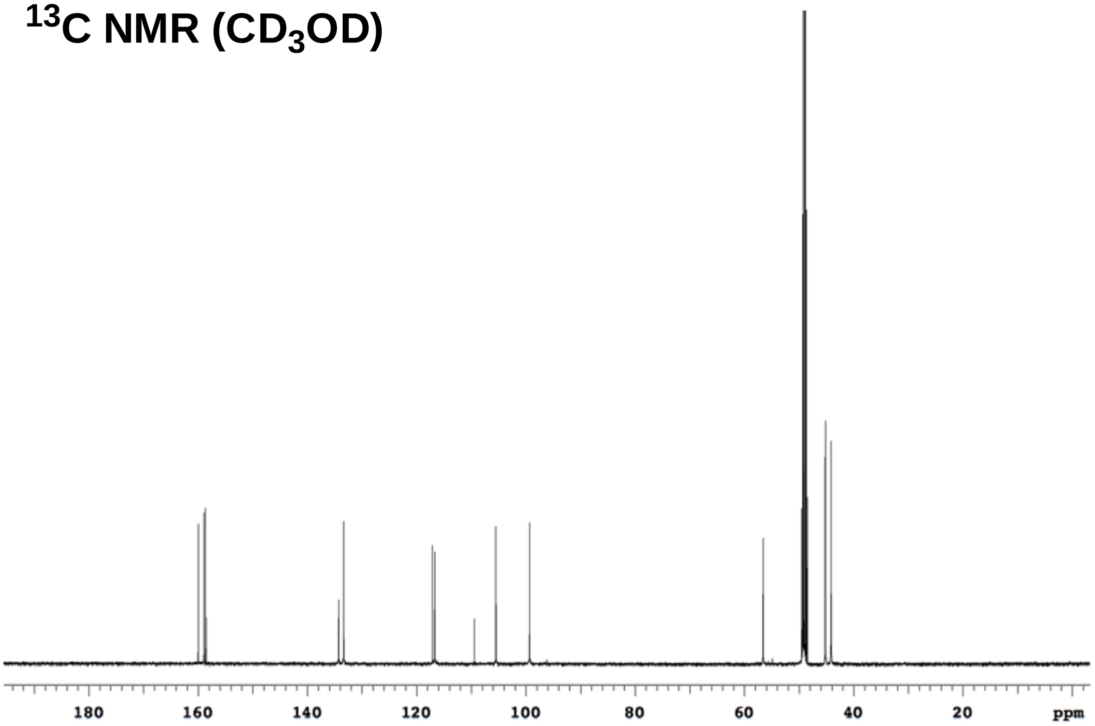


Synthesis of rosamine 15h

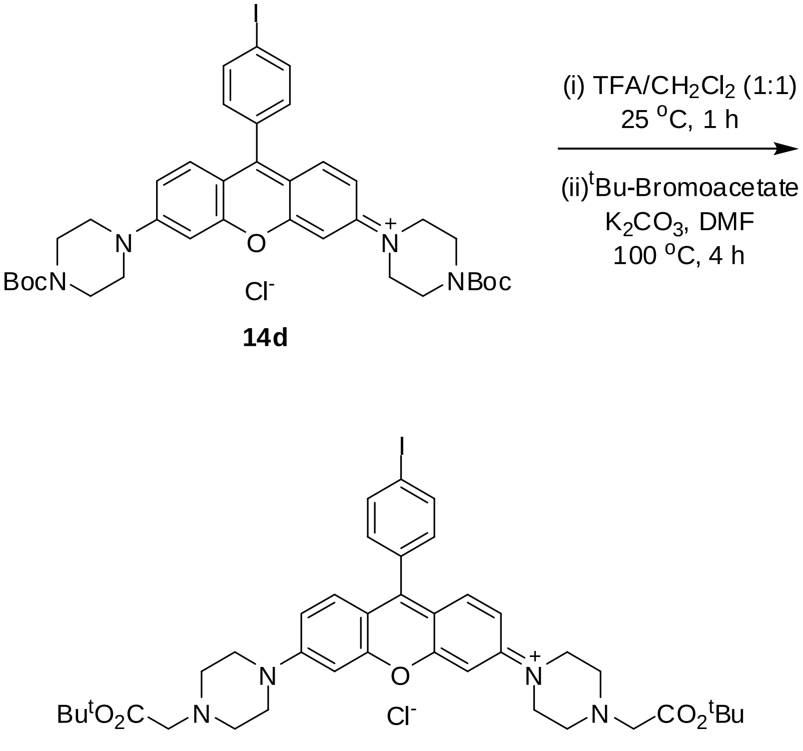


A solution of rosamine **14h** (50 mg, 0.07 mmol) in 5 mL TFA/CH_2Cl_2 (1:1) was stirred at room temperature for 1 h. The solvents were removed with a N_2 stream. The residue was dissolved in 15 mL ^{i}PrOH/CH_2Cl_2 (1:1) and washed with saturated $NaHCO_3$ $_{(aq.)}$, water, and brine then dried over Na_2SO_4. The solvents were removed under reduced pressure. The residue was dissolved in 10 mL MeOH and 0.5 g Amberlite IRA-400 (Cl) ion exchange resin was added. The mixture was stirred at room temperature for 1h and filtered through celite. The solvent was removed under reduced pressure. The ion-exchange process was repeated twice. The crude product was purified by reverse phase MPLC (H_2O - 50 % CH_3CN/H_2O) to give the pure product **15h** (30 mg, 83 %) as a purple solid. ^{1}H NMR (500 MHz, CD_3OD) δ 7.66 (t, 1H, *J* = 8.6 Hz), 7.44 (d, 2H, *J* = 9.3 Hz), 7.37 (dd, 2H, *J* = 9.3, 2.2Hz), 7.35 (d, 2H, *J* = 2.2 Hz), 6.94 (d, 2H, *J* = 8.6 Hz), 4.10 (t, 8H, *J* = 5.3 Hz), 3.67 (s, 6H), 3.46 (t, 8H, *J* = 5.3 Hz); ^{13}C NMR (125 MHz, CD_3OD) δ 160.0, 158.9, 158.6, 158.5, 134.3, 133.4, 117.1, 116.7, 109.4, 105.5, 99.3, 56.6, 45.2, 44.1; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{29}H_{33}N_4O_3$ 485.2553; found 485.2557.


MeO
OMe
HN
N
O
N
NH
Cl
^{1}H NMR (CD$_3$OD)
9
8
7
6
5
4
3
2
1
ppm


^{13}C NMR (CD$_3$OD)
180
160
140
120
100
80
60
40
20
ppm

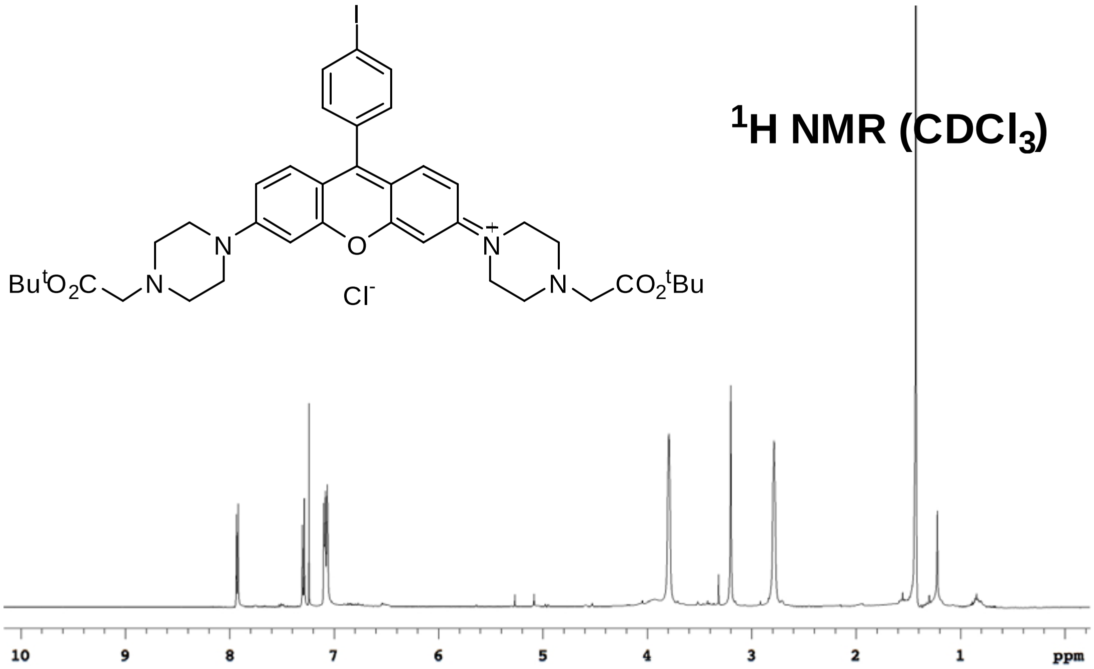
Synthesis of rosamine 23



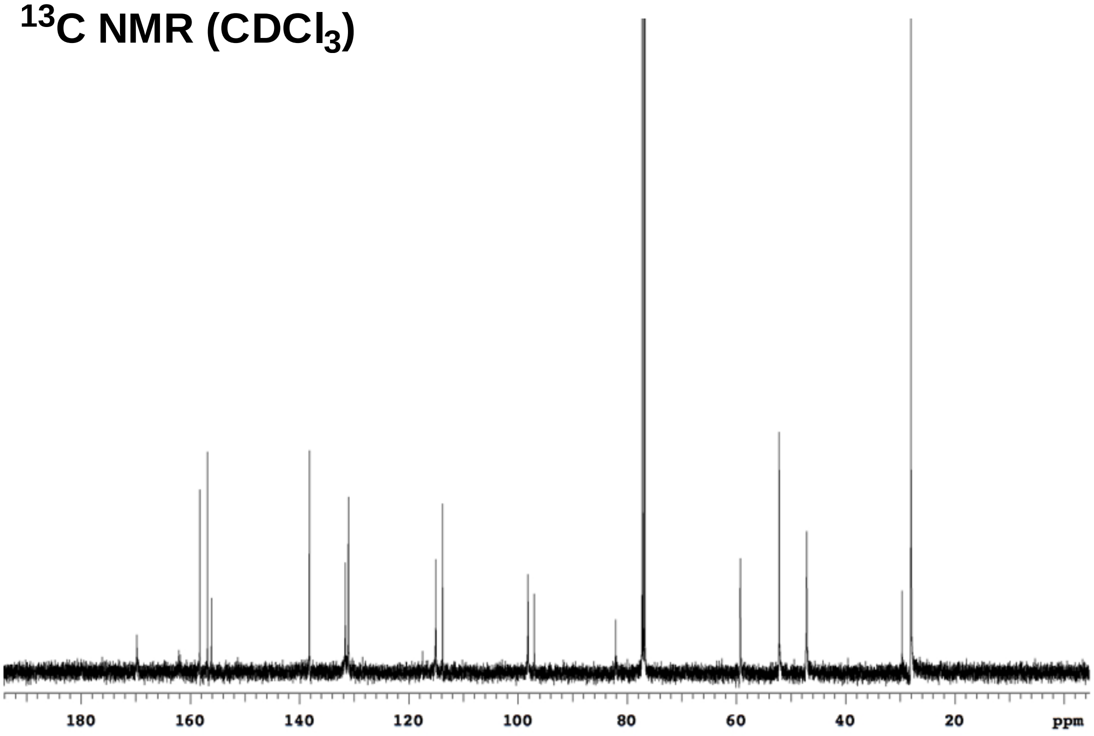
14d

23

Rosamine **14d** (39 mg, 0.05 mmol) was dissolved in 2 mL TFA/CH_2Cl_2 (1:1). The solution was stirred at 25 °C for 1 h. The solvents were removed with nitrogen stream. The residue was dissolved in 2 mL DMF then K_2CO_3 (69 mg, 0.5 mmol) and tert-butyl bromoacetate (74 μL, 0.5 mmol) were added. The mixture was heated to 100 °C and stirred for 4 h. After cooling to room temperature, the mixture was diluted with CH_2Cl_2 and washed with water, dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatography (10 % MeOH/CH_2Cl_2, $R_f = 0.28$) to afford the pure product **23** (31 mg, 76 %) as a purple solid. ^1H NMR (500 MHz, $CDCl_3$) δ 7.93 (d, 2H, $J = 8.4$ Hz), 7.29 (d, 2H, $J = 9.4$ Hz), 7.10-7.06 (m, 6H), 3.79 (br, 8H), 3.20 (s, 4H), 2.79 (br, 8H), 1.43 (s, 18H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 169.8, 158.2, 156.8, 156.1, 138.2, 131.6, 131.1, 131.0, 115.0, 113.8, 98.1, 97.0, 82.1, 59.3, 52.2, 47.1, 28.0; IR (thin film) 1735, 1688, 1644, 1594, 1482, 1391, 1236, 1195, 1152 cm^{-1}; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{39}H_{48}IN_4O_5$ 779.2669; found 779.2671.

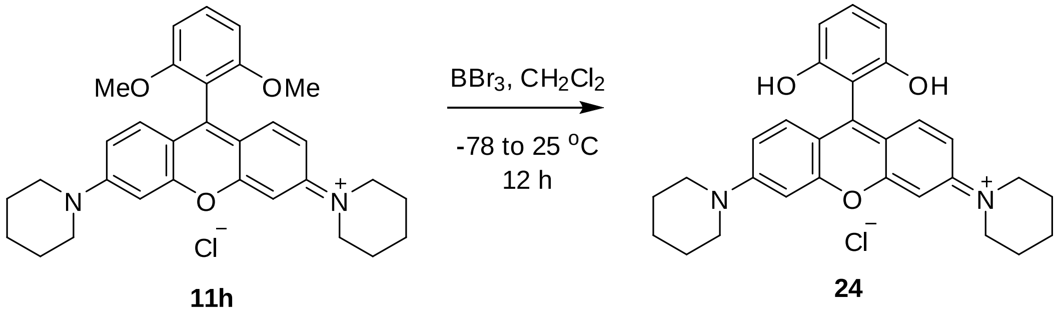


^{1}H NMR (CDCl$_3$)

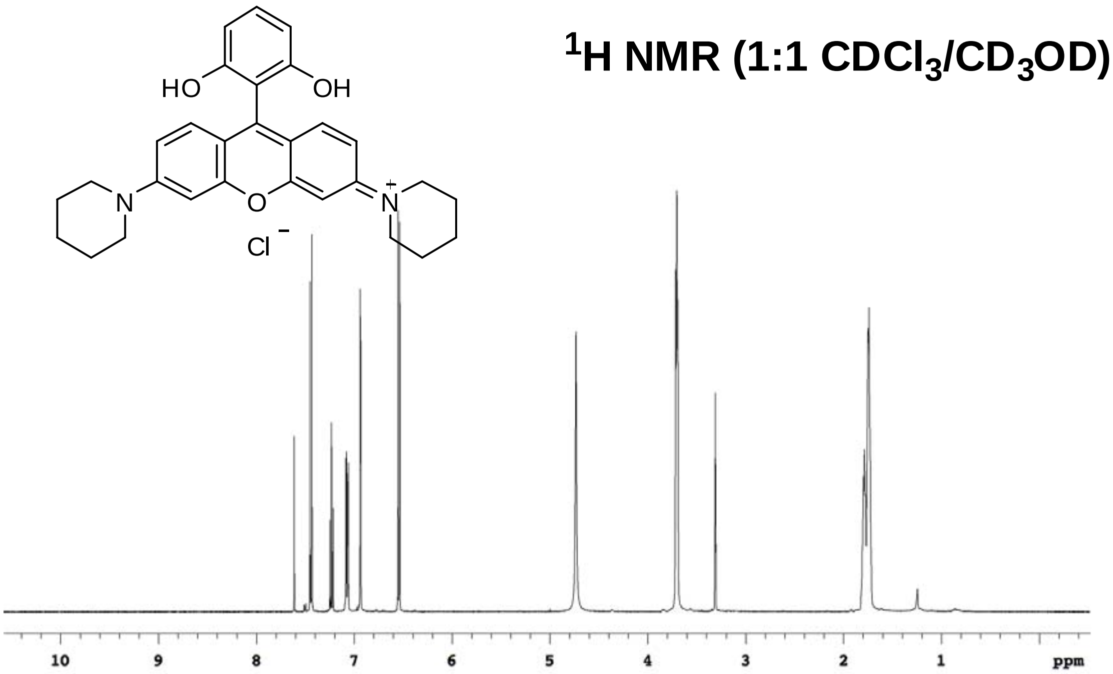


^{13}C NMR (CDCl$_3$)

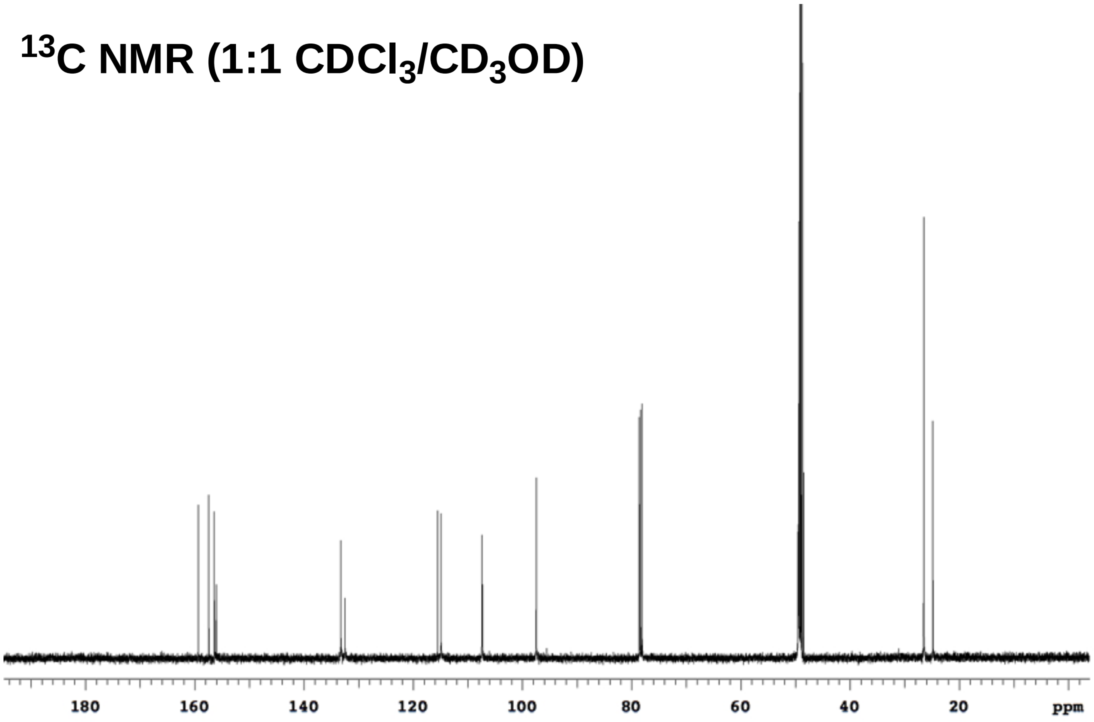
Synthesis of Rosamine 24



BBr_3 (0.19 mL, 2.0 mmol) was added dropwise over 1 min to the solution of **11h** (104 mg, 0.2 mmol) in 4 mL CH_2Cl_2 at -78 °C. The solution was warmed to room temperature slowly and stirred for 12 h. The reaction was quenched with ice-water and the mixture was extracted with 1:1 ^{i}PrOH/CH_2Cl_2 (3 x 15 mL). The organic layer was washed with water (1 x 20 mL), brine (1 x 20 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The residue was dissolved in 10 mL MeOH and 0.5 g Amberlite IRA-400 (Cl) ion exchange resin was added. The mixture was stirred at room temperature for 1h and filtered through celite. The solvent was removed under reduced pressure. The ion-exchange process was repeated twice. The crude product was purified by flash chromotagraphy (5 % to 10 % MeOH/CH_2Cl_2) to afford the product **24** (85 mg, 87 %) as a green solid. R_f = 0.15 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$/CD_3OD 1:1) δ 7.44 (d, 2H, *J* = 9.6 Hz), 7.23 (t, 1H, *J* = 8.3 Hz), 7.07 (dd, 2H, *J* = 9.6, 2.6 Hz), 6.94 (d, 2H, *J* = 2.6 Hz), 6.54 (d, 2H, *J* = 8.3 Hz), 3.71-3.69 (m, 8H), 1.79-1.74 (m, 12H); ^{13}C NMR (125 MHz, $CDCl_3$/CD_3OD 1:1) δ 159.3, 157.4, 156.4, 156.1, 133.2, 132.5, 115.5, 114.9, 107.4, 107.3, 97.5, 49.4, 26.5, 24.8; HRMS (ESI) m/z calcd for $(M-Cl)^+$ $C_{29}H_{31}N_2O_3$ 455.2335; found 455.2340.

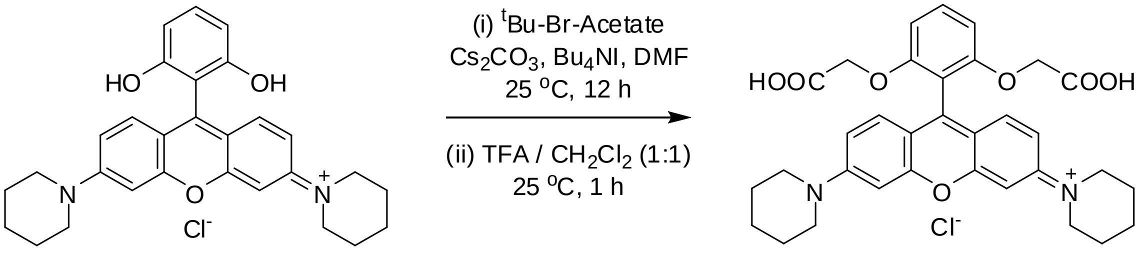


^{1}H NMR (1:1 $CDCl_3/CD_3OD$)



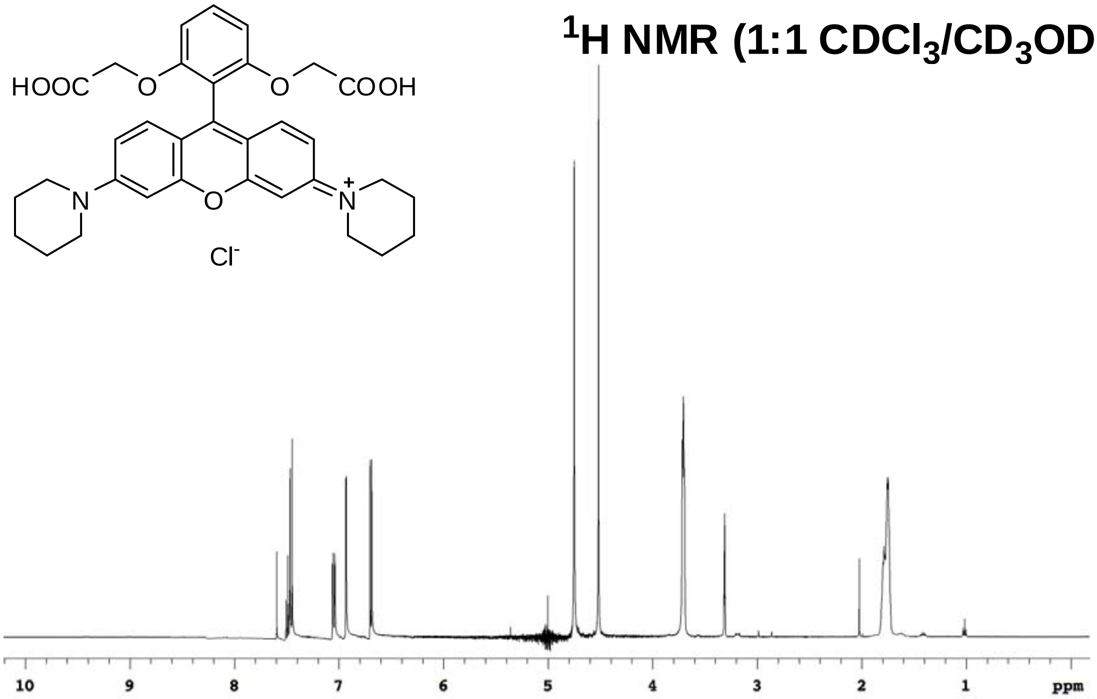
^{13}C NMR (1:1 $CDCl_3/CD_3OD$)

Synthesis of Rosamine 25



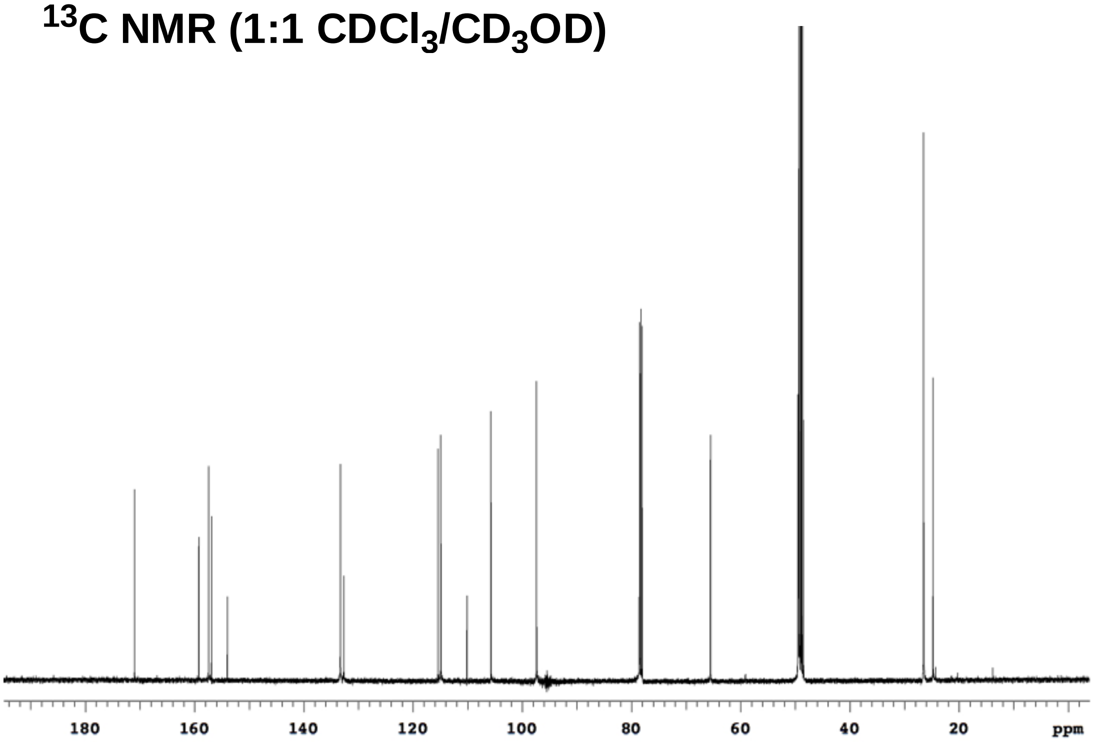
24 **25**

Rosamine **24** (137 mg, 0.28 mmol), Cs_2CO_3 (456 mg, 1.4 mmol), Bu_4NI (310 mg, 0.84 mmol) were dissolved in 5 mL DMF and tert-butyl bromoacetate (0.41 mL, 2.8 mmol) was added. The reaction mixture was stirred at 25 ^{o}C for 12 h then diluted with 30 mL CH_2Cl_2, washed with H_2O (3 x 20 mL), brine (1 x 20 mL) and dried over Na_2SO_4. The solvents were removed under reduced pressure and the residue was passed through a short pad of silica gel eluting with 5 % MeOH/ CH_2Cl_2 to give the crude product which was used in the next step without further purification. The crude material was dissolved in 10 mL TFA/CH_2Cl_2 (1:1) and stirred at 25 ^{o}C for 1 h. The solvents were removed with a N_2 stream. The residue was dissolved in 30 mL CH_2Cl_2, washed with H_2O (2 x 20 mL), brine (1 x 20 mL) and dried over Na_2SO_4. The solvents were removed under reduced pressure. The residue was dissolved in 20 mL MeOH/ CH_2Cl_2 (1:1) and 1.0 g Amberlite IRA-400 (Cl) ion exchange resin was added. The mixture was stirred at room temperature for 1 h and filtered through celite. The solvent was removed under reduced pressure. The ion-exchange process was repeated twice. The crude product was purified by reverse phase MPLC (H_2O – 60 % CH_3CN/H_2O) to afford the pure product **25** as a green solid (128 mg, 75 %). ^{1}H NMR (500 MHz, $CDCl_3/CD_3OD$ 1:1) δ 7.49 (t, 1H, J = 8.5 Hz), 7.46 (d, 2H, J = 9.6 Hz), 7.05 (dd, 2H, J = 9.6, 2.5Hz), 6.93 (d, 2H, J = 2.5 Hz), 6.69 (d, 2H, J = 8.5 Hz), 4.52 (s, 4H), 3.72-3.70 (m, 8H), 1.80-1.75 (m, 12H); ^{13}C NMR (125 MHz, $CDCl_3/CD_3OD$ 1:1) δ 171.0, 159.2, 157.4, 156.9, 154.0, 133.3, 132.7, 115.4, 114.9, 110.1, 105.7, 97.4, 65.5, 49.4, 26.5, 24.8. HRMS (ESI) m/z calcd for $(M\text{-}Cl)^{+}$ $C_{33}H_{35}N_2O_7$ 571.2444; found 571.2451.

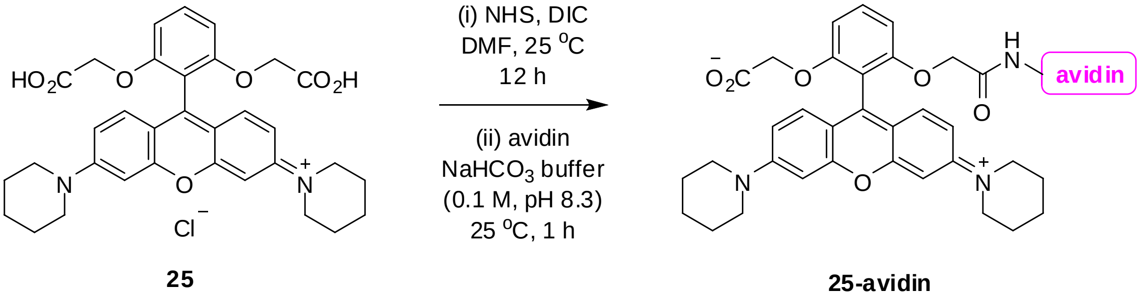


^{1}H NMR (1:1 $CDCl_3/CD_3OD$)

^{13}C NMR (1:1 $CDCl_3/CD_3OD$)



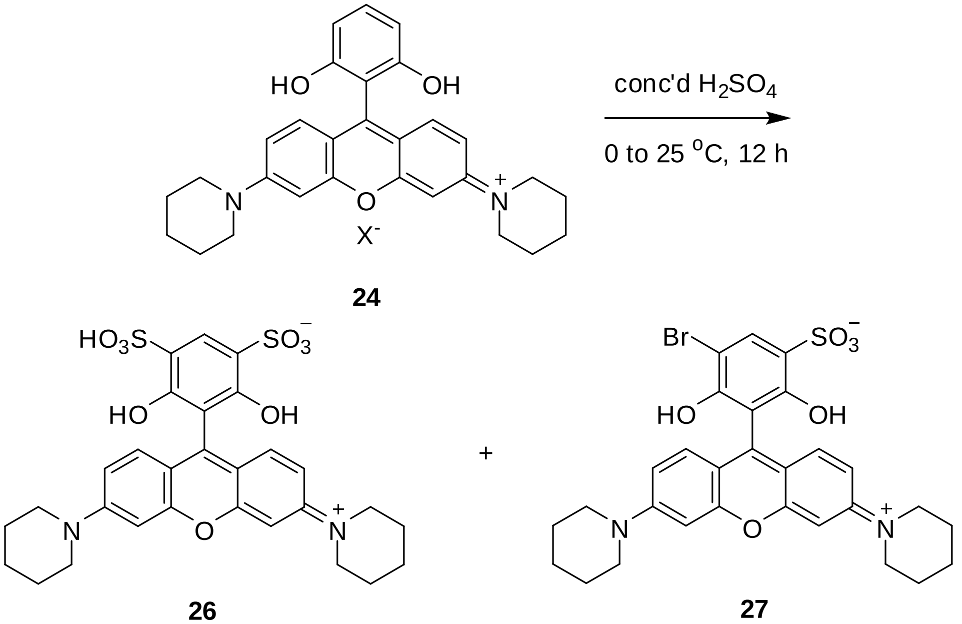
Labeling of Avidin with Rosamine 25


(i) NHS, DIC
DMF, 25 °C
12 h
(ii) avidin
NaHCO3 buffer
(0.1 M, pH 8.3)
25 °C, 1 h
avidin
25
25-avidin

Rosamine **25** (6.1 mg, 0.01 mmol) and NHS (1.2 mg, 0.01 mmol) were dissolved in 0.3 mL DMF, then DIC (1.5 µL, 0.01 mmol) was added. The mixture was stirred at room temperature in the dark for overnight.

The above solution (15 µL, 5 eq.) was added to the solution of avidin (6.6 mg, 1 eq.) in 1 mL sodium bicarbonate buffer (0.1 M, pH 8.3). The mixture was stirred at room temperature in the dark for 1h. The unreacted dye was removed by PD-10 (Sephadex G-25) column to afford the labeled avidin.

Synthesis of water-soluble rosamines 26 and 27



Rosamine **24** (68 mg, 0.15 mmol) was dissolved in 2 mL concentrated H_2SO_4. The mixture was stirred at 0 °C for 4 h then warmed to 25 °C and stirred for 12 h. The reaction mixture was poured into ice-water and extracted with 1:1 $^{i}PrOH/CH_2Cl_2$. The organic layer was evaporated to dryness under reduced pressure and the residue was purified by flash chromatography (5 to 15 % $MeOH/CH_2Cl_2$) to afford product **26** (32 mg, 35 %) as a purple solid and **27** (39 mg, 42 %) as a green solid.

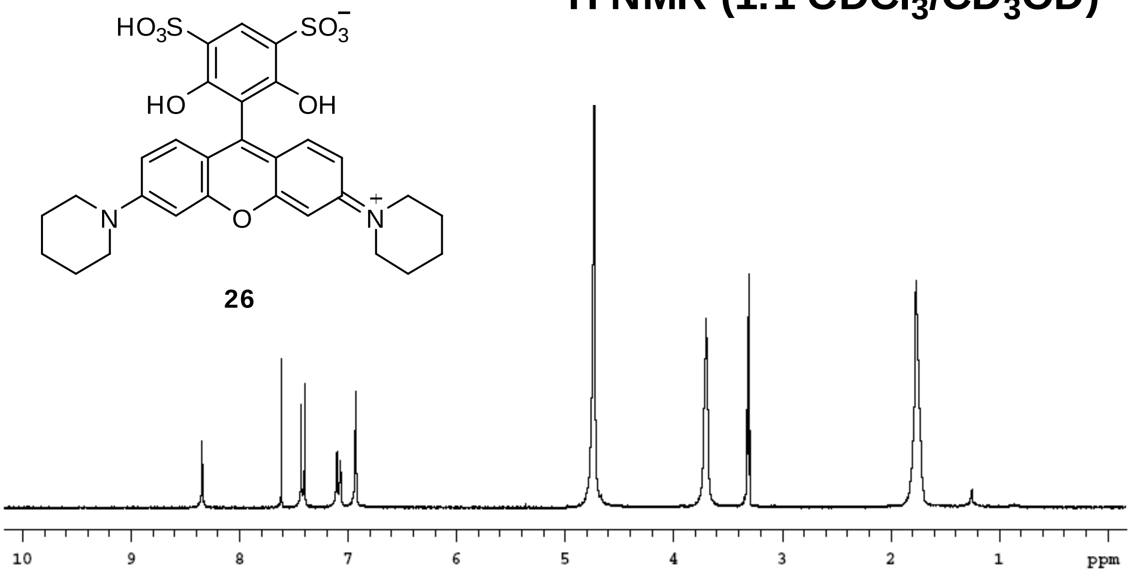
26: ^1H NMR (300 MHz, 1:1 $CD_3OD/CDCl_3$) δ 8.34 (s, 1H), 7.42 (d, 2H, J = 9.6 Hz), 7.09 (dd, 2H, J = 9.6, 2.4 Hz), 6.93 (d, 2H, J = 2.4 Hz), 3.72-3.69 (m, 8H), 1.76 (br, 12H); ^{13}C NMR (125 MHz, 1:1 $CD_3OD/CDCl_3$) δ 159.1, 157.3, 154.8, 152.9, 132.7, 130.2, 121.8, 115.3, 115.2, 107.6, 97.7, 49.6, 26.5, 24.7; MS (ESI) m/z calcd for $(M+H)^+$ $C_{29}H_{31}N_2O_9S_2$ 615.15; found 615.15.

27: ^1H NMR (500 MHz, 1:1 $CD_3OD/CDCl_3$) δ 8.05 (s, 1H), 7.38 (d, 2H, J = 9.6 Hz), 7.07 (dd, 2H, J = 9.6, 2.5 Hz), 6.92 (d, 2H, J = 2.5 Hz), 3.72-3.70 (m, 8H), 1.80-1.75 (m, 12H); ^{13}C NMR (125 MHz, 1:1 $CD_3OD/CDCl_3$) δ 159.2, 157.3, 154.5, 153.2, 152.3, 133.7, 132.6, 123.3, 115.1 (2 peaks: 115.13, 115.10), 109.1, 101.2, 97.4, 49.4, 26.4, 24.7; MS (ESI) m/z calcd for $(M-H)^-$ $C_{29}H_{28}BrN_2O_6S$ 611.09; found 611.02.

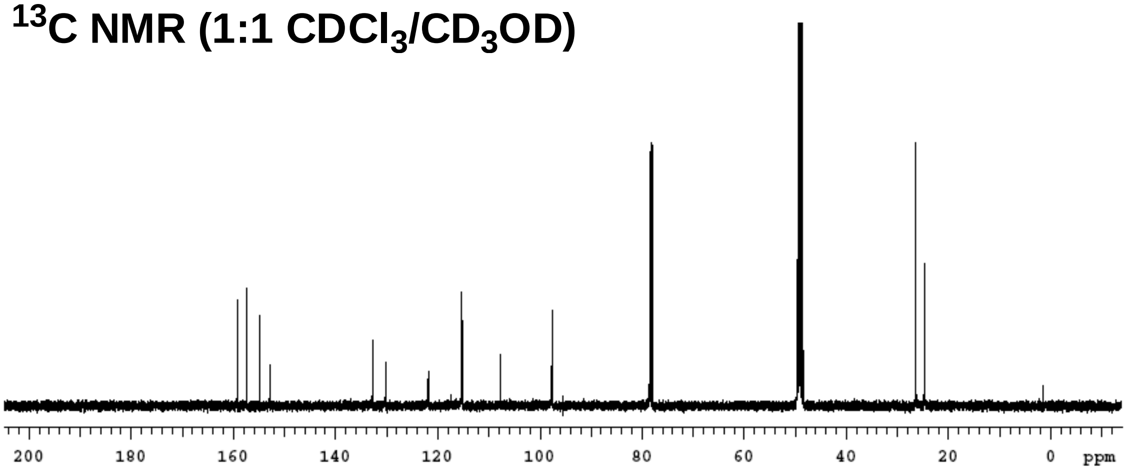
HO$_3$S
SO$_3^-$
HO
OH
N
O
N+

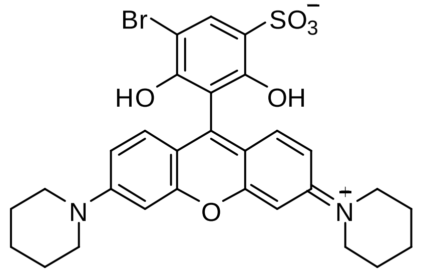
26

^{1}H NMR (1:1 CDCl$_3$/CD$_3$OD)



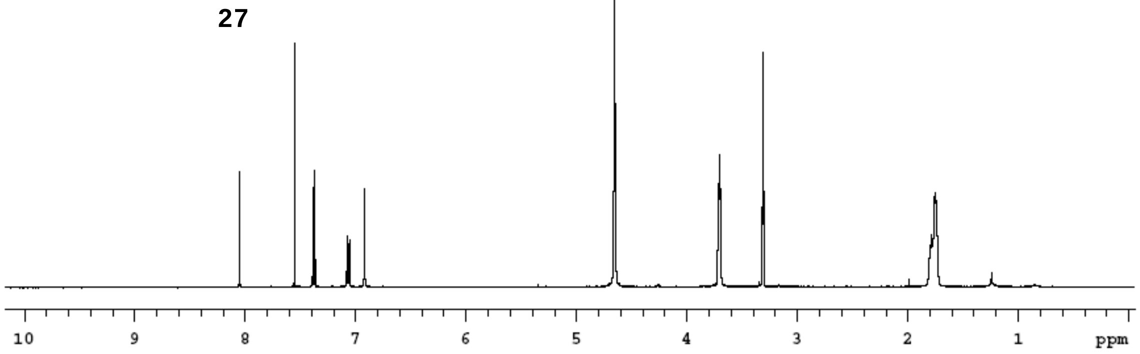
^{13}C NMR (1:1 CDCl$_3$/CD$_3$OD)



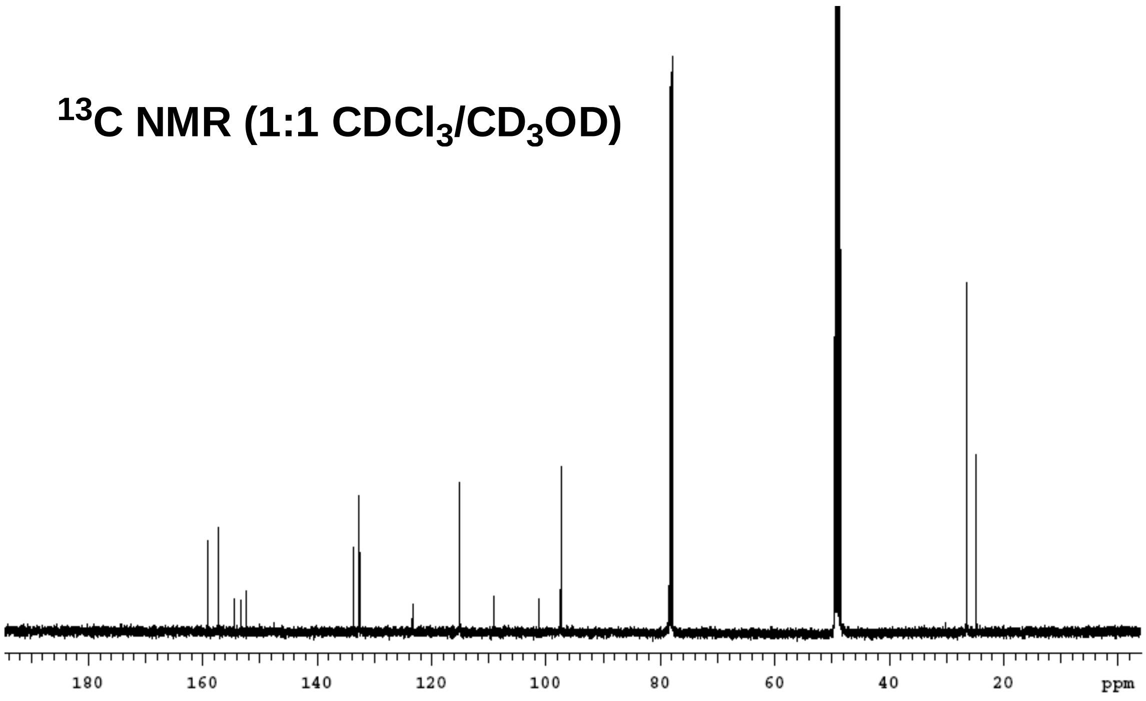

Br
SO3−
HO
OH
N
O
N+

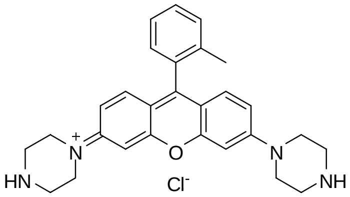
27

^{1}H NMR (1:1 $CDCl_3$/CD_3OD)



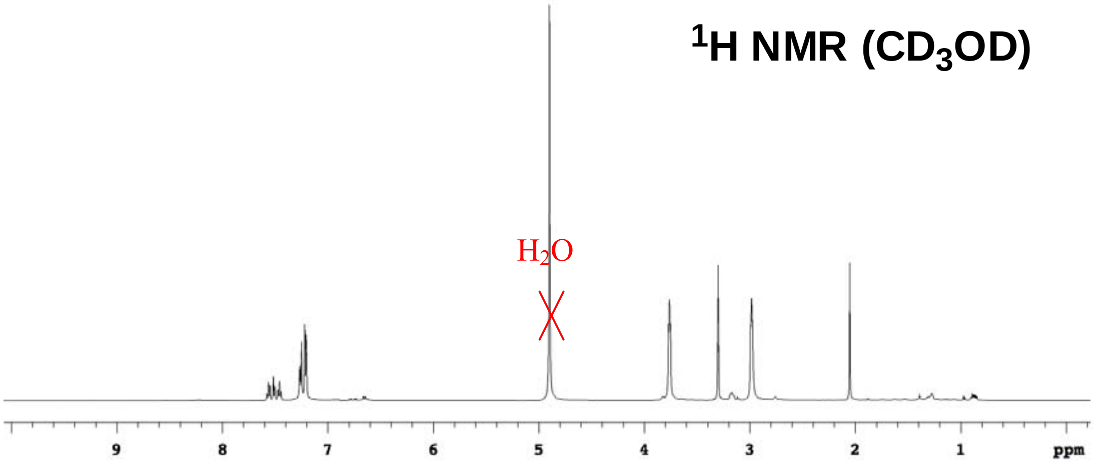
^{13}C NMR (1:1 $CDCl_3$/CD_3OD)

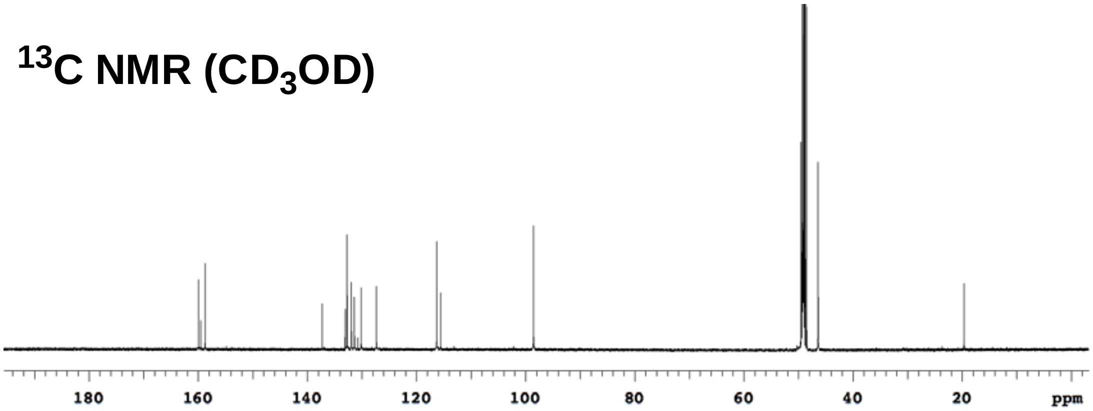




28

^{1}H NMR (500 MHz, CD_3OD) δ 7.58-7.55 (m, 1H), 7.52-7.50 (m, 1H), 7.47-7.44 (m, 1H), 7.27-7.25 (m, 3H), 7.22-7.20 (m, 4H), 3.76 (t, 8H, *J* = 4.9 Hz), 2.98 (t, 8H, *J* = 4.9 Hz), 2.05 (s, 3H); ^{13}C NMR (125 MHz, CD_3OD) δ 159.9, 159.5, 158.7, 137.2, 133.0, 132.7, 131.9, 131.4, 130.1, 127.3, 116.3, 115.5, 98.5, 46.4 (2C), 19.7; HRMS (ESI) m/z calcd for $(M\text{-}Cl)^+$ $C_{28}H_{31}N_4O$ 439.2498; found 439.2500.





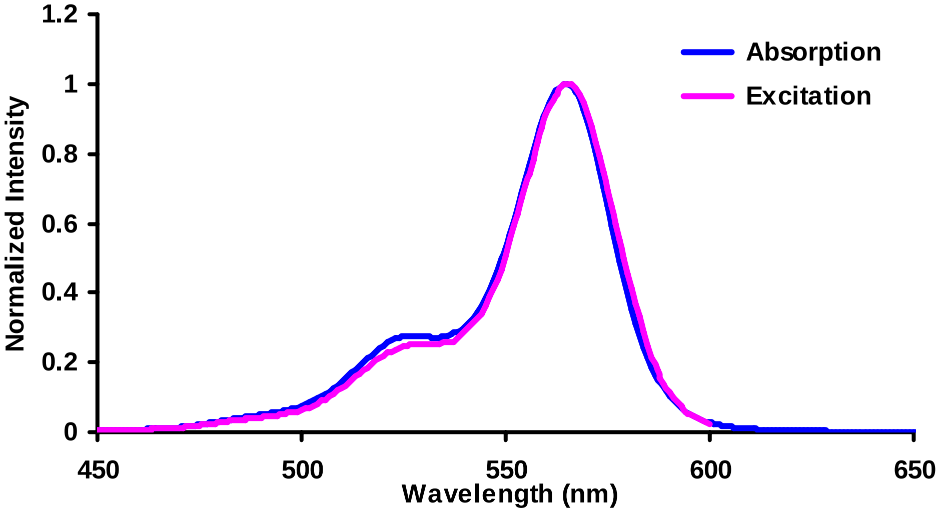


Figure D1. Absorption (10^{-6} M) and excitation (10^{-7} M, emission detected at 582 nm) spectra for rosamine **11i** in CH_2Cl_2.

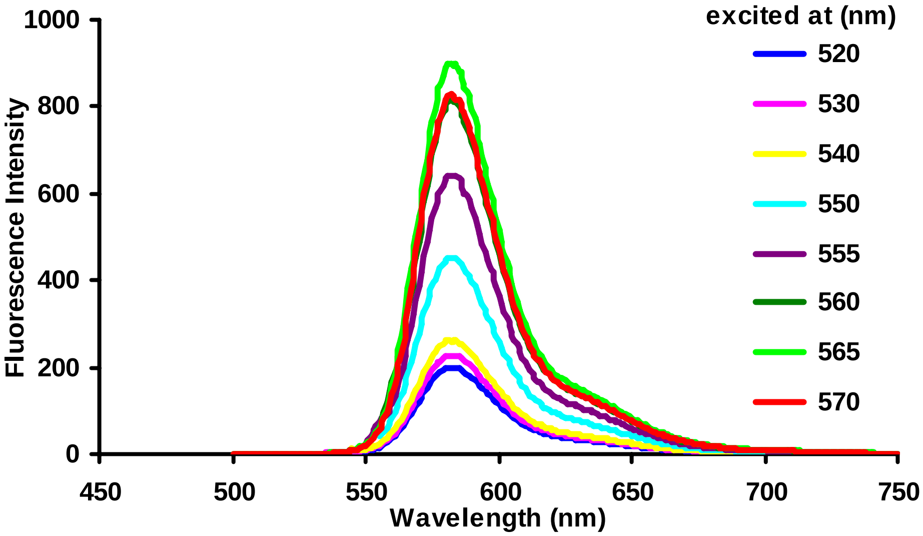


Figure D2. Fluorescence spectra of **11i** (10^{-7} M) at different excitation wavelengths in CH_2Cl_2.

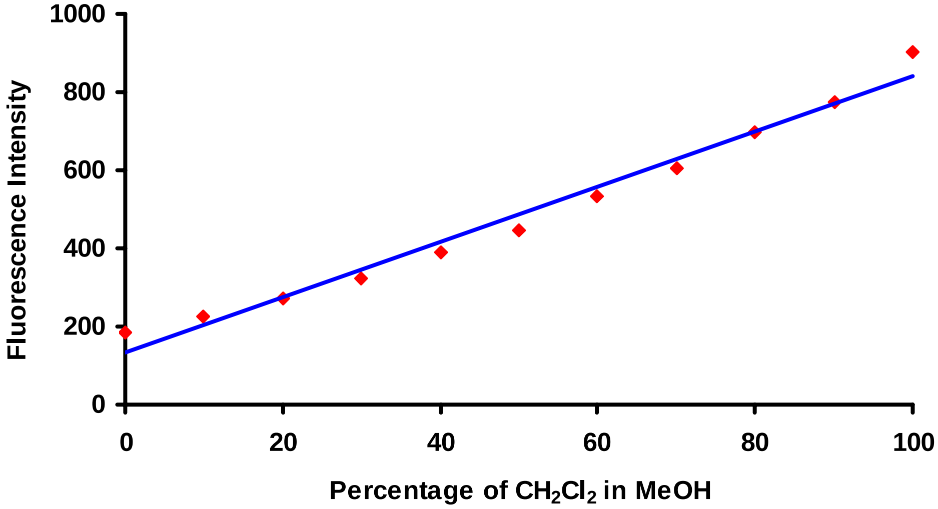


Figure D3. Fluorescence intensity (ex. 562 nm, em. 582nm) of **11i** at 1.4 x 10^{-7} M in a CH_2Cl_2/MeOH mixed solvent system.

Biological activity studies:

Materials. ER-Tracker Blue-White DPX, LysoTracker Blue DND-22, rhodamine 123 (Rh123), Sytox Green were purchased from Molecular Probes, Invitrogen (Oregon, USA). Annexin V-PE Apoptosis Detection Kit 1 BD Biosciences (CA, USA). Cell culture reagents were purchased from Gibco, Invitrogen (Auckland, NZ). RNase A, propidium iodide and MTT were purchased from Sigma (St Loist, USA).

Cell cultures. HSC2 oral cavity human squamous carcinoma cells were obtained from Health Science Research Resources Bank (Japan). HK1 cell-line is a gift from the University of Hong Kong. Both cell-lines were grown in MEM medium supplemented with 10% FBS. MCF-7 breast carcinoma and HCT-116 colon carcinoma were obtained from American Tissue Culture Collection (Virginia, USA) and maintained in RPMI 1640 medium supplemented with 10% FBS. OKF6, an immortalized human oral keratinocyte cell-line and NP69, an immortalized human nasopharyngeal epithelial cell-line were obtained from BWH Cell Culture and Microscopy Core at Harvard Institutes of Medicine and the University of Hong Kong respectively, and were maintained in keratinocyte serum-free medium supplemented with epidermal growth factor, bovine pituitary extract and a final Ca^{2+} concentration of 0.3 mM.

***In vitro* proliferation assay.** Approximately 3 000 - 5 000 cells/well exponentially growing cells were seeded into 96-well plate and allowed to adhere overnight. Cells were treated with each compound at concentrations ranging from 0.01-10 μM. At the end of incubation period, 15 μl of MTT solution (5.0 mg/ml in PBS) was added and incubated for an additional 4 h. Medium and excessive MTT were aspirated and formazan formed was solubilized with 100 μl of DMSO. Absorbance, as a measurement of viable cell number was read at 570 nm with ThermoLabsystems OpsysMR microplate spectrometer. At least three independent experiments were performed and results are presented as an average.

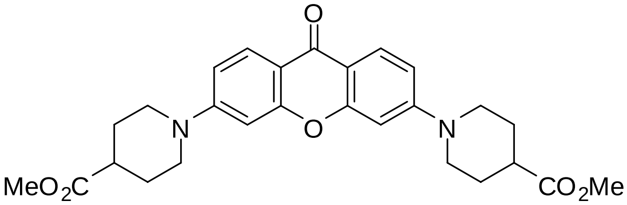
Cellular localization. HSC2 cells grown on round glass coverslips in 12-well plate were co-incubated with 100 nM of compound **11** together with organelle-specific fluorescence probes. The endoplasmic reticulum was labeled with 100 nM of ER-Tracker Blue-White DPX, the lysosomes were stained with 500 nM of LysoTracker Blue DND-22, and the mitochrondria was tracked with 100 nM of Rh123 respectively for 15-30 min of incubation at room temperature. After incubation, cells were gently rinsed in PBS to remove free dyes, and the stained cells were observed using Olympus DSU spinning disk confocal microscope configured with a PlanApo ×63 oil objective and iXon EM + (Andor Technology) digital camera. Fluorescent images of *X*-*Y* sections at 0.2 μm were collected sequentially using Olympus Cell^{R} software. Organelle-specific fluorescence probes were respectively excited at 365 nm to illuminate ER-tracker and LysoTracker, at 494 nm for Rh123 and at 575 nm for compound **11**.

AnnexinV-FITC apoptosis analysis. HSC2 cells grown in 60-mm dishes at 50% confluency were treated with 0.5 μM of compound **11**. At various treatment intervals, floating cells in the medium were pooled together with the adherent cells after trypsinization and were washed twice with cold PBS. The cells were resuspended with 1X binding buffer at 1 x 10^6 cells/ml. A 100μl of cell suspension was transferred to a 5 ml tube followed by 5 μl of AnnexinV-FITC and 5 μl of propidium iodide (200 μg/ml in PBS). The cells were gently mixed and incubated for 15 min at RT in the dark before analysed on a FACSCalibur flow cytometer with 488 nm argon laser. The fluorescence data of 10 000 cells were collected with the FL1 detector with 530/30 band pass filter to collect Annexin-FITC fluorescence, and the FL3 detector with a 630 nm long pass filter to collect propidium iodide fluorescence.

Cell cycle analysis. HSC2 cells were treated and collected as above. Cells were then fixed in 70% ice-cold ethanol (v/v in PBS) overnight at 4 ºC. Following fixation, the cells were washed twice in cold PBS. The pellet was then resuspended in PBS solution containing 20 μg/ml RNase A and 1 μM SYTOX Green for 30 min. The cells were analysed on a FACSCalibur flow cytometer with 488 nm argon laser. The DNA-SYTOX Green fluorescence of 10 000 cells were collected with the FL1 detector with 530/30 band pass filter.

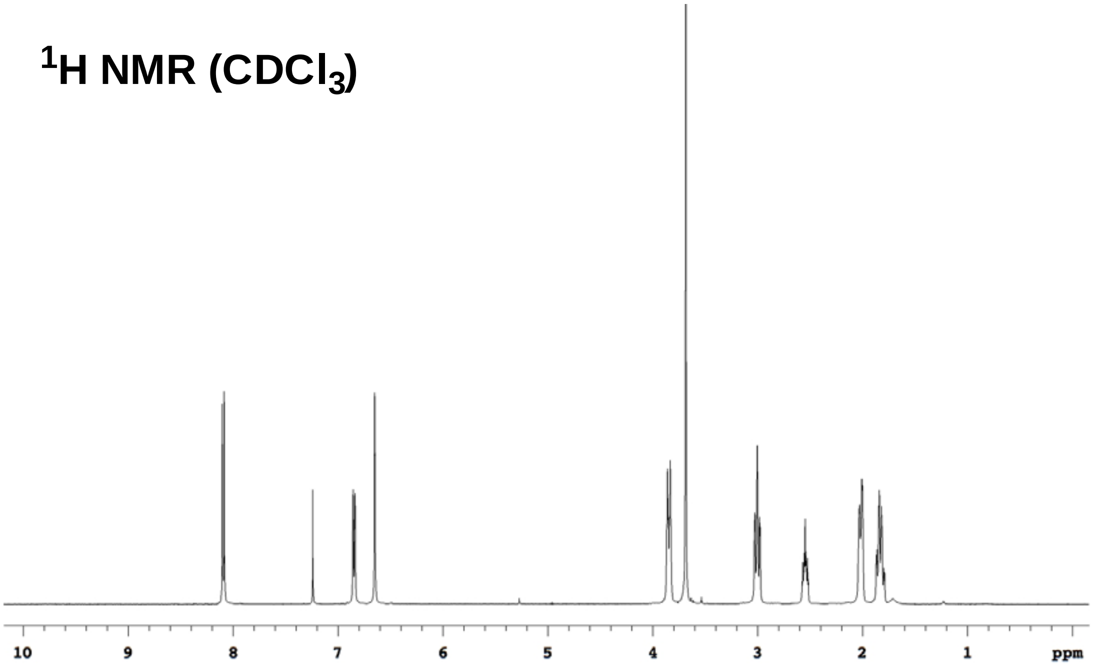
APPENDIX E
EXPERIMENTAL DATA FOR CHAPTER V

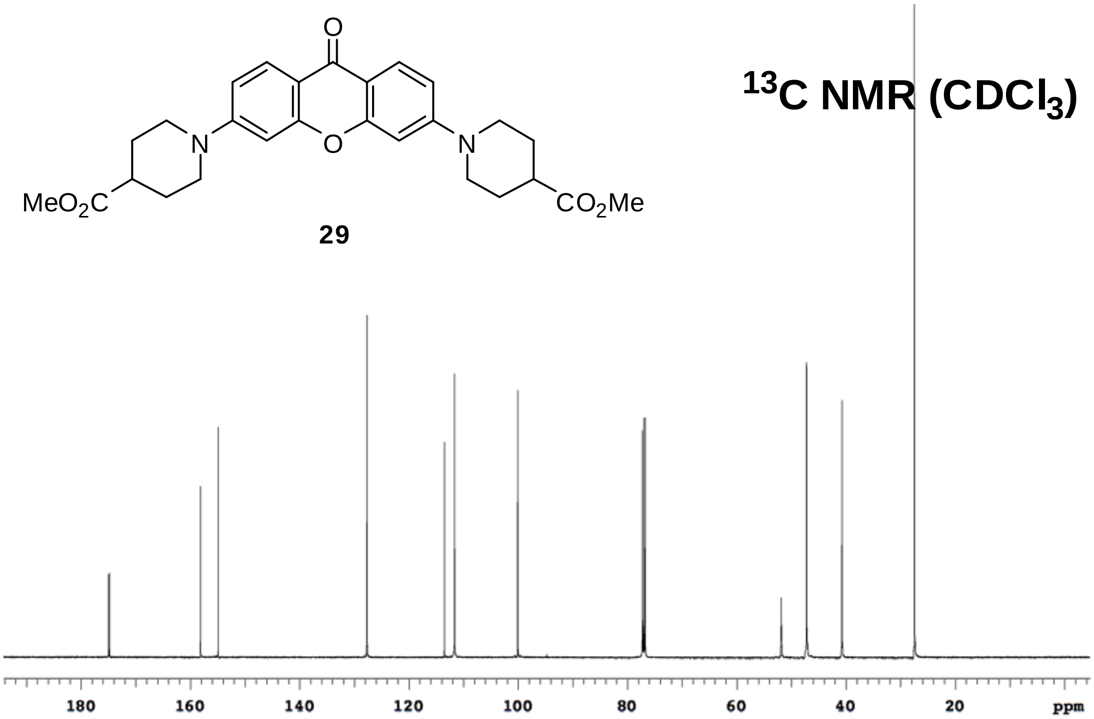
Compound 29



29

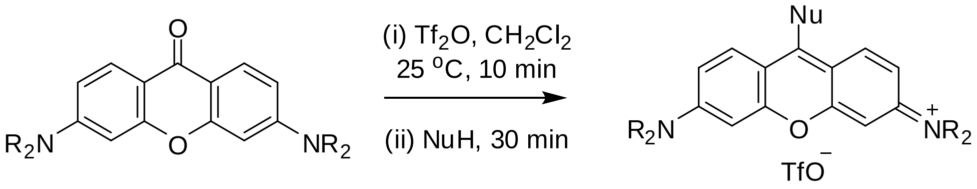
Synthesized according to the procedure described in Appendix D. Light yellow solid (500 mg, 51 %). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.09 (d, 2H, *J* = 8.9 Hz), 6.85 (dd, 2H, *J* = 8.9, 2.3 Hz), 6.65 (d, 2H, *J* = 2.3 Hz), 3.87-3.83 (m, 4H), 3.69 (s, 6H), 3.03-2.98 (m, 4H), 2.58-2.52 (m, 2H), 2.04-2.00 (m, 4H), 1.87-1.79 (m, 4H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 175.0, 174.8, 158.1, 154.9, 127.7, 113.5, 111.6, 100.1, 51.8, 47.2, 40.7, 27.5. MS (ESI) m/z calcd for $(M+H)^+$ $C_{27}H_{31}N_2O_6$ 479.22; found 479.23.





29

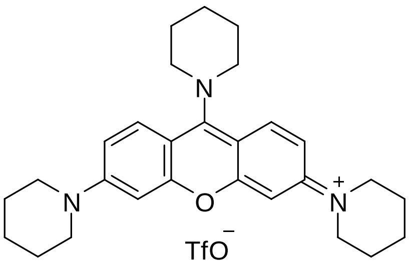
Syntheses of Pyronins 30


R$_2$N
NR$_2$
O
(i) Tf$_2$O, CH$_2$Cl$_2$
25 ^{o}C, 10 min
(ii) NuH, 30 min
Nu
$\overset{+}{N}R_2$
TfO$^-$

30

Tf$_2$O (1.0 mmol) was added dropwise to a solution of 3,6-diamino-xanthen-9-one (0.2 mmol) in 2 mL CH$_2$Cl$_2$ at 25 ^{o}C over 1 min. The reaction mixture was stirred for 10 min then the corresponding nucleophile (NuH) was added dropwise over 1 min. The mixture was stirred for 30 min at 25 ^{o}C. (If the reaction was incomplete, another portion of Tf$_2$O and nucleophile was added). The solvents were removed under reduced pressure and the residue was purified by flash chromatography (2 % to 10 % MeOH/CH$_2$Cl$_2$) to afford the pure product **30**.

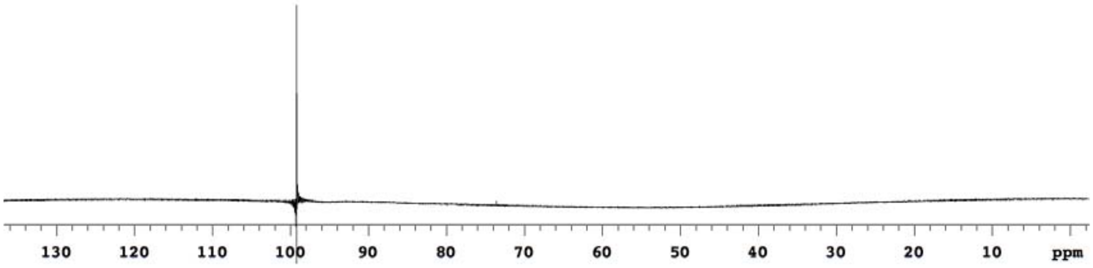
Pyronin 30a

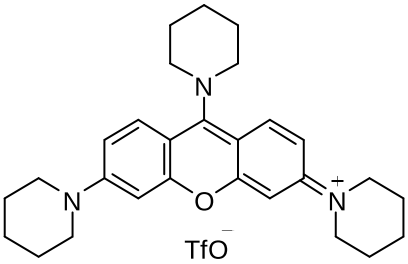


30a

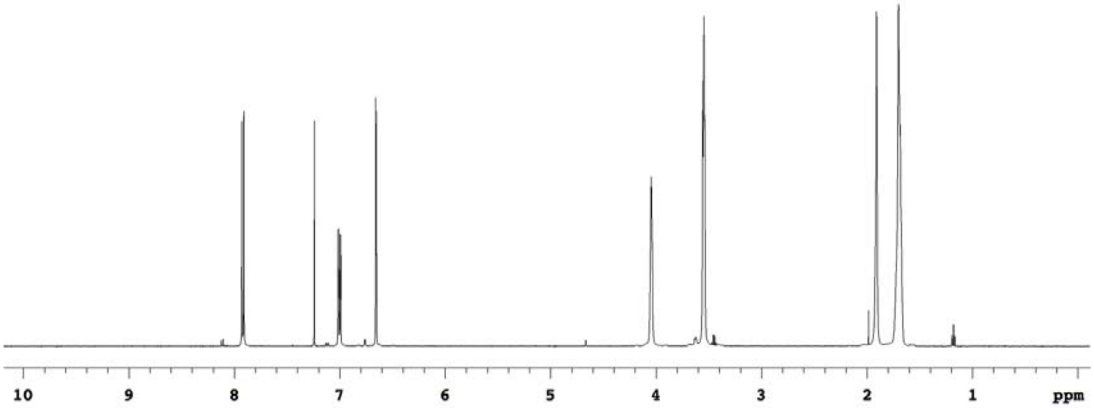
Red solid (94 mg, 81 %). R_f = 0.38 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92 (d, 2H, *J* = 9.7 Hz), 7.00 (dd, 2H, *J* = 9.7, 2.7 Hz), 6.65 (d, 2H, *J* = 2.7 Hz), 4.05 (br, 4H), 3.56-3.54 (m, 8H), 1.91 (br, 6H), 1.70 (br, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 160.6, 158.2, 155.1, 129.8, 120.9 (q, *J* = 322.4 Hz), 112.5, 107.2, 97.8, 57.0, 48.2, 27.0, 25.4, 24.2, 23.7; ^{19}F NMR (282 MHz, $CDCl_3$) δ 99.2 (s). MS (ESI) m/z calcd for $(M\text{-}OTf)^+$ $C_{28}H_{36}N_3O$ 430.29; found 430.28.

^{19}F NMR ($CDCl_3$)

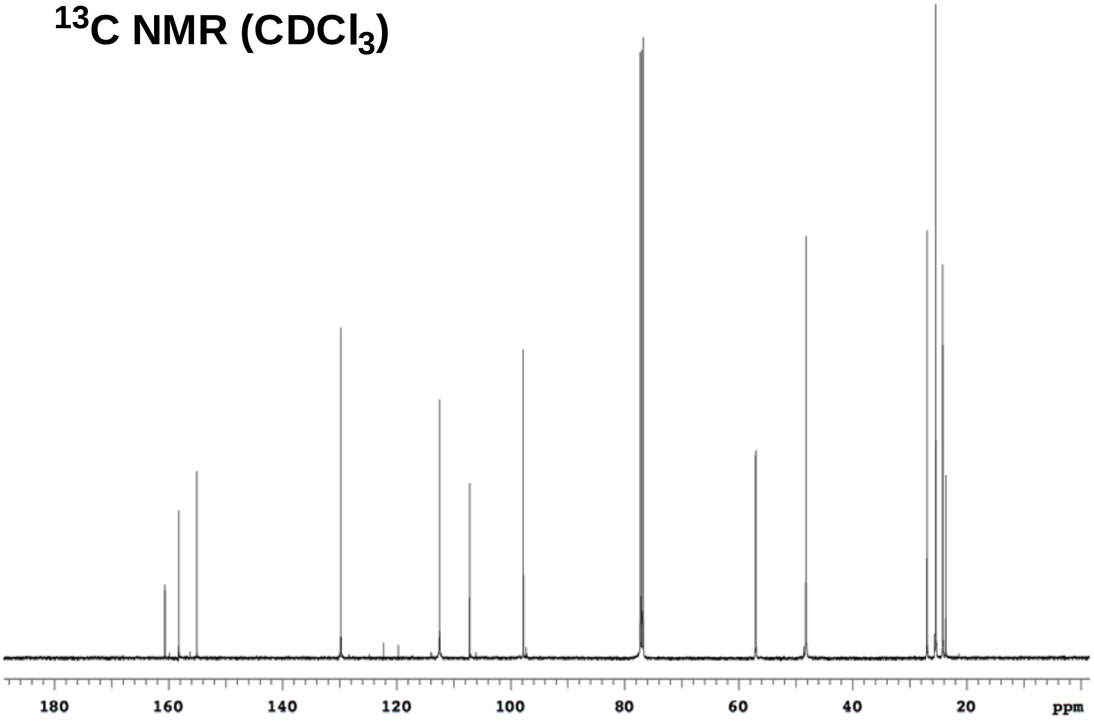




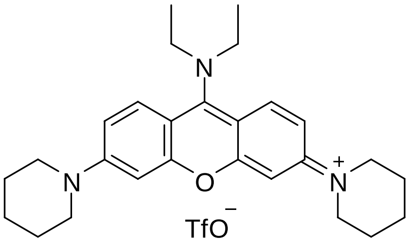
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

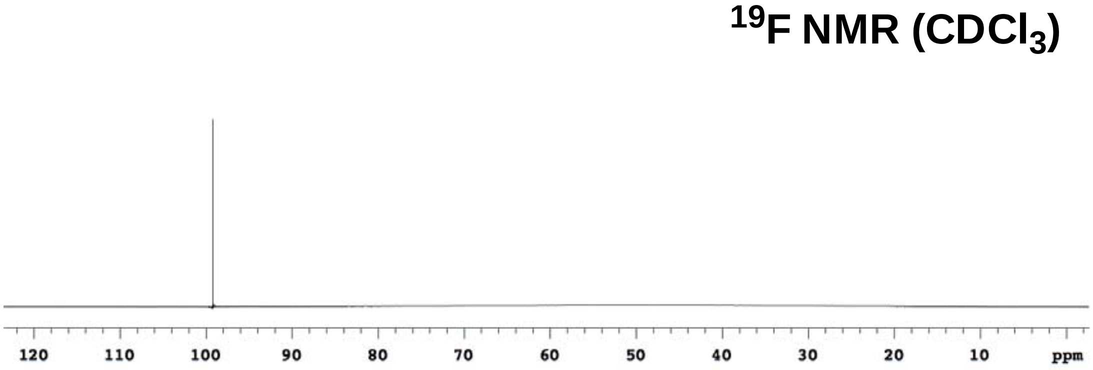


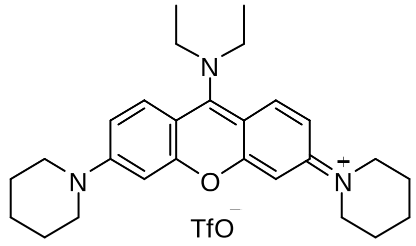
Pyronin 30b



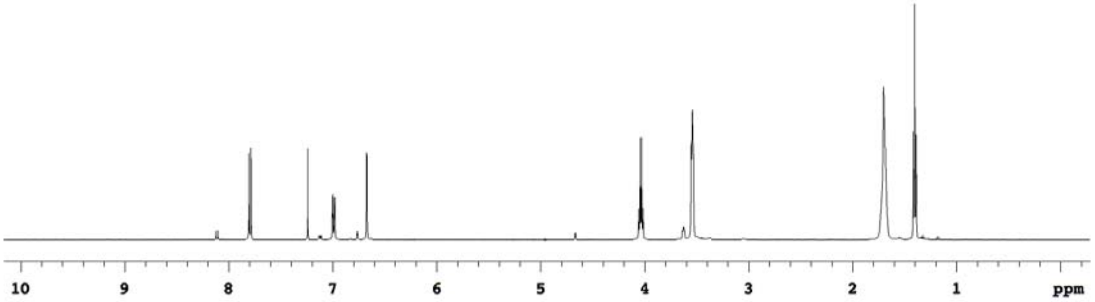
30b

Red solid (69 mg, 61 %). R_f = 0.44 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (d, 2H, *J* = 9.6 Hz), 6.99 (dd, 2H, *J* = 9.6, 2.7 Hz), 6.67 (d, 2H, *J* = 2.7 Hz), 4.04 (q, 4H, *J* = 7.0 Hz), 3.55-3.53 (m, 8H), 1.70 (br, 12H), 1.40 (t, 6H, *J* = 7.0 Hz); ^{13}C NMR (125 MHz, $CDCl_3$) δ 161.7, 158.1, 155.0, 129.6, 120.9 (q, *J* = 320.1 Hz), 112.3, 107.8, 97.8, 50.6, 48.2, 25.4, 24.2, 14.1; ^{19}F NMR (282 MHz, $CDCl_3$) δ 99.2 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ $C_{27}H_{36}N_3O$ 418.29; found 418.28.

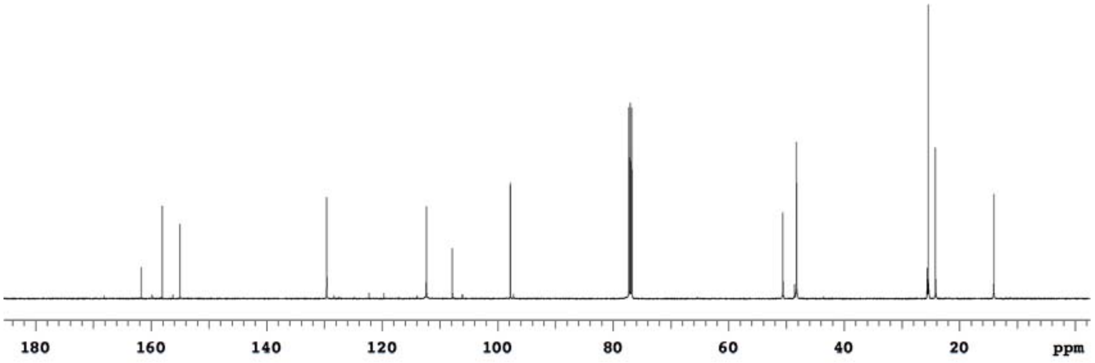




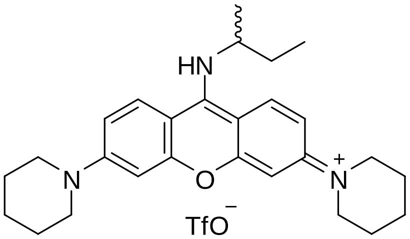
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)



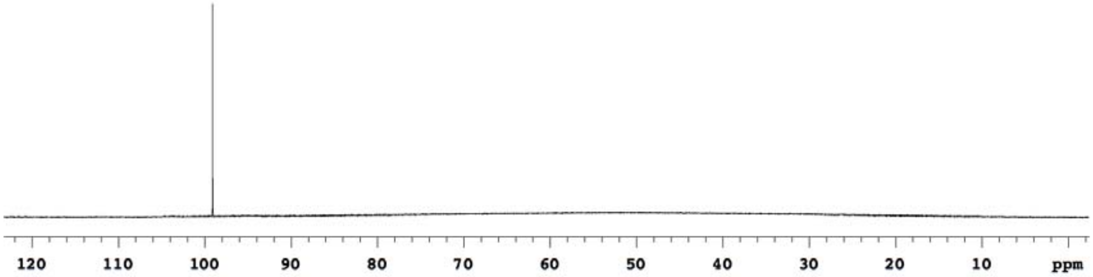
Pyronin 30c

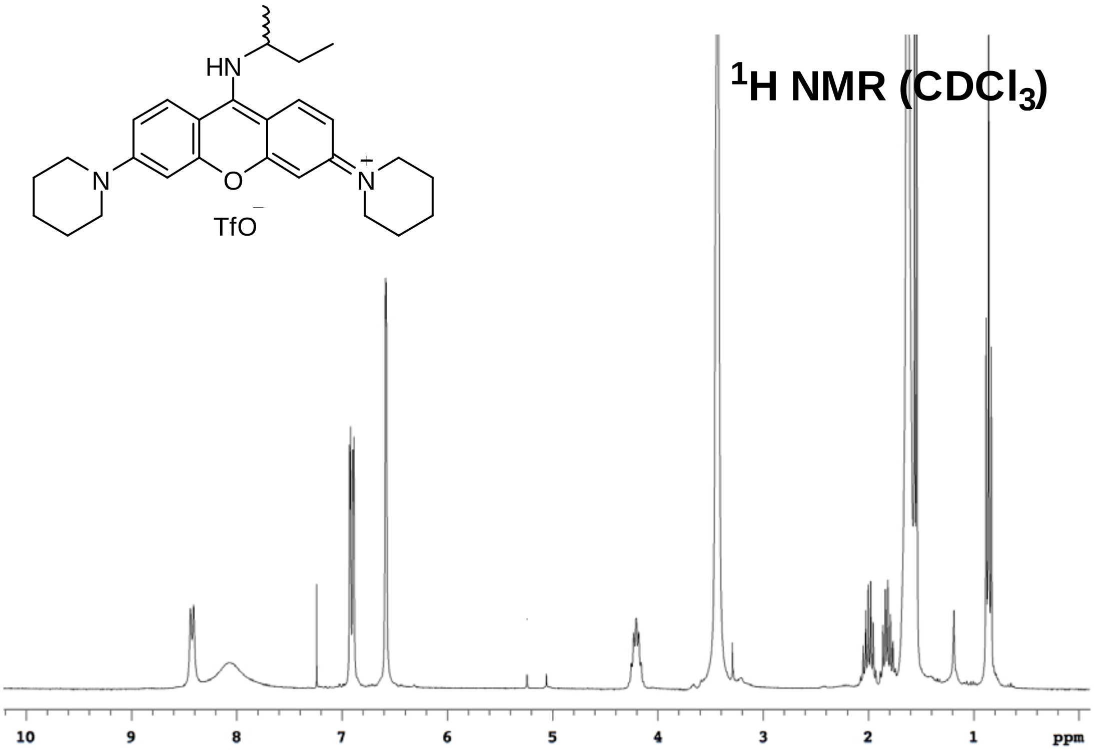


30c

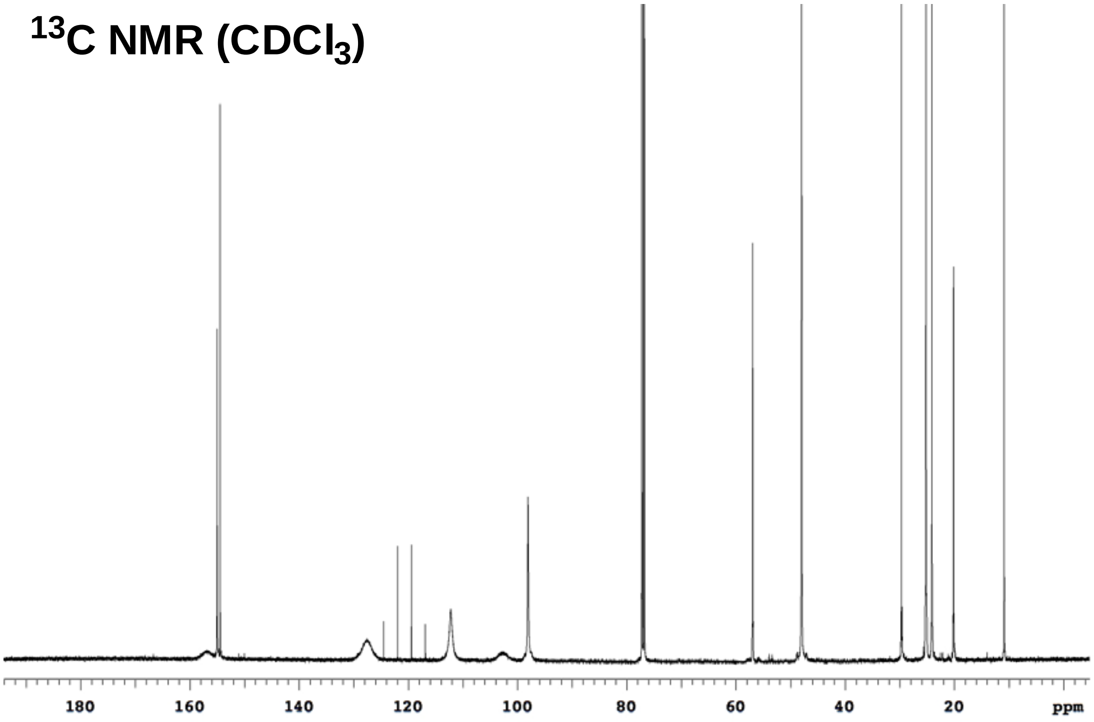
Yellow solid (112 mg, 98 %) [racemic *sec*-BuNH_2 was used as a nucleophile]. R_f = 0.36 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.42 (d, 1H, *J* = 9.0 Hz), 8.06 (br, 2H), 6.91 (dd, 2H, *J* = 9.6, 2.4 Hz), 6.58 (d, 2H, *J* = 2.4 Hz), 4.27-4.13 (m, 1H), 3.43 (br, 8H), 2.07-1.93 (m, 1H), 1.88-1.74 (m, 1H), 1.63 (br, 12H), 1.55 (d, 3H, *J* = 6.4 Hz), 0.86 (t, 3H, *J* = 7.4 Hz); ^{13}C NMR (125 MHz, $CDCl_3$) δ 157.0, 155.0, 154.5, 127.6, 120.6 (q, *J* = 320.3 Hz), 112.2, 102.6, 98.1, 56.9, 48.0, 29.7, 25.2, 24.1, 20.2, 10.9; ^{19}F NMR (282 MHz, $CDCl_3$) δ 99.1 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ $C_{27}H_{36}N_3O$ 418.29; found 418.28.

^{19}F NMR ($CDCl_3$)



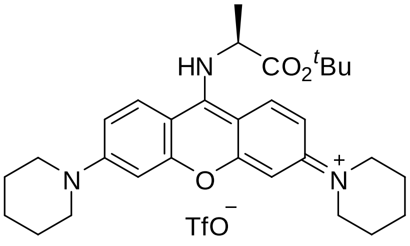


^{1}H NMR (CDCl$_3$)



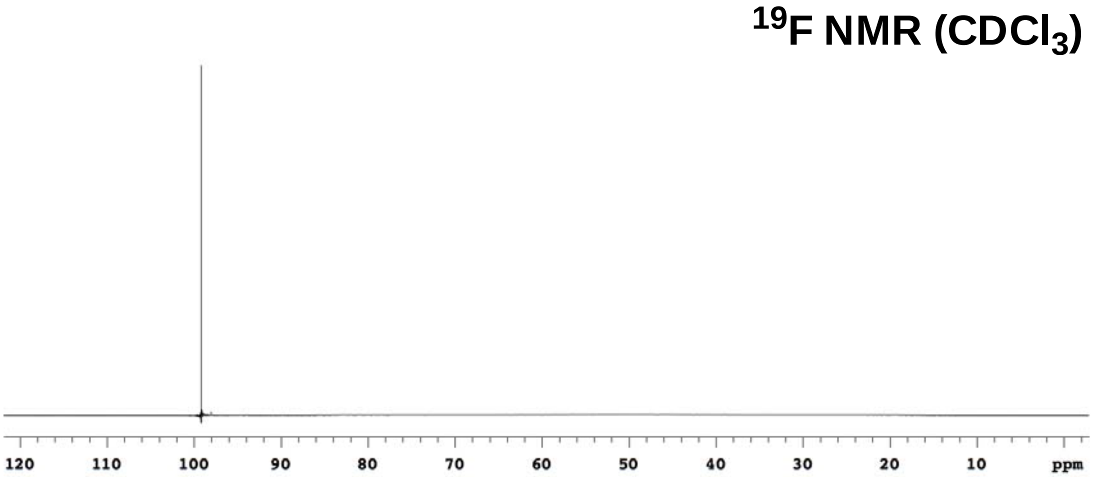
^{13}C NMR (CDCl$_3$)

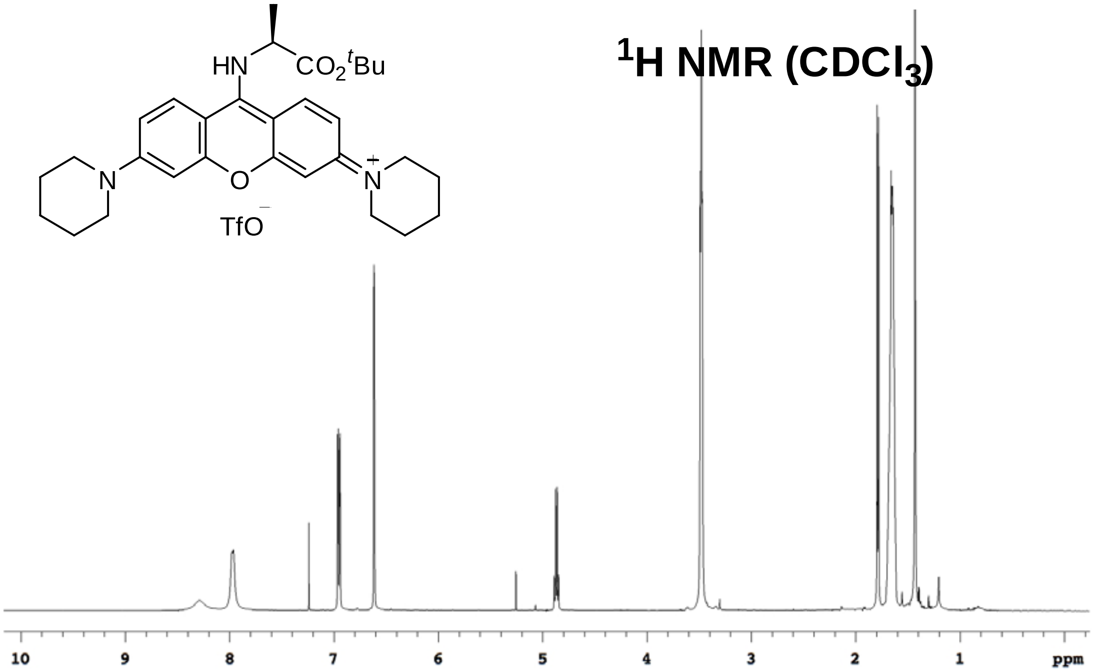
Pyronin 30d

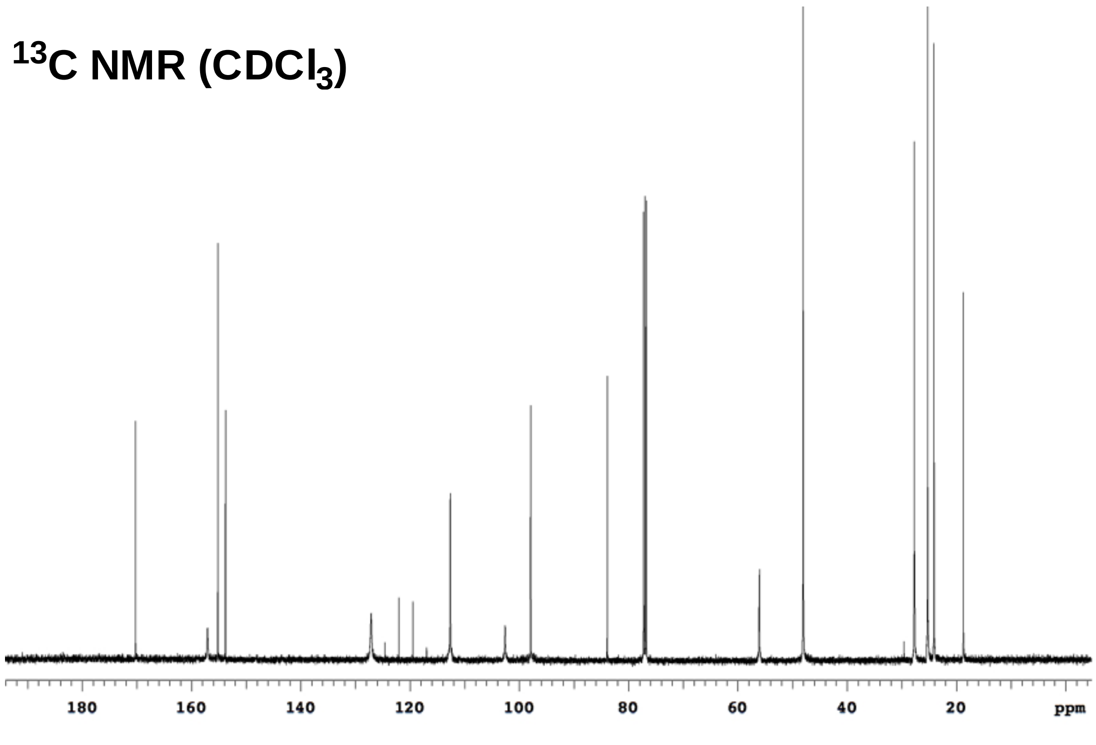


30d

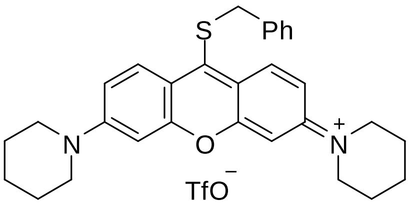
Yellow solid (101 mg, 80 %). R_f = 0.35 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.29 (br, 1H), 7.97 (br, 2H), 6.95 (dd, 2H, *J* = 9.5, 2.7 Hz), 6.61 (d, 2H, *J* = 2.7 Hz), 4.87 (q, 1H, *J* = 7.0 Hz), 3.49-3.47 (m, 8H), 1.79 (d, 3H, *J* = 7.0 Hz), 1.66-1.64 (m, 12H), 1.43 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.2, 157.1, 155.2, 153.8, 127.1, 120.7 (q, *J* = 320.4 Hz), 112.6, 102.6, 97.9, 83.9, 56.1, 48.0, 27.7, 25.2, 24.1, 18.7; ^{19}F NMR (282 MHz, $CDCl_3$) δ 99.2 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ $C_{30}H_{40}N_3O_3$ 490.31; found 490.30.




HN
CO2tBu
N
O
N+
TfO−
1H NMR (CDCl3)
10
9
8
7
6
5
4
3
2
1
ppm

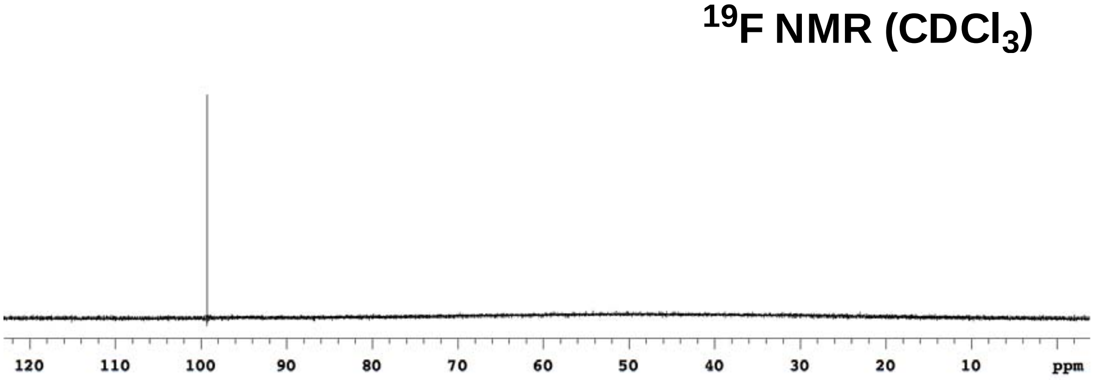

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm

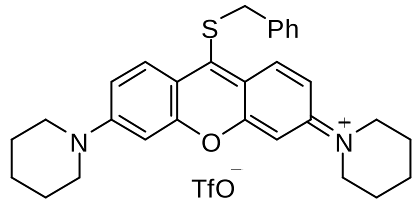
Pyronin 30e



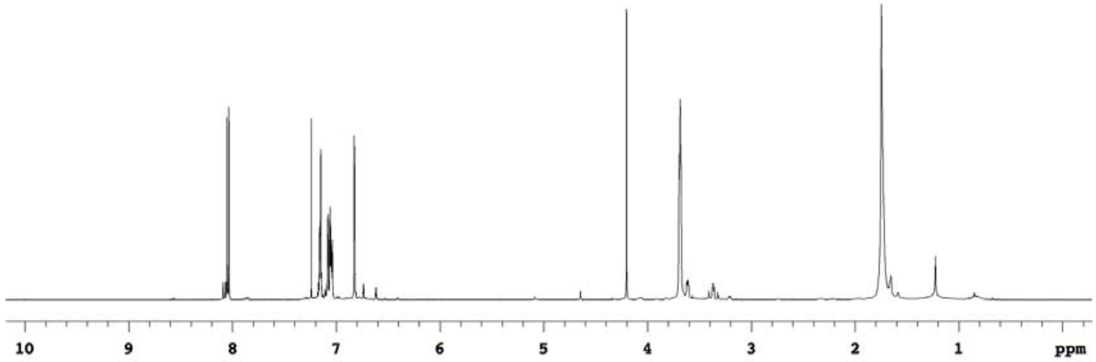
30e

Green solid (28 mg, 23 %). R_f = 0.29 (10 % MeOH/CH_2Cl_2). [1.5 eq Tf_2O and 3 eq BnSH were used in the reaction and stirred for 3h at 25°C]. ^1H NMR (500 MHz, $CDCl_3$) δ 8.04 (d, 2H, J = 9.6 Hz), 7.16-7.15 (m, 3H), 7.08-7.03 (m, 4H), 6.82 (d, 2H, J = 2.6 Hz), 4.20 (s, 2H), 3.70-3.68 (m, 8H), 1.75 (m, 12H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 157.1, 156.4, 155.6, 136.0, 131.4, 128.8, 128.7, 128.0, 120.9 (q, J = 320.3 Hz), 116.9, 114.8, 97.2, 49.1, 43.2, 25.9, 24.1; ^{19}F NMR (282 MHz, $CDCl_3$) δ 99.3 (s). MS (ESI) m/z calcd for $(M\text{-}OTf)^+$ $C_{30}H_{33}N_2OS$ 469.23; found 469.23.

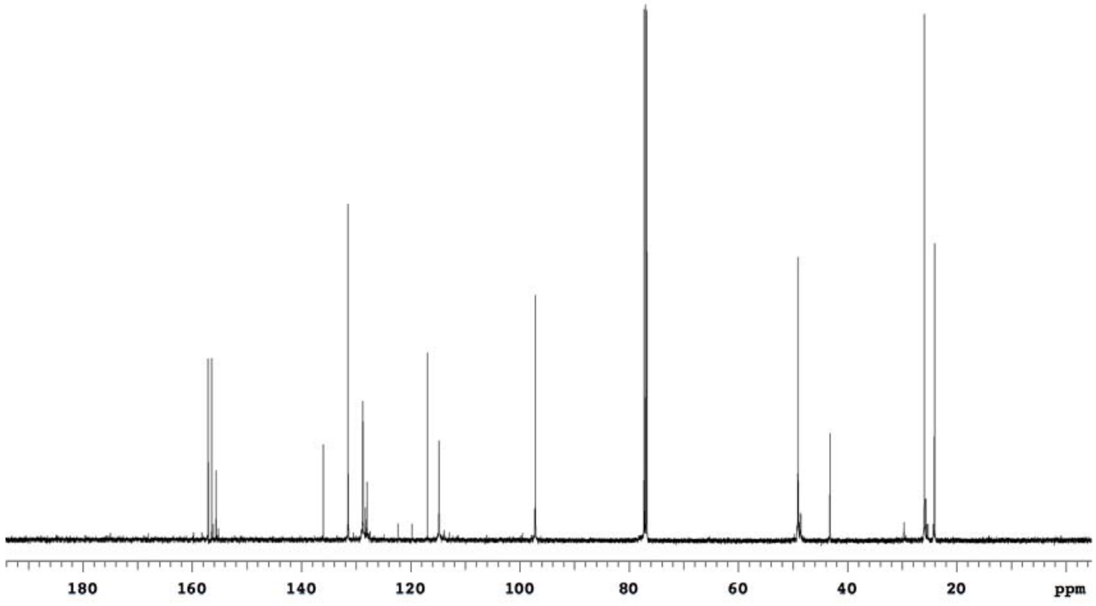




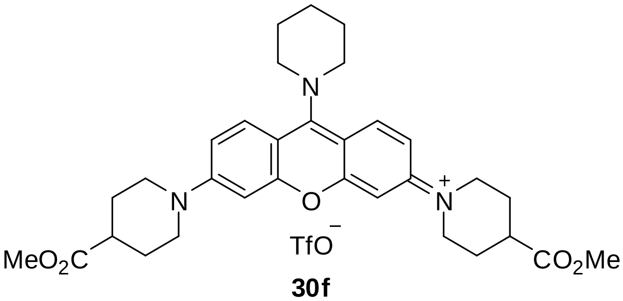
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)



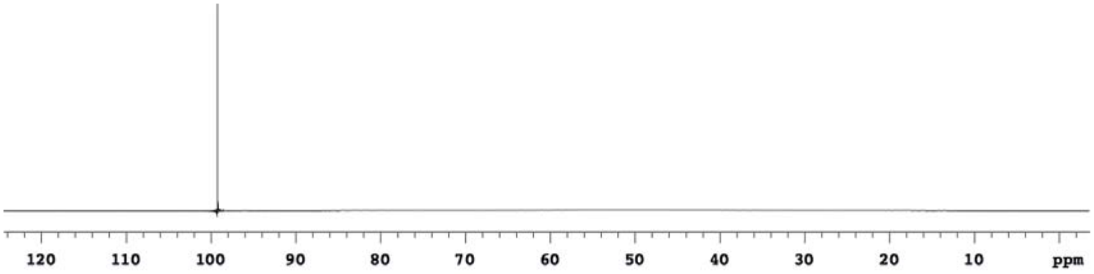
Pyronin 30f

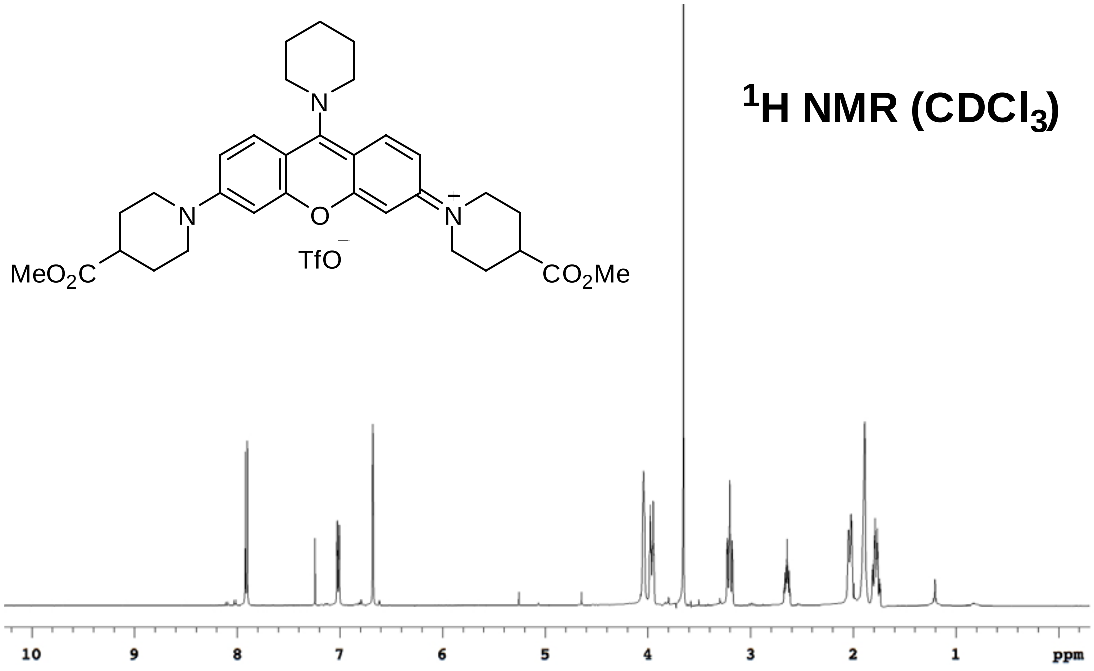


30f

Red solid (120 mg, 86 %). R_f = 0.36 (10 % MeOH/CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.91 (d, 2H, *J* = 9.6 Hz), 7.02 (dd, 2H, *J* = 9.6, 2.7 Hz), 6.68 (d, 2H, *J* = 2.7 Hz), 4.04 (br, 4H), 3.96 (dt, 4H, *J* = 13.7, 4.0 Hz), 3.65 (s, 6H), 3.23-3.17 (m, 4H), 2.67-2.61 (m, 2H), 2.03 (dd, 4H, *J* = 14.0, 3.6 Hz), 1.89 (br, 6H), 1.81-1.74 (m, 4H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 174.3, 160.6, 158.0, 154.8, 129.7, 120.8 (q, *J* = 320.6 Hz), 112.6, 107.4, 98.2, 57.0, 51.9, 46.3, 40.2, 27.3, 26.9, 23.5; ^{19}F NMR (282 MHz, $CDCl_3$) δ 99.2 (s). MS (ESI) m/z calcd for $(M\text{-}OTf)^+$ $C_{32}H_{40}N_3O_5$ 546.30; found 546.31.

^{19}F NMR ($CDCl_3$)





^{1}H NMR ($CDCl_3$)

^{13}C NMR ($CDCl_3$)

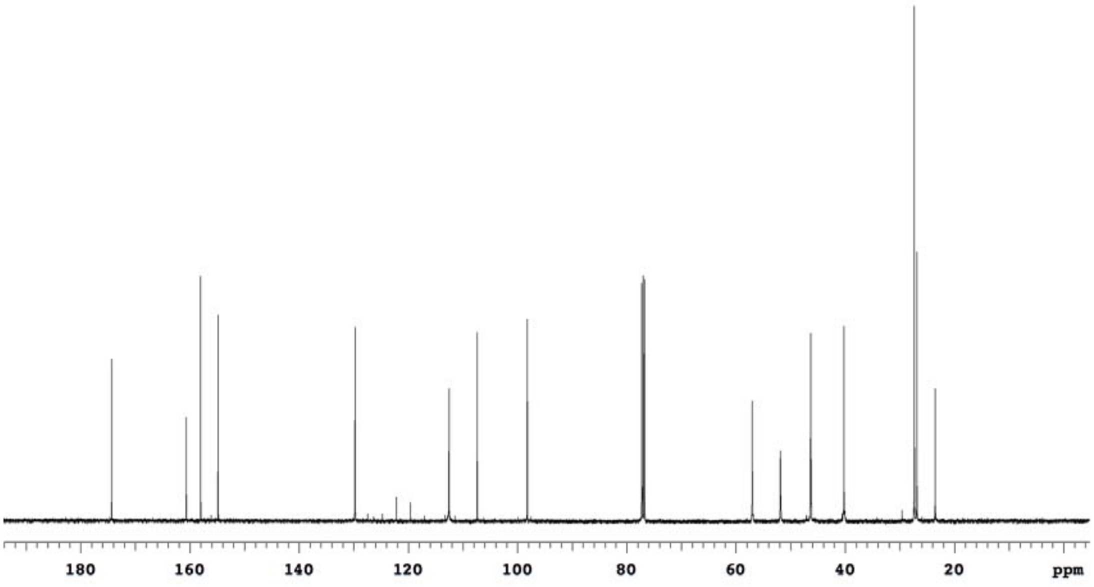
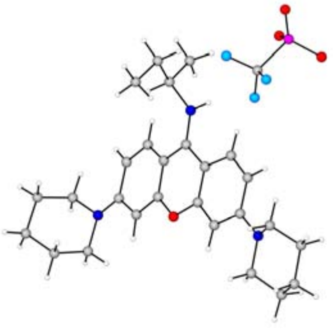
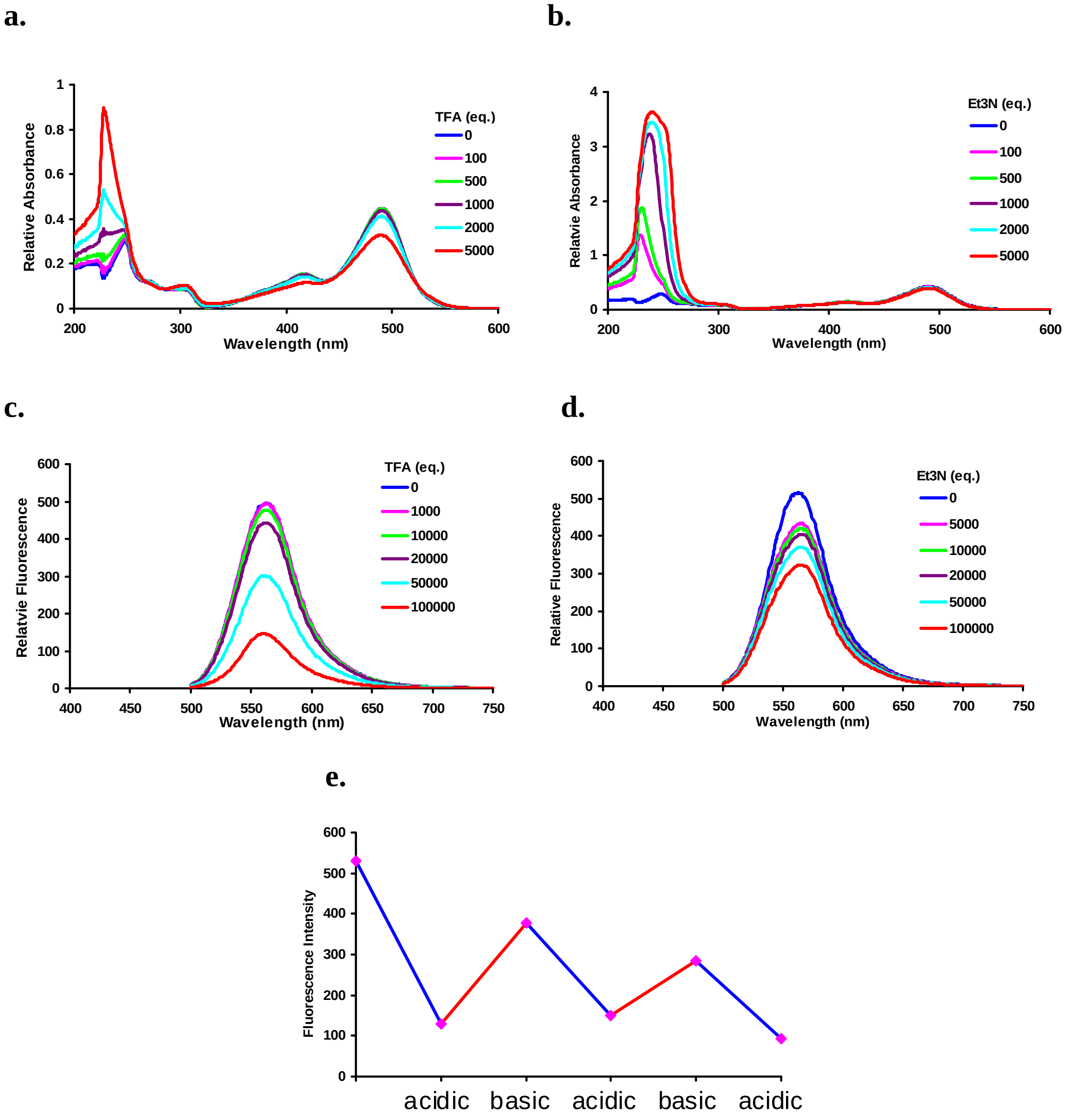


Figure E1. X-Ray Structure of Pyronin 30c



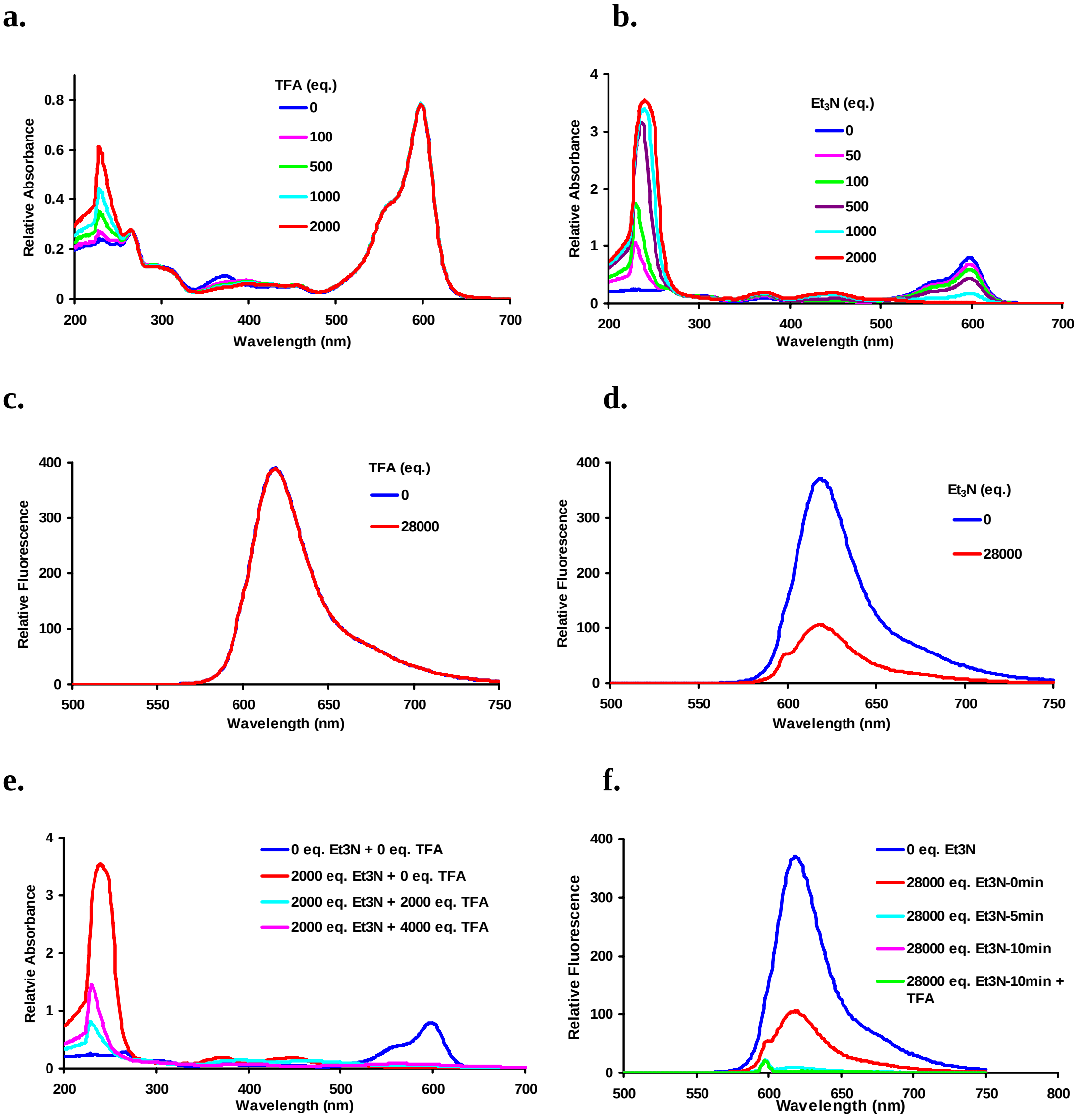
Empirical formula	C28 H36 F3 N3 O4 S	
Formula weight	567.66	
Temperature	110(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2(1)/n	
Unit cell dimensions	a = 12.2520(19) Å	α= 90°.
	b = 14.376(2) Å	β= 106.676(2)°.
	c = 16.502(3) Å	γ = 90°.
Volume	2784.4(7) Å^3	
Z	4	
Density (calculated)	1.354 Mg/m^3	
Absorption coefficient	0.175 mm^{-1}	
F(000)	1200	
Crystal size	0.50 x 0.43 x 0.40 mm^3	
Theta range for data collection	1.84 to 25.00°.	
Index ranges	-14<=h<=14, -17<=k<=17, -19<=l<=19	
Reflections collected	25786	
Independent reflections	4882 [R(int) = 0.0456]	
Completeness to theta = 25.00°	99.6 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.9332 and 0.9175	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	4882 / 0 / 354	
Goodness-of-fit on F^2	1.101	
Final R indices [I>2sigma(I)]	R1 = 0.0864, wR2 = 0.2285	
R indices (all data)	R1 = 0.1013, wR2 = 0.2543	
Largest diff. peak and hole	0.893 and -0.336 e.Å^{-3}	

Figure E2. pH Dependence of Photophysical Properties for 30a



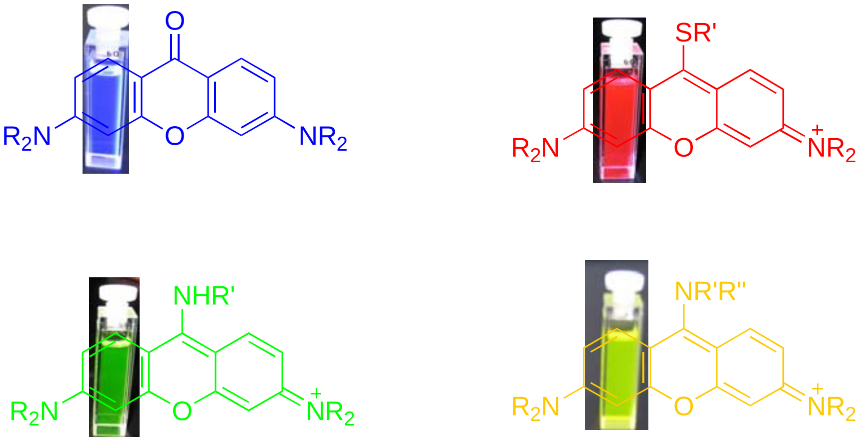
pH dependence of UV absorbance of **30a** under: **a** acidic; **b** basic conditions (CH_2Cl_2, 7.0 x 10^{-6} M). Fluorescence pH dependence of **30a** under: **c** acidic; **d** basic conditions (CH_2Cl_2, 7.0 x 10^{-7} M). **e** Fluorescence intensity of **30a** under alternating acidic (TFA) and basic (Et_3N) conditions (CH_2Cl_2, 7.0 x 10^{-7} M).

Figure E3. pH Dependence of Photophysical Properties for 30e



pH dependence of UV absorbance of **30e** under: **a** acidic; **b** basic conditions (CH_2Cl_2, 9.6 x 10^{-6} M). Fluorescence pH dependence of **30e** under: **c** acidic; **d** basic conditions (CH_2Cl_2, 1.4 x 10^{-7} M). **e** UV absorbance of **30e** under alternating acidic and basic conditions (CH_2Cl_2, 9.6 x 10^{-6} M) [absorbance decreased at λ_{max} with addition of Et_3N and won't recover after addition of TFA]. **f** Fluorescence of **30e** under alternating acidic and basic conditions (CH_2Cl_2, 1.4 x 10^{-7} M). [Completely lost fluorescence within 10 min after addition of Et_3N and did not recover after addition of TFA]

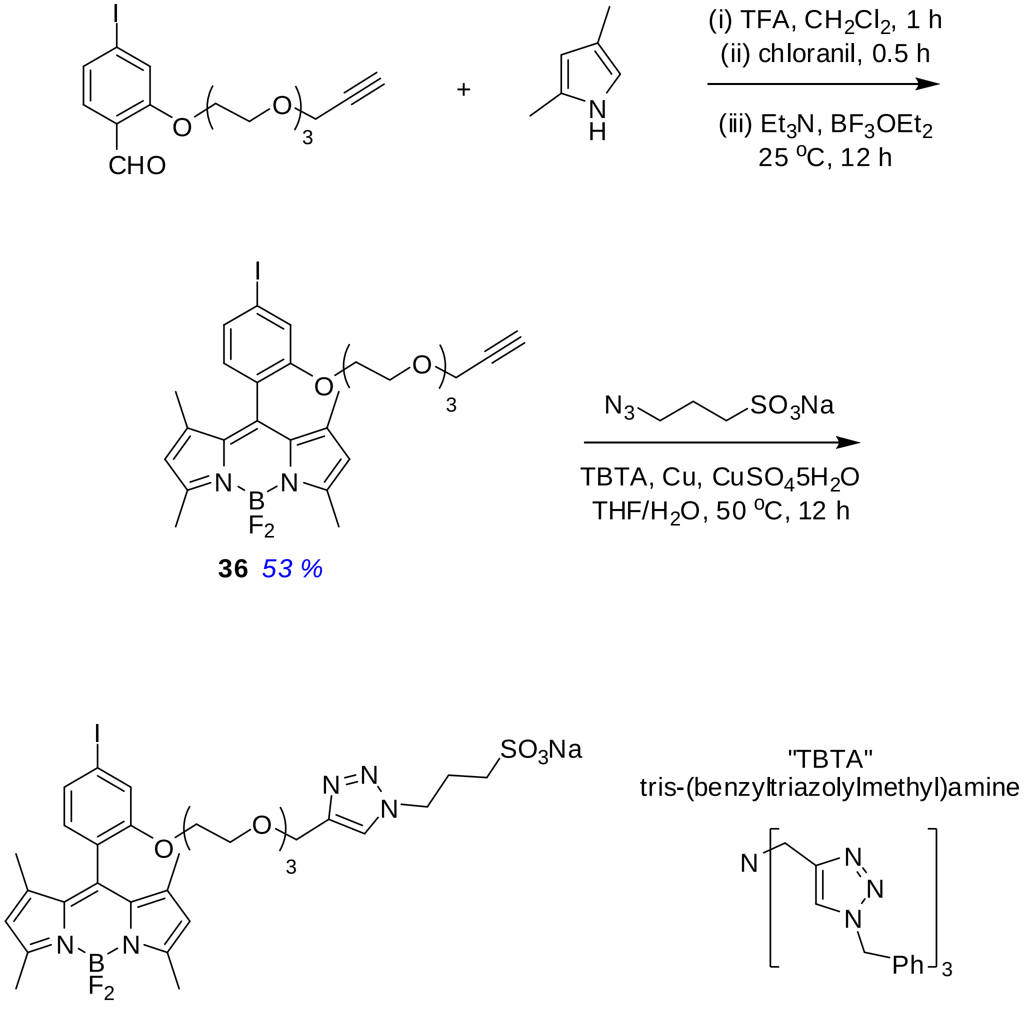
Figure E4. Photograph of 16 and 30 in CH_2Cl_2 Under UV Irradiation


O
R2N
O
NR2
SR'
R2N
O
+NR2
NHR'
R2N
O
+NR2
NR'R"
R2N
O
+NR2

APPENDIX F

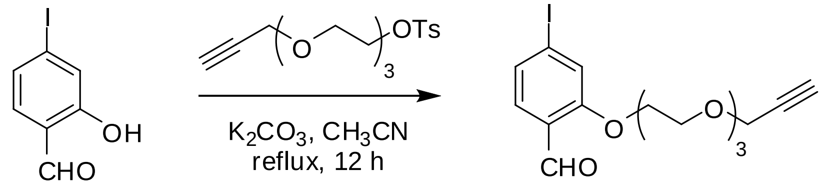
EXPERIMENTAL DATA FOR CHAPTER VI

Synthesis of BODIPY donor 35

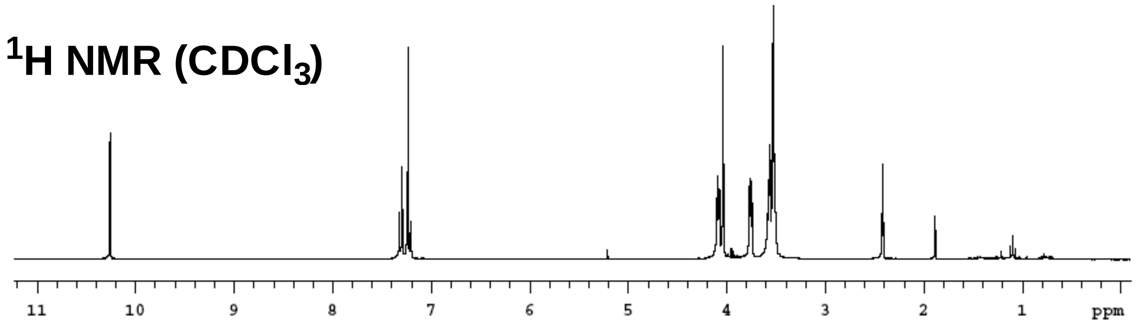

I
CHO
O
O
3
+
N
H
(i) TFA, CH$_2$Cl$_2$, 1 h
(ii) chloranil, 0.5 h
(iii) Et$_3$N, BF$_3$OEt$_2$
25 ^{o}C, 12 h
N$_3$
SO$_3$Na
TBTA, Cu, CuSO$_4$5H$_2$O
THF/H$_2$O, 50 ^{o}C, 12 h
N
B
F$_2$
N=N
"TBTA"
tris-(benzyltriazolylmethyl)amine
Ph
3

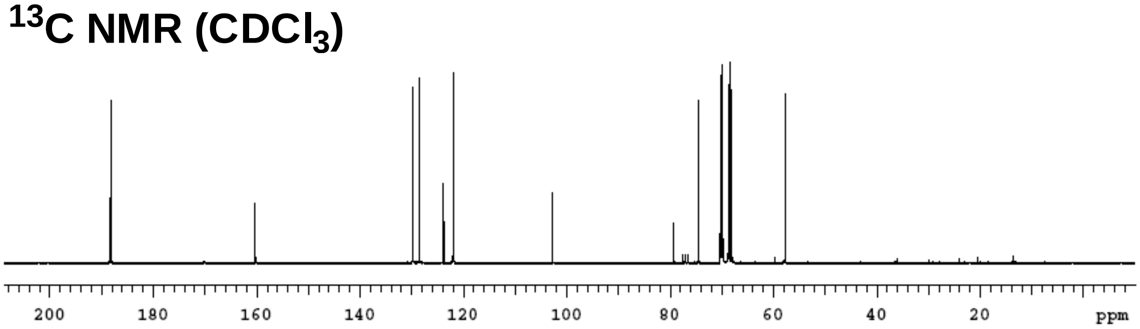
36 *53 %*

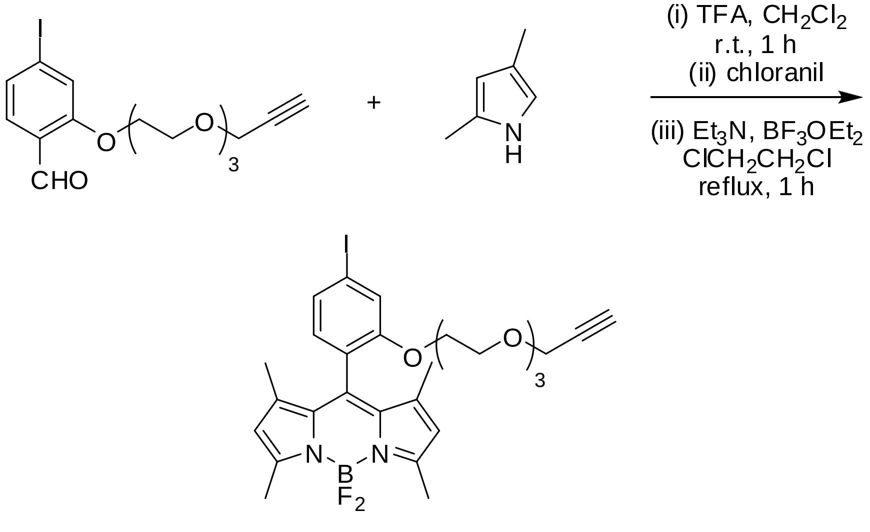
35 *95 %*



A mixture of 2-hydroxy-4-iodobenzaldehyde (501 mg, 2.02 mmol) and K_2CO_3 (332 mg, 2.4 mmol) in 20 mL CH_3CN was stirred at 25 °C for 10 min then a solution of $CH{\equiv}CCH_2(OCH_2CH_2)_3OTs$ (685 mg, 2.0 mmol) in 20 mL CH_3CN was added. The reaction mixture was refluxed for 12 h then cooled to 25 °C and diluted with water. The mixture was extracted with EtOAc three times. The organic layer was washed with brine, dried over Na_2SO_4 and evaporated to dryness under reduced pressure. The residue was purified by flash chromatography (30 % EtOAc/hexanes) to afford the pure product (634 mg, 76 %) as light yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.26 (s, 1H), 7.32-7.30 (m, 1H), 7.24-7.21 (m, 2H), 4.10-4.07 (m, 2H), 4.04 (d, 2H, *J* = 2.4 Hz), 3.77-3.74 (m, 2H), 3.59-3.50 (m, 8H), 2.42 (t, 1H, *J* = 2.4 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 188.2, 160.2, 129.7, 128.5, 123.8, 121.9, 102.8, 79.2, 74.4, 70.2, 70.0, 69.8, 68.7, 68.4, 68.1, 57.7.

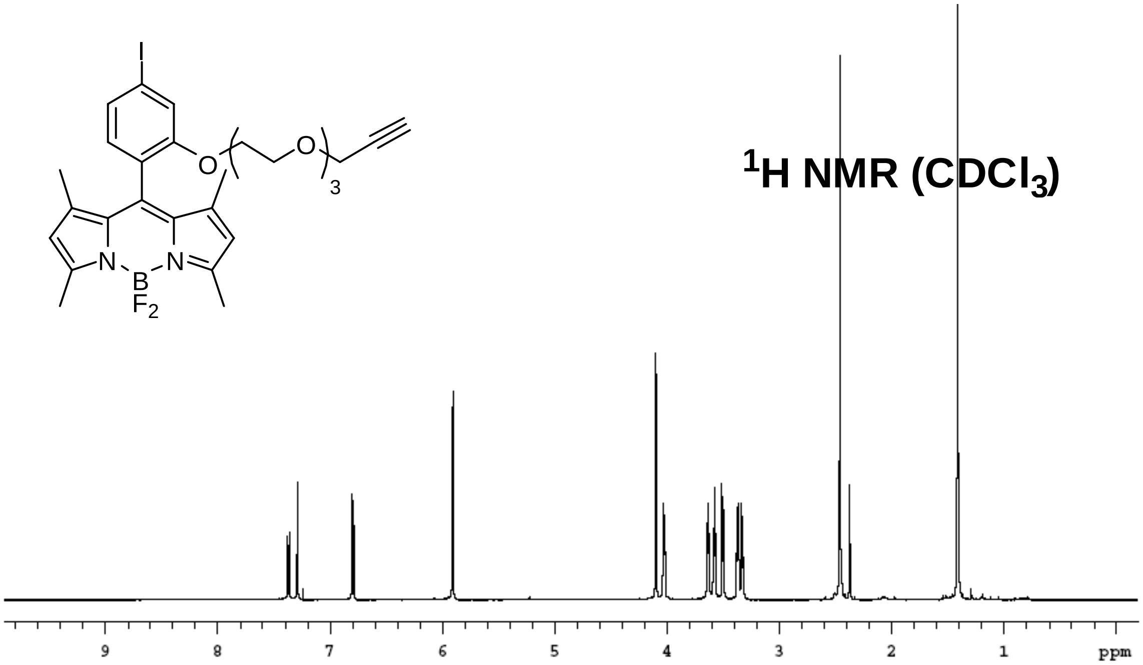




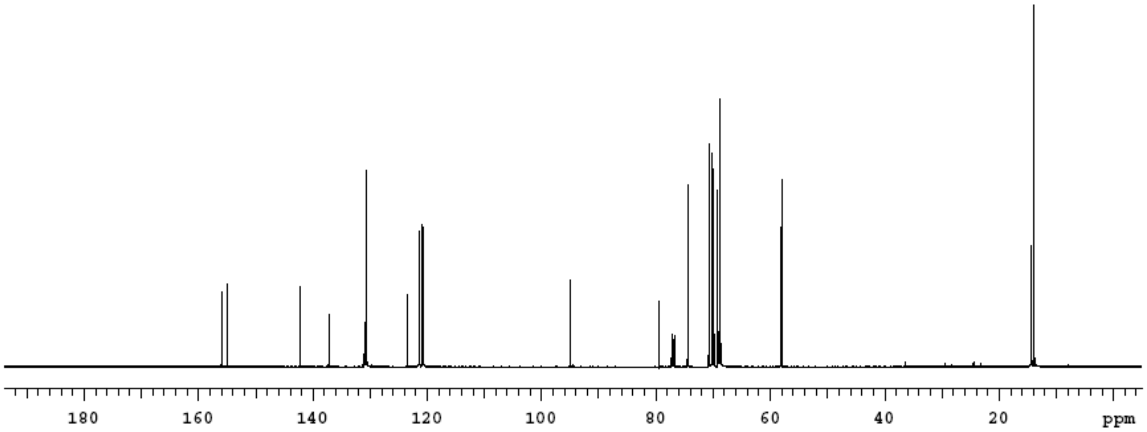


36

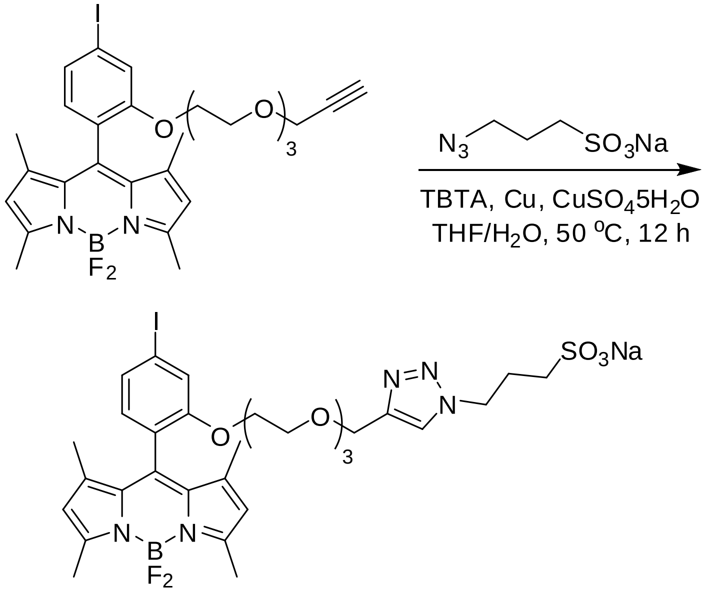
To a solution of the benzaldehyde derivative (634 mg, 1.52 mmol) and 2,4-dimethylpyrrole (0.33 mL, 3.2 mmol) in 10 mL CH_2Cl_2 was added one drop of TFA. The solution was stirred at 25 °C for 4 h. Chloranil (411 mg, 1.67 mmol) was added and stirring was continued for 1 h. The mixture was passed through a short pad of basic alumina eluting with 40 % EtOAc/hexanes. The solvents were removed under reduced pressure. The residue was dissolved in 10 mL CH_2Cl_2 and Et_3N (0.63 mL, 4.56 mmol) was added followed by BF_3OEt_2 (0.96 mL, 7.6 mmol) after 10 min. The solution was stirred at 25 °C for 12 h then washed with water and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (30 % EtOAc/hexanes) to afford the pure product **36** (508 mg, 53 %) as an orange sticky solid. ^1H NMR (300 MHz, $CDCl_3$) δ 7.37 (dd, 1H, *J* = 7.8, 1.4 Hz), 7.29 (d, 1H, *J* = 1.4 Hz), 6.80 (d, 1H, *J* = 7.8 Hz), 5.91 (s, 2H), 4.10 (d, 2H, *J* = 2.4 Hz), 4.04-4.02 (m, 2H), 3.64-3.62 (m, 2H), 3.59-3.57 (m, 2H), 3.51-3.49 (m, 2H), 3.38-3.36 (m, 2H), 3.34-3.32 (m, 2H), 2.46 (s, 6H), 2.37 (t, 1H, *J* = 2.4 Hz), 1.41 (s, 6H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 156.0, 154.9, 142.2, 137.2, 130.9, 130.6 (2 peaks: 130.62, 130.60), 123.4, 121.4, 120.8, 95.0, 79.4, 74.4, 70.7, 70.3, 69.9, 69.1, 68.8, 68.7, 58.0,14.3, 13.9; MS (MALDI) m/z calcd for $(M+Na)^+$ $C_{28}H_{32}BF_2INaN_2O_4$ 659.14; found 659.18.



^{1}H NMR (CDCl$_3$)

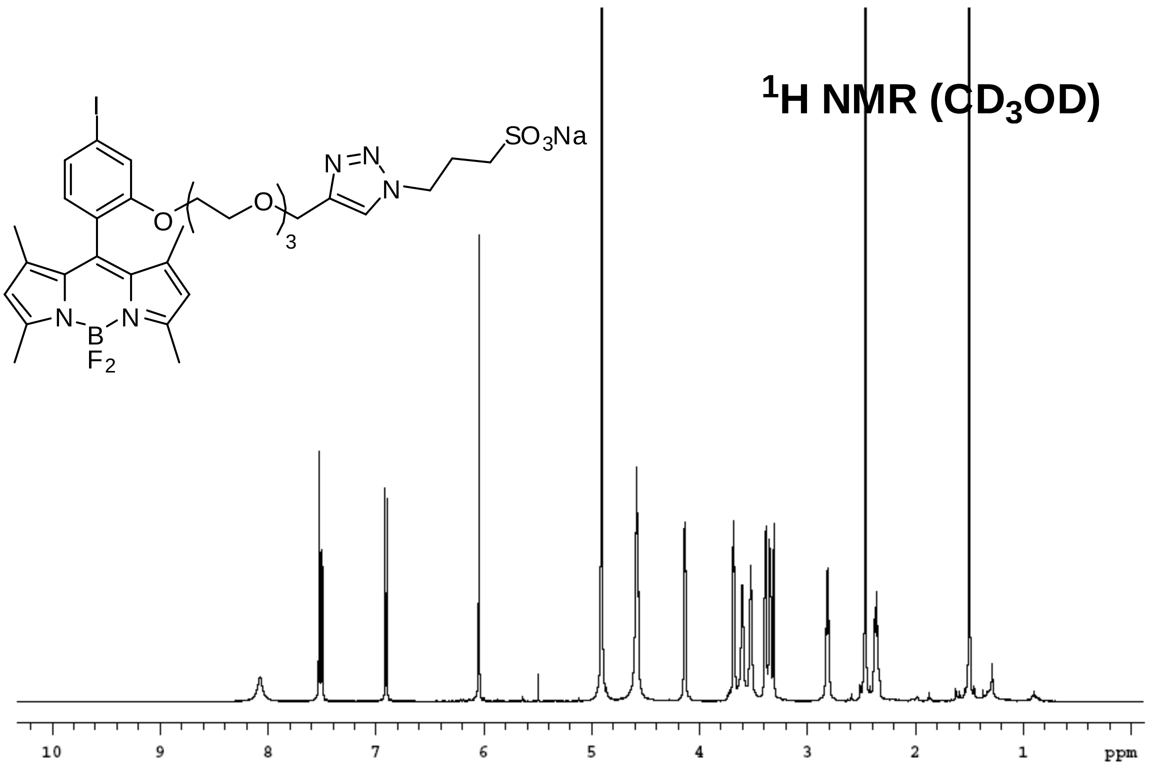


^{13}C NMR (CDCl$_3$)



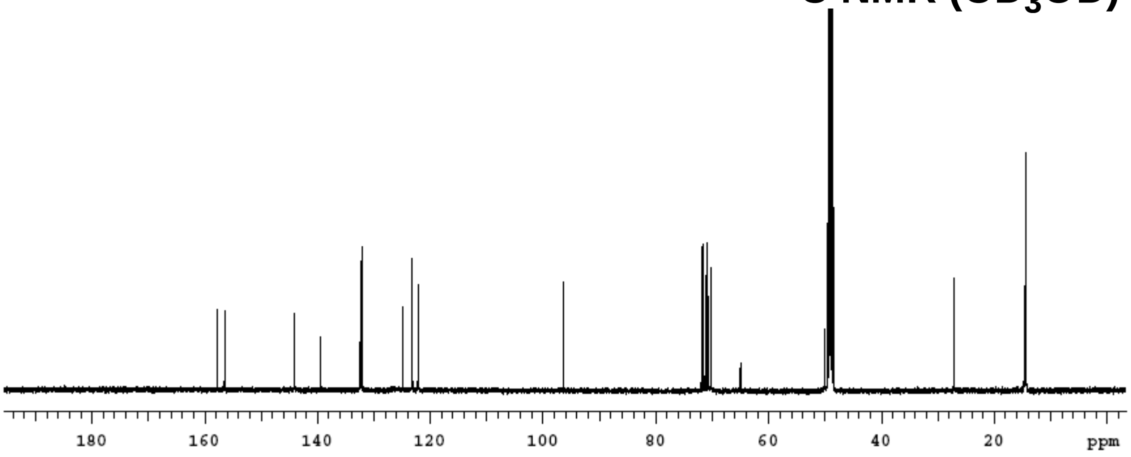
35

BODIPY **36** (77 mg, 0.12 mmol), 3-azido-1-propanesulfonic acid, sodium salt (28 mg, 0.15 mmol), Cu (8 mg, 0.12 mmol), $CuSO_45H_2O$ (5 mg, 0.02 mmol) and TBTA ligand (3 mg, 0.006 mmol) were dissolved in 10 mL THF and 5 ml water. The mixture was stirred at 50 °C for 12 h then evaporated to dryness under reduced pressure. The residue was purified by flash chromatography (20 % MeOH/CH_2Cl_2) to afford the product **35** (95 mg, 95 %) as an orange semi-solid. ^{1}H NMR (500 MHz, CD_3OD) δ 8.07 (s, 1H), 7.53-7.52 (m, 1H), 7.51-7.49 (m, 1H), 6.90 (d, 1H, J = 7.8 Hz), 6.04 (s, 2H), 4.59-4.56 (m, 4H), 4.14-4.13 (m, 2H), 3.69-3.67 (m, 2H), 3.61-3.59 (m, 2H), 3.53-3.51 (m, 2H), 3.40-3.38 (m, 2H), 3.35-3.34 (m, 2H), 2.83-2.80 (m, 2H), 2.46 (s, 6H), 2.39-2.33 (m, 2H), 1.50 (s, 6H); ^{13}C NMR (125 MHz, CD_3OD) δ 157.7, 156.4, 144.1, 139.4, 132.5, 132.2, 132.1, 124.8, 123.2, 122.1, 96.4, 71.9, 71.5, 71.2, 70.8, 70.7, 70.2, 65.0, 50.1, 27.2, 14.6, 14.4, several peaks account for more than one carbon; ^{19}F NMR (282 MHz, CD_3OD) δ 31.81-31.02 (m); MS (ESI) m/z calcd for $(M\text{-}Na)^-$ $C_{31}H_{38}BF_2IN_5O_7S$ 800.16; found 800.16.

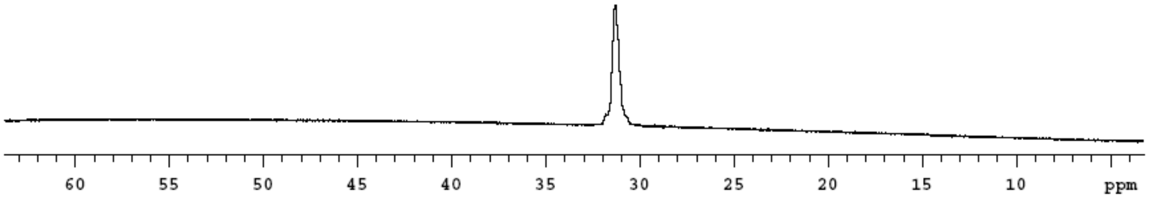


^{1}H NMR (CD$_3$OD)

^{13}C NMR (CD$_3$OD)

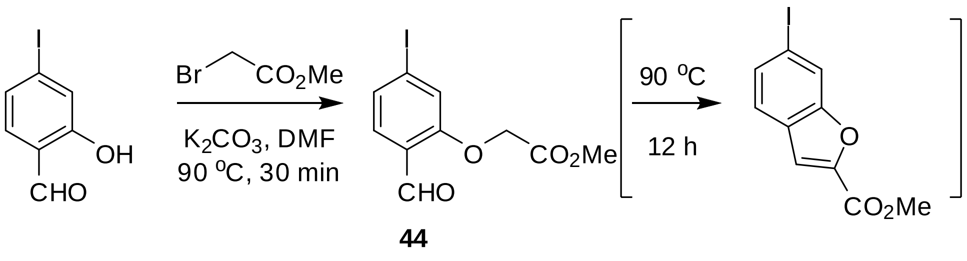


^{19}F NMR (CD$_3$OD)



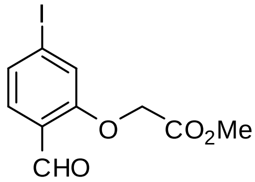
Synthesis of BODIPY acceptors

(1) Syntheses of benzaldehyde derivatives

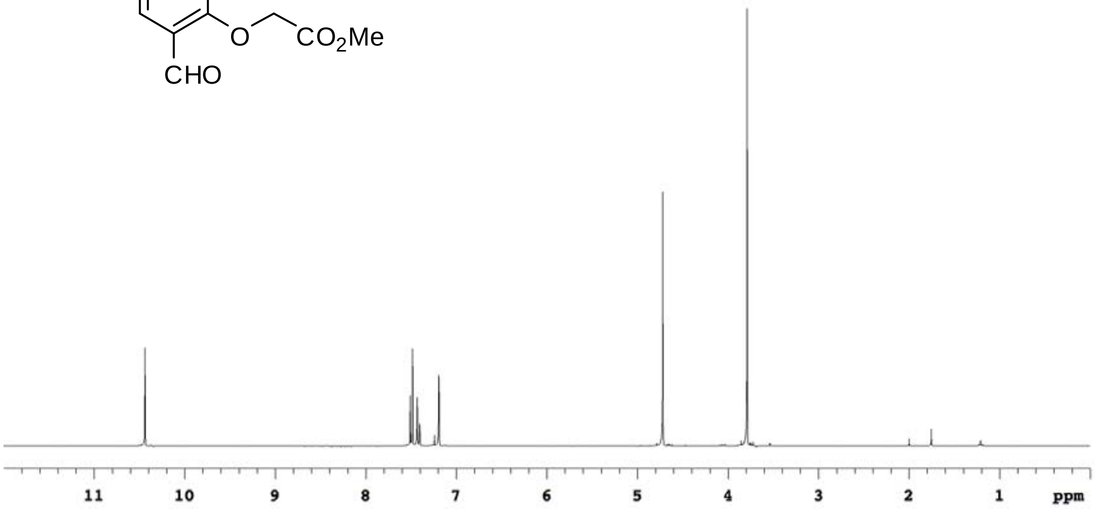


2-Hydroxy-4-iodobenzaldehyde (1.24 g, 5.0 mmol) and K_2CO_3 (1.38 g, 10.0 mmol) were dissolved in 25 mL DMF and stirred for 10 min then methyl bromoacetate (0.7 mL, 7.5 mmol) was added. The mixture was stirred at 90 °C for 30 min then cooled to room temperature and diluted with EtOAc. The mixture was washed with H_2O and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (10 to 20 % EtOAc/Hexanes) to afford the product **44** as white solid (1.41 g, 88 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.44 (s, 1H), 7.49 (d, 1H, *J* = 8.1 Hz), 7.43-7.40 (m, 1H), 7.19 (d, 1H, *J* = 1.4 Hz), 4.72 (s, 2H), 3.79 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 188.8, 168.0, 159.6, 131.4, 129.4, 124.8, 122.1, 102.8, 65.5, 52.5.

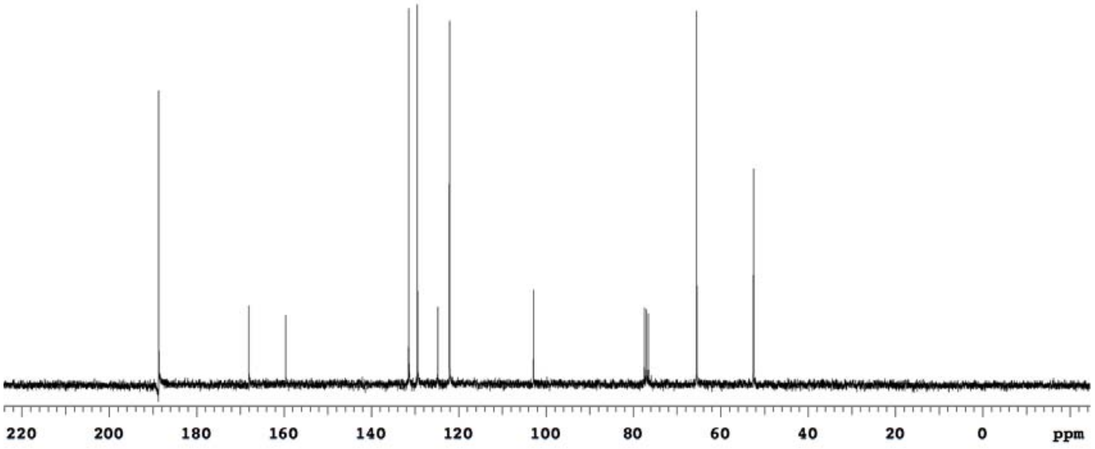
If the mixture was stirred at 90 °C for 12 h, the reaction gave a different product which was found to be a benzofuran derivative in 62 % yield. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.91 (s, 1H), 7.57 (dd, 1H, *J* = 8.2, 1.4 Hz), 7.43 (d, 1H, *J* = 1.4 Hz), 7.38 (d, 1H, *J* = 8.2 Hz), 3.94 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 159.5, 155.8, 145.5, 133.0, 126.4, 124.0, 121.6, 113.6, 91.9, 52.5.

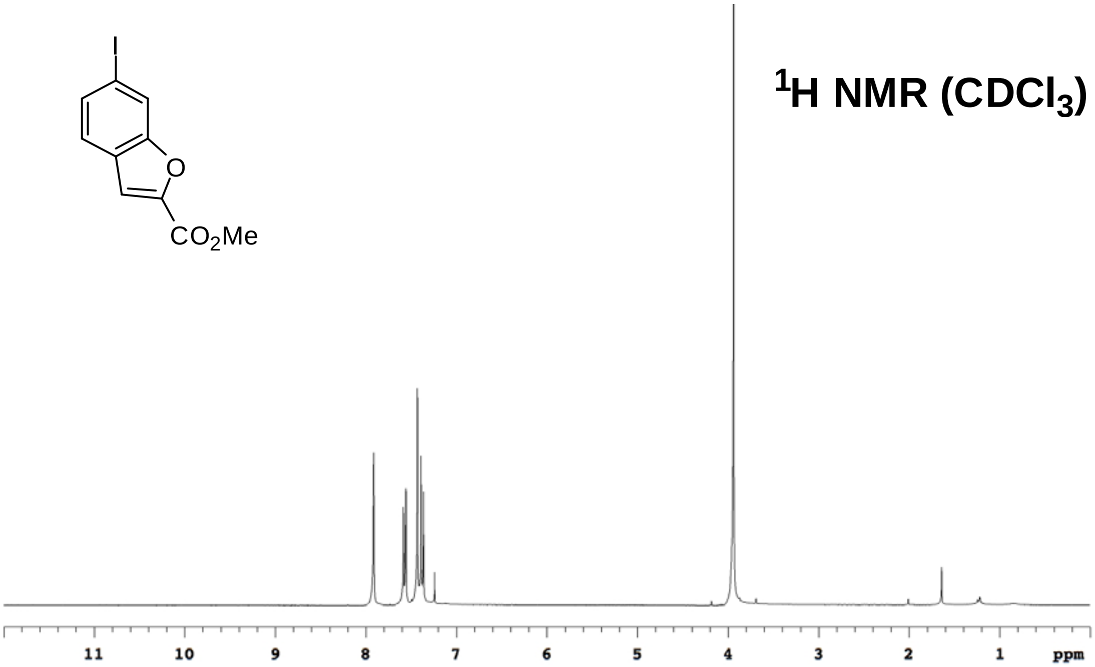

I
O
CO2Me
CHO

^{1}H NMR (CDCl$_3$)



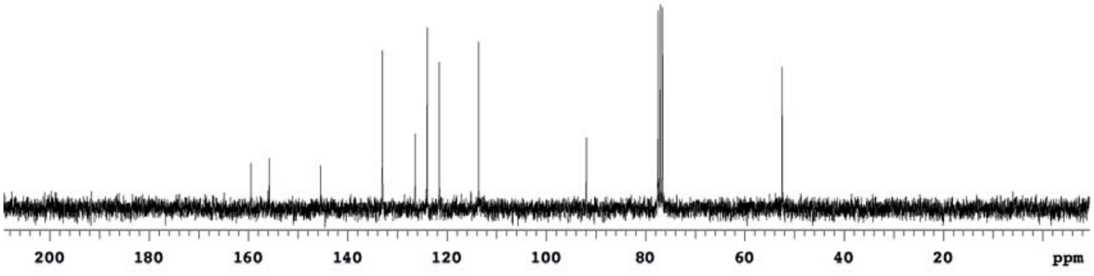
^{13}C NMR (CDCl$_3$)

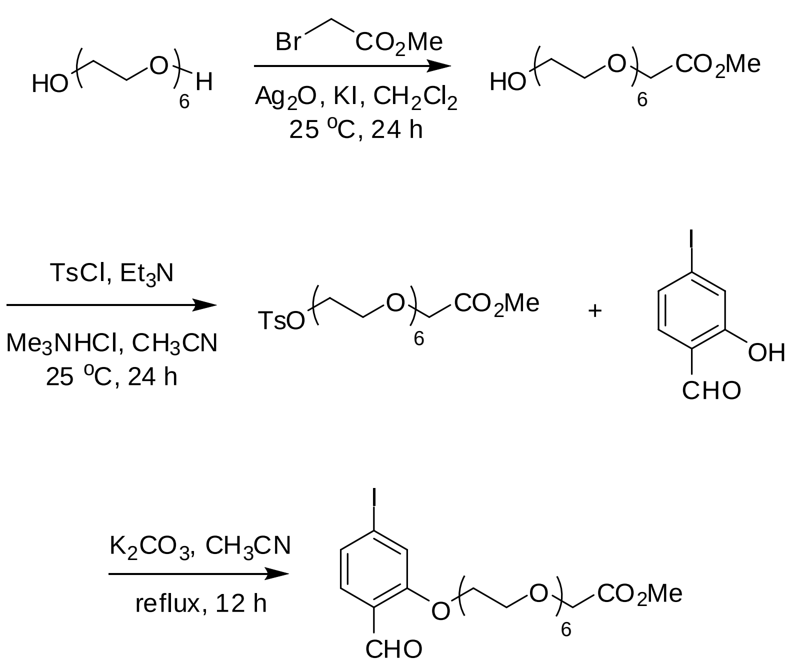




^{1}H NMR (CDCl$_3$)

^{13}C NMR (CDCl$_3$)



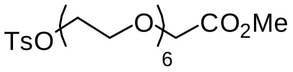


45

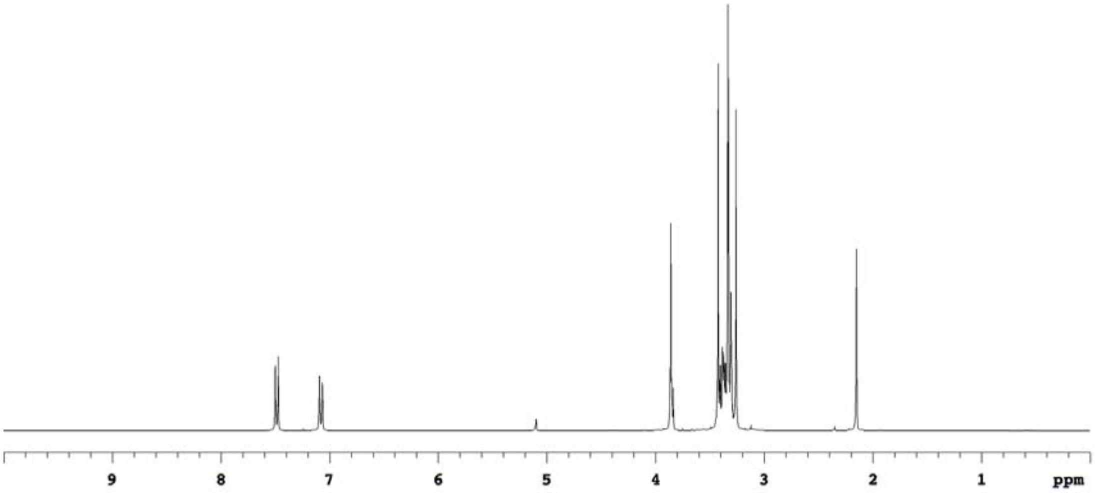
(i) Hexaethylene glycol (8.9 g, 30 mmol) was dissolved in 60 mL dichloromethane and Ag_2O (7.7 g, 33 mmol) was added in three portions over 10 min. The mixture was cooled to 0 °C and stirred for 10min. Bromo-methyl acetate (3.3 mL, 36 mmol) was added dropwise to the mixture over 5 min followed by KI (3.0 g, 18 mmol) in one portion. The mixture was stirred at 0 °C for 1 h then warmed to room temperature. After stirred for 24 h, the mixture was filtered through celite and concentrated under reduced pressure. The residue was purified by flash chromatograph (2 to 5 % $MeOH/CH_2Cl_2$) to afford the product as yellow oil (5.2 g, 49 %) which was used in the next step without further purification.

(ii) A solution of $HO(CH_2CH_2O)_6CH_2CO_2Me$ (2.4 g, 6.8 mmol) and Me_3NHCl (67 mg, 0.7 mmol) in 30 mL acetonitrile was cooled to 0 °C and Et_3N (1.9 mL, 13.6 mmol) was added then a solution of TsCl (1.7 g, 8.8 mmol) in 20 mL acetonitrile was added dropwise over 10 min. The mixture was stirred at 0 °C for 1 h then room temperature for 24 h. The mixture was diluted with EtOAc, washed with water and brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatograph (2 to 5 % MeOH/CH_2Cl_2) to afford the pure product as yellow oil (2.5 g, 73 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.49 (d, 2H, *J* = 8.2 Hz), 7.08 (d, 2H, *J* = 8.2 Hz), 3.87-3.84 (m, 4H), 3.42-3.26 (m, 25 H), 2.15 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.9, 144.0, 132.1, 129.1, 127.0, 69.9, 69.7, 69.6 (2C: 69.64, 69.61), 69.5, 68.6, 67.7, 67.5, 50.8, 20.7, several peaks account for more than one carbon.

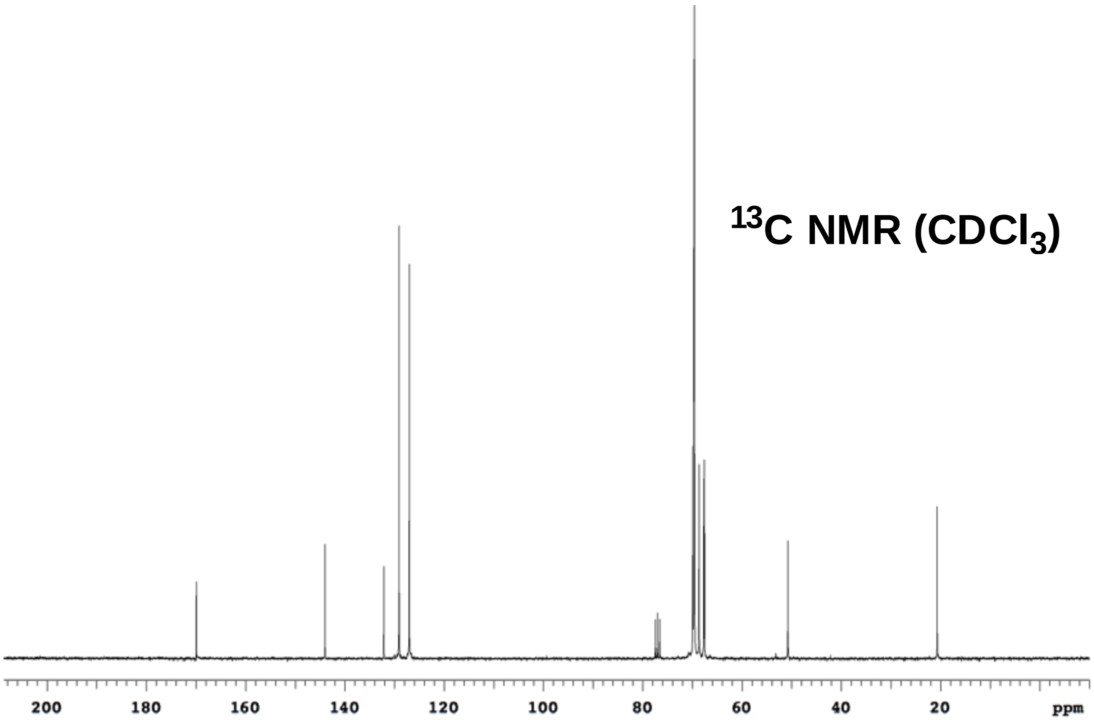
(iii) A solution of 2-hydroxy-4-iodobenzaldehyde (1.1 g, 4.3 mmol) and K_2CO_3 (774 mg, 5.6 mmol) in 30 mL CH_3CN was stirred at room temperature for 10 min then a solution of $TsO(CH_2CH_2O)_6CH_2CO_2Me$ in 20 mL CH_3CN was added dropwise over 10 min. After refluxing for 12 h, the mixture was diluted with EtOAc, washed with water and brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatograph (2 to 5 % MeOH/CH_2Cl_2) to afford the pure product as yellow oil (2.4 g, 92 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.38 (s, 1H), 7.45 (d, 1H, *J* = 8.6 Hz), 7.37-7.34 (m, 2H), 4.19 (t, 2H, *J* = 4.6 Hz), 4.11 (s, 2H), 3.86-3.83 (m, 2H), 3.69-3.59 (m, 23 H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 189.0, 170.8, 160.8, 130.4, 129.1, 124.3, 122.4, 103.1, 70.9, 70.8, 70.5, 70.4, 69.3, 68.5, 51.7, several peaks account for more than one carbon. MS (ESI) m/z calcd for $(M+Li)^+$ $C_{22}H_{33}ILiO_{10}$ 591.13; found 591.13.

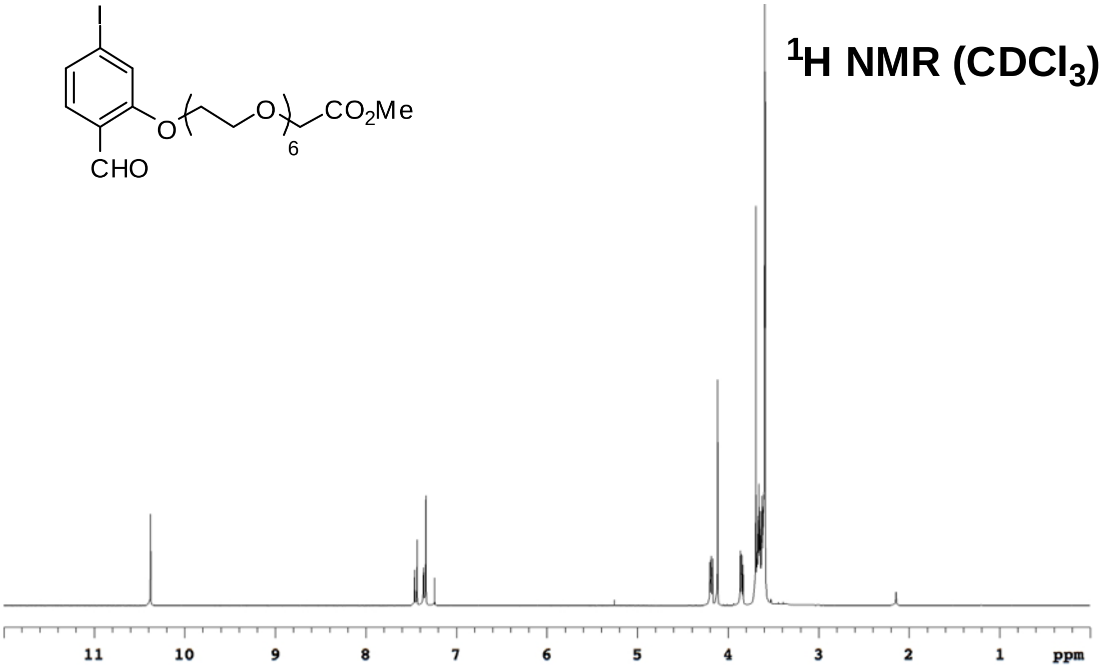

TsO
O
6
CO2Me

^{1}H NMR ($CDCl_3$)

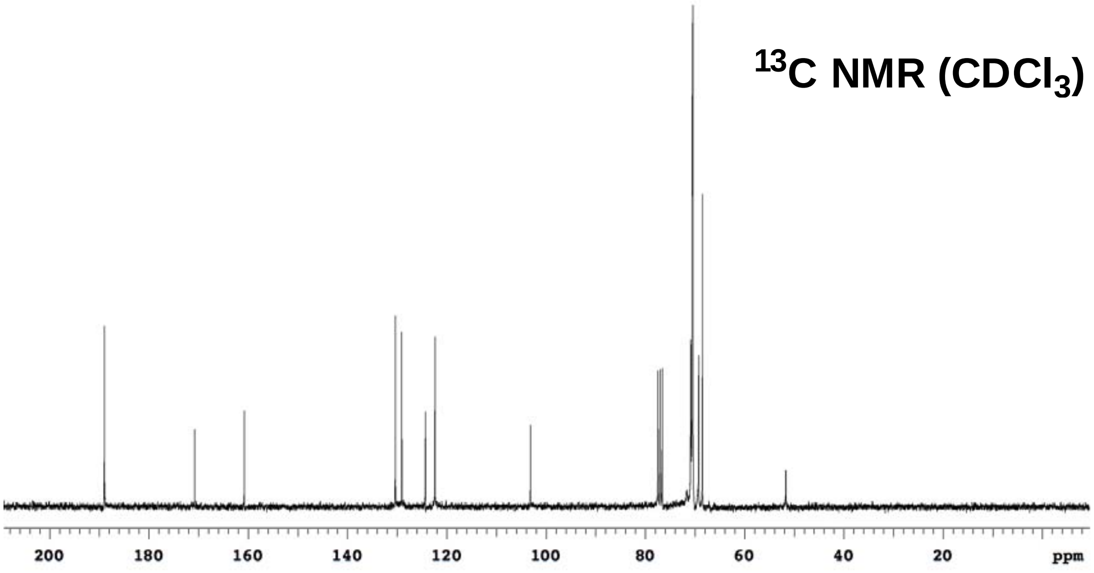


^{13}C NMR ($CDCl_3$)

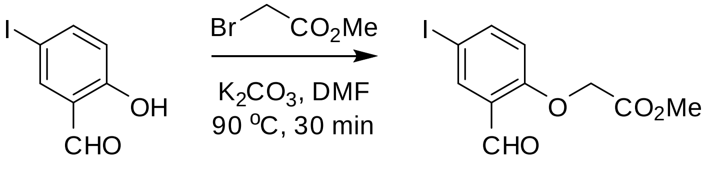




^{1}H NMR (CDCl$_3$)



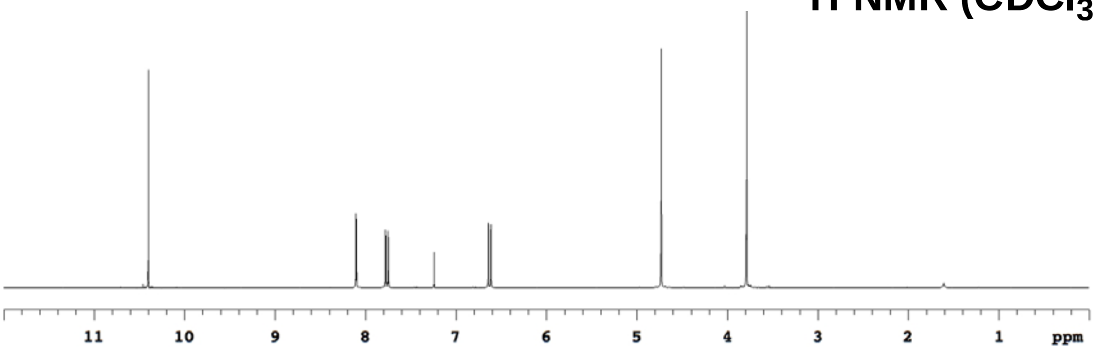
^{13}C NMR (CDCl$_3$)



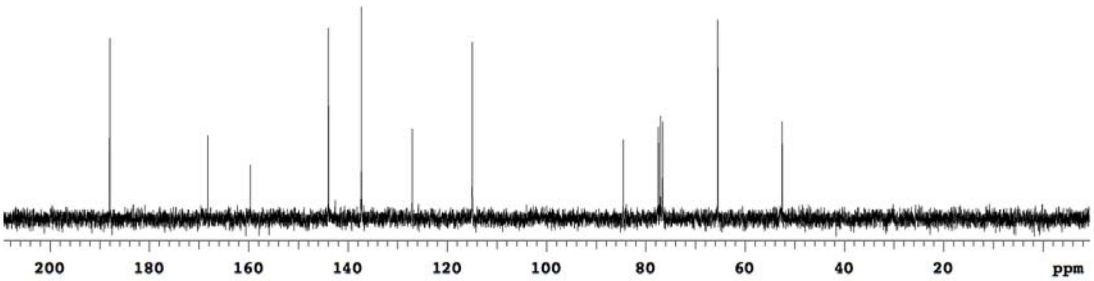
46

2-Hydroxy-5-iodobenzaldehyde (1.24 g, 5.0 mmol) and K_2CO_3 (1.38 g, 10.0 mmol) were dissolved in 25 mL DMF and stirred for 10 min then methyl bromoacetate (0.7 mL, 7.5 mmol) was added. The mixture was stirred at 90 °C for 30 min then cooled to room temperature and diluted with EtOAc. The mixture was washed with H_2O and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (20 % EtOAc/Hexanes) to afford the product as white solid (1.40 g, 87 %). ^1H NMR (300 MHz, $CDCl_3$) δ 10.40 (s, 1H), 8.10 (d, 1H, *J* = 2.3 Hz), 7.76 (dd, 1H, *J* = 8.8, 2.3 Hz), 6.63 (d, 1H, *J* = 8.8 Hz), 4.73 (s, 2H), 3.79 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 188.0, 168.2, 159.7, 143.9, 137.3, 127.0, 114.9, 84.5, 65.5, 52.5.

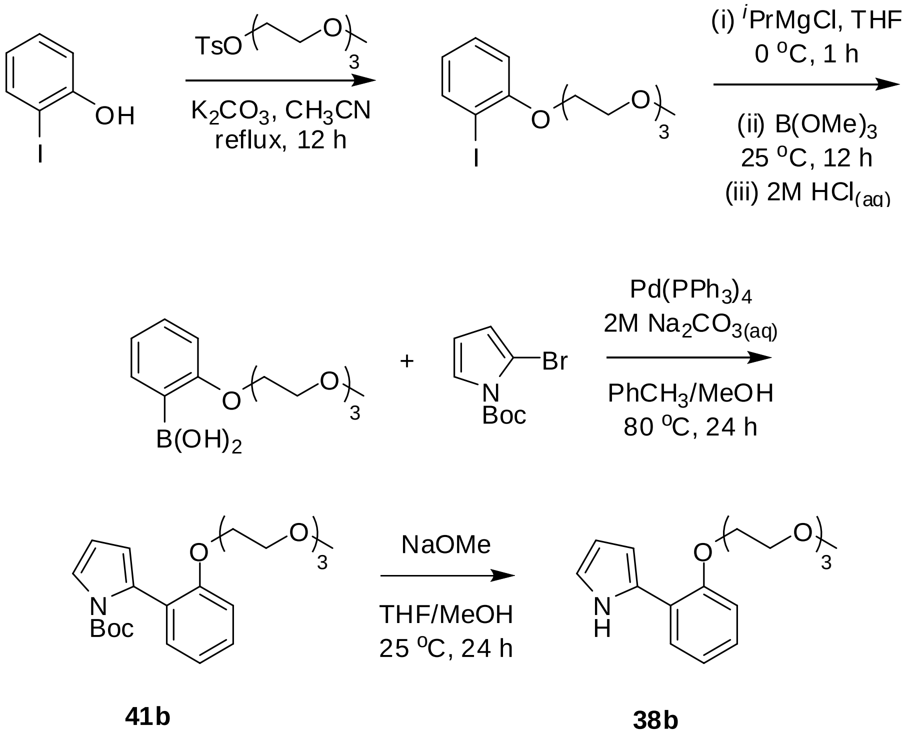
^1H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)



(2) Syntheses of pyrrole fragments 38-40

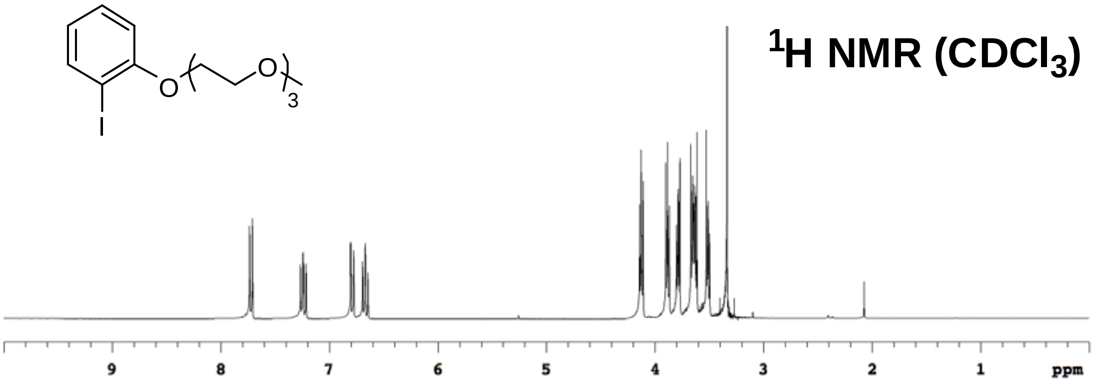

TsO
O
3
K2CO3, CH3CN
reflux, 12 h
OH
I
(i) iPrMgCl, THF
0 °C, 1 h
(ii) B(OMe)3
25 °C, 12 h
(iii) 2M HCl(aq)
B(OH)2
+
Br
N
Boc
Pd(PPh3)4
2M Na2CO3(aq)
PhCH3/MeOH
80 °C, 24 h
NaOMe
THF/MeOH
25 °C, 24 h
N
H
41b
38b

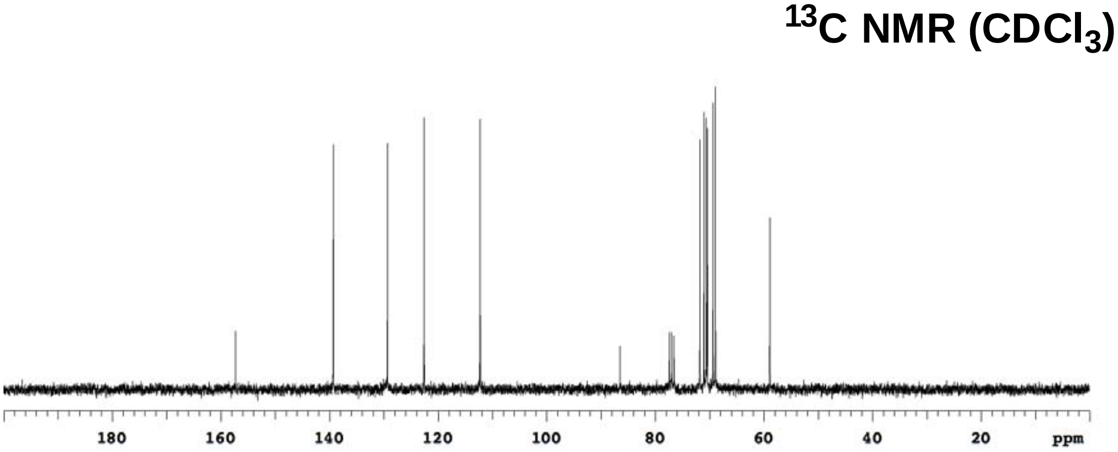
(i) A solution of 2-iodophenol (3.3 g, 15.0 mmol) and K_2CO_3 (2.7 g, 19.5 mmol) in 30 mL CH_3CN was stirred at room temperature for 10min then a solution of $TsO(CH_2CH_2O)_3CH_3$ (5.3 g, 16.5 mmol) in 20 mL CH_3CN was added dropwise over 5 min. After refluxing for 12 h, the mixture was diluted with EtOAc, washed with water and brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatograph (20 to 40 % EtOAc/Hexanes) to afford the pure product as yellow oil (5.0 g, 91 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.72 (dd, 1H, *J* = 7.8, 1.7 Hz), 7.27-7.21 (m, 1H), 6.79 (dd, 1H, *J* = 8.2, 1.3 Hz), 6.67 (td, 1H, *J* = 7.8, 1.3 Hz), 4.14-4.11 (m, 2H), 3.90-3.87 (m, 2H), 3.80-3.77 (m, 2H), 3.67-3.61 (m, 4H), 3.53-3.50 (m, 2H), 3.34 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 157.3, 139.3, 129.3, 122.6, 112.2, 86.5, 71.8, 71.1, 70.6, 70.4, 69.4, 68.9, 58.9.

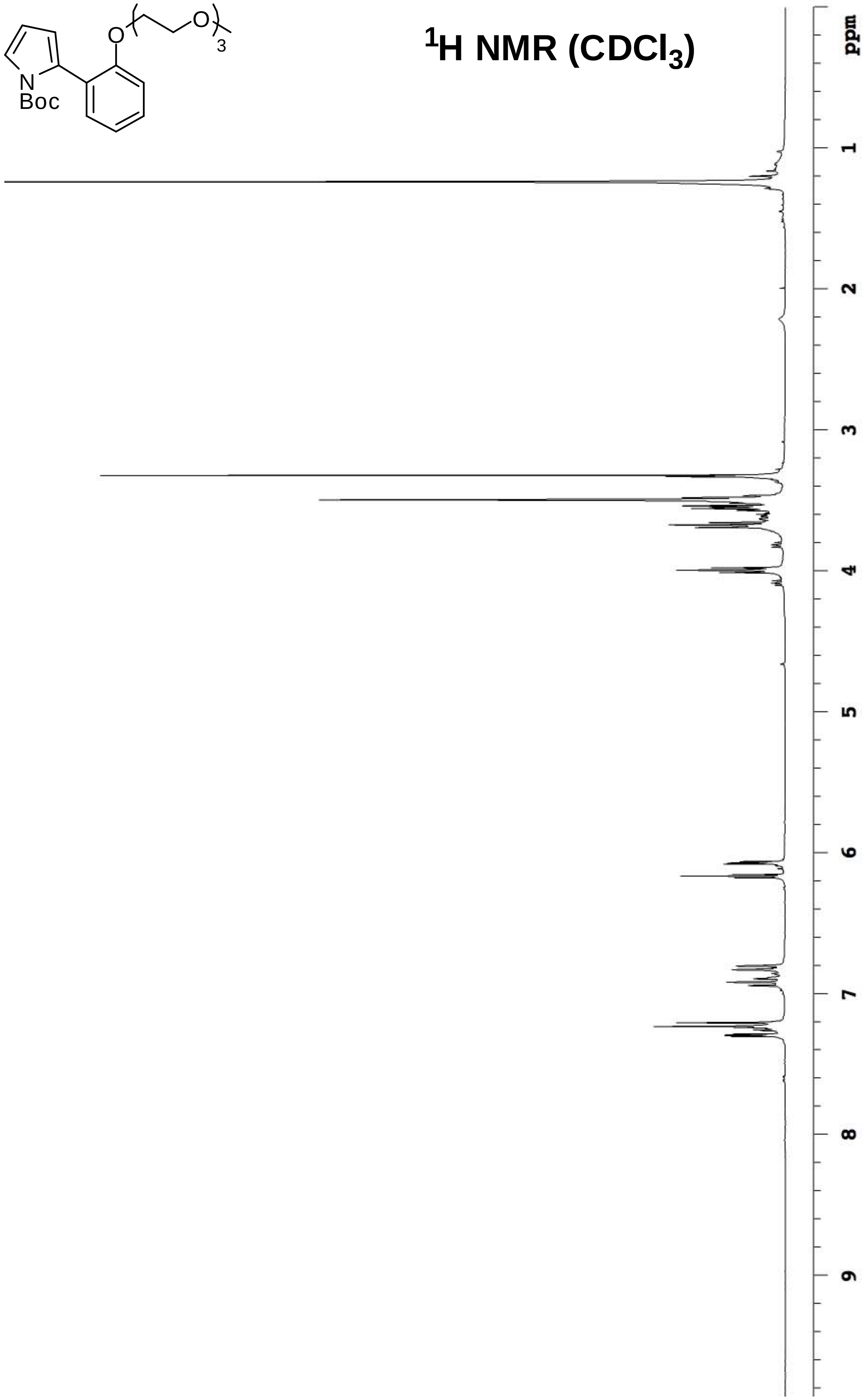
(ii) A solution of the previous product (366 mg, 1.0 mmol) in 5 mL THF was cooled to 0 ^{o}C and ^{i}PrMgCl (2M in THF, 0.70 mL, 1.4 mmol) was added dropwise over 2 min. The solution was stirred at 0 ^{o}C for 1 h then $B(OMe)_3$ (0.22 mL, 2 mmol) was added. After stirring at room temperature for 12 h, the reaction was quenched with 2M HCl at 0 ^{o}C and stirred for 20 min. The mixture was extracted with EtOAc. The organic layer was washed with water, brine and dried over Na_2SO_4 then concentrated under reduced pressure to give the product (212 mg, 75 %) as light yellow oil which was used in the next step without further purification.

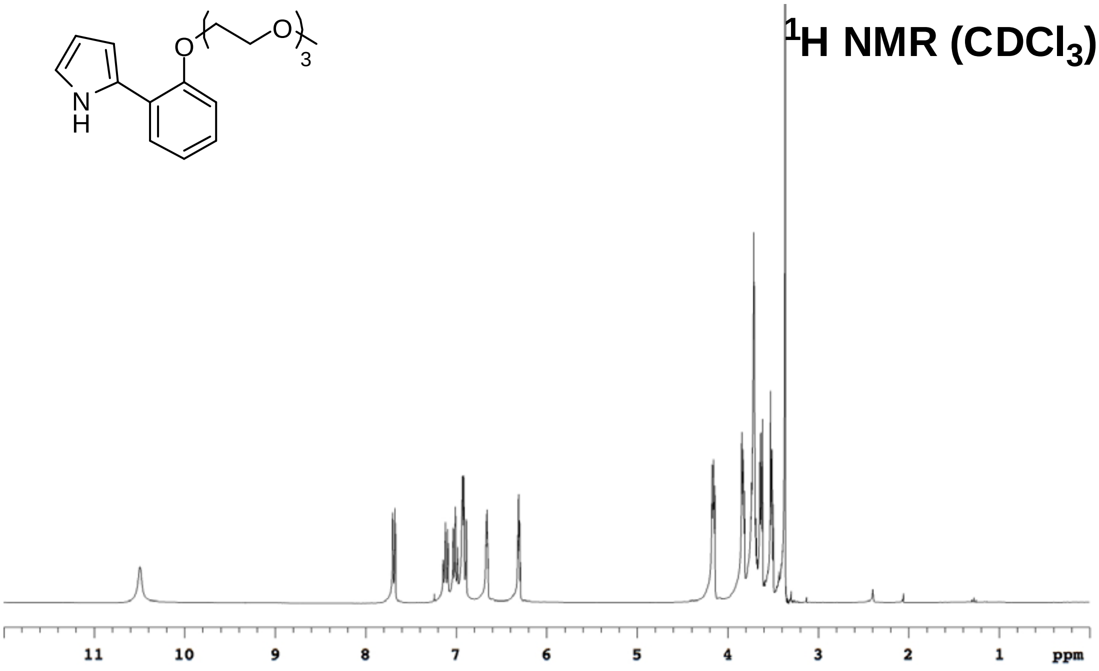
(iii) Na_2CO_3 (2M in water, 5.4 mL) was added to a solution of the pyrrole derivative (1.1 g, 4.5 mmol), arylboronic acid (1.5 g, 5.4 mmol) and $Pd(PPh_3)_4$ (104 mg, 0.1 mmol) in 20 mL $PhCH_3$ and 4 mL MeOH. The mixture was degassed and filled with N_2. After stirring at 80 ^{o}C for 24 h, the mixture was cooled to room temperature and extracted with EtOAc. The organic layer was washed with saturated $Na_2CO_{3(aq)}$, water, brine and dried over Na_2SO_4 then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et_3N/Hexanes) silica gel (40 % EtOAc/Hexanes) to afford the product **41b** (1.7 g, 94 %) as yellow oil which was used in the next step without further purification. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.30-7.28 (m, 1H), 7.26-7.21 (m, 2H), 6.95-6.80 (m, 2H), 6.18-6.15 (m, 1H), 6.08-6.06 (m, 1H), 4.01-3.98 (m, 2H), 3.69-3.66 (m, 2H), 3.57-3.47 (m, 8H), 3.32 (s, 3H), 1.24 (s, 9H).

(iv) A solution of **41b** (1.8 g, 4.5 mmol) in 20 mL THF and 5 mL MeOH was cooled to 0 ^{o}C and NaOMe (1.2 g, 22.3 mmol) was added in three portions over 15 min. The reaction mixture was stirred at room temperature for 24 h then quenched with water at 0 ^{o}C. The mixture was extracted with EtOAc. The organic layer was washed with water, brine and dried over Na_2SO_4 then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et_3N/Hexanes) silica gel (40 % EtOAc/Hexanes) to afford the product **38b** (1.2 g, 87 %) as yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.49 (br, 1H), 7.70-7.67 (m, 1H), 7.14-7.09 (m, 1H), 7.03-6.98 (m, 1H), 6.94-6.88 (m, 2H), 6.67-6.65 (m, 1H), 6.32-6.29 (m, 1H), 4.17-4.14 (m, 2H), 3.84-3.81 (m, 2H), 3.74-3.69 (m, 4H), 3.65-3.61 (m, 2H), 3.53-3.50 (m, 2H), 3.36 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 153.4, 129.1, 126.2, 126.0, 121.8, 121.5, 118.0, 113.3, 108.0, 105.7, 71.4, 70.1 (2C: 70.14, 70.11), 69.9, 68.9, 67.1, 58.5.

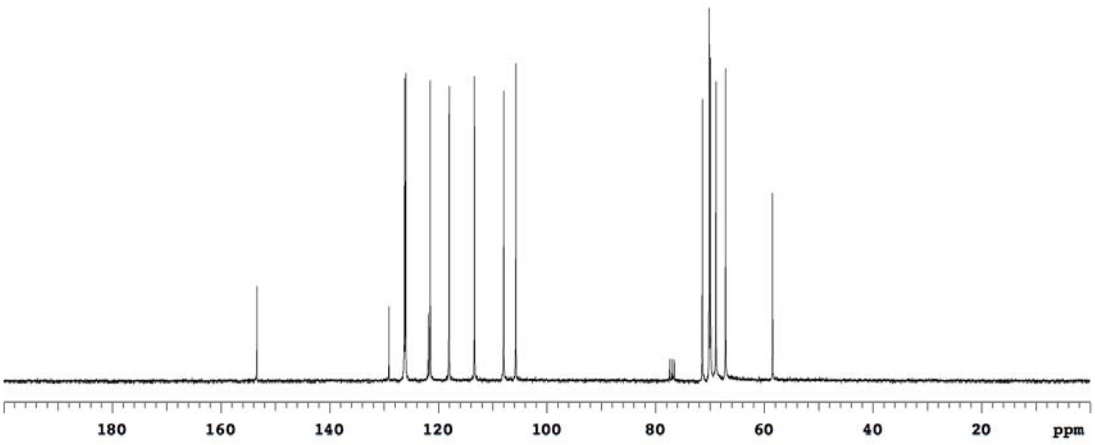




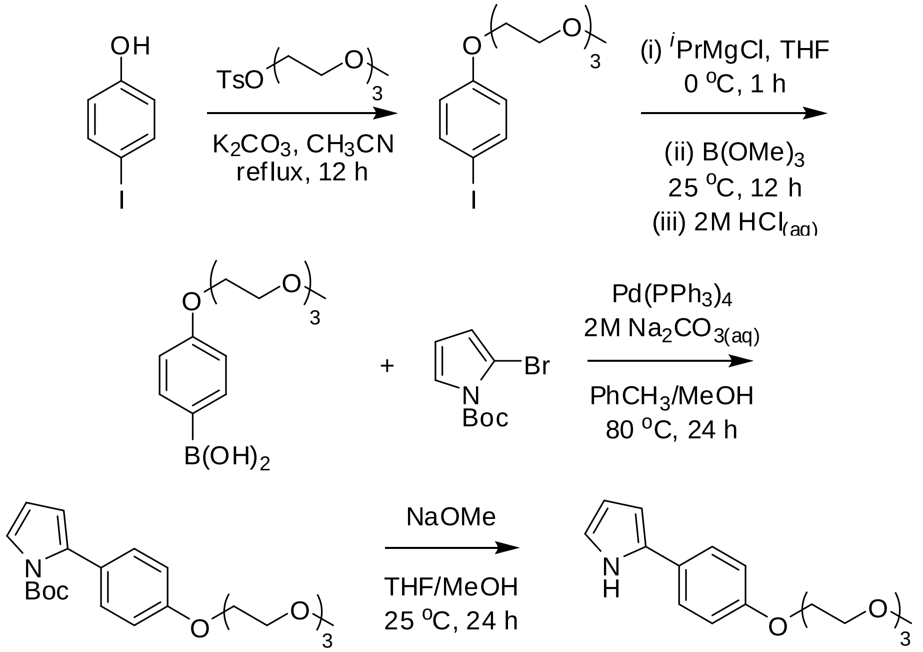

N
Boc
O
O
3
¹H NMR (CDCl₃)
ppm
1
2
3
4
5
6
7
8
9



^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)


OH
I
TsO
O
3
K2CO3, CH3CN
reflux, 12 h
O
O
3
I
(i) iPrMgCl, THF
0 °C, 1 h
(ii) B(OMe)3
25 °C, 12 h
(iii) 2M HCl(aq)
O
O
3
B(OH)2
+
Br
N
Boc
Pd(PPh3)4
2M Na2CO3(aq)
PhCH3/MeOH
80 °C, 24 h
N
Boc
O
O
3
NaOMe
THF/MeOH
25 °C, 24 h
N
H
O
O
3

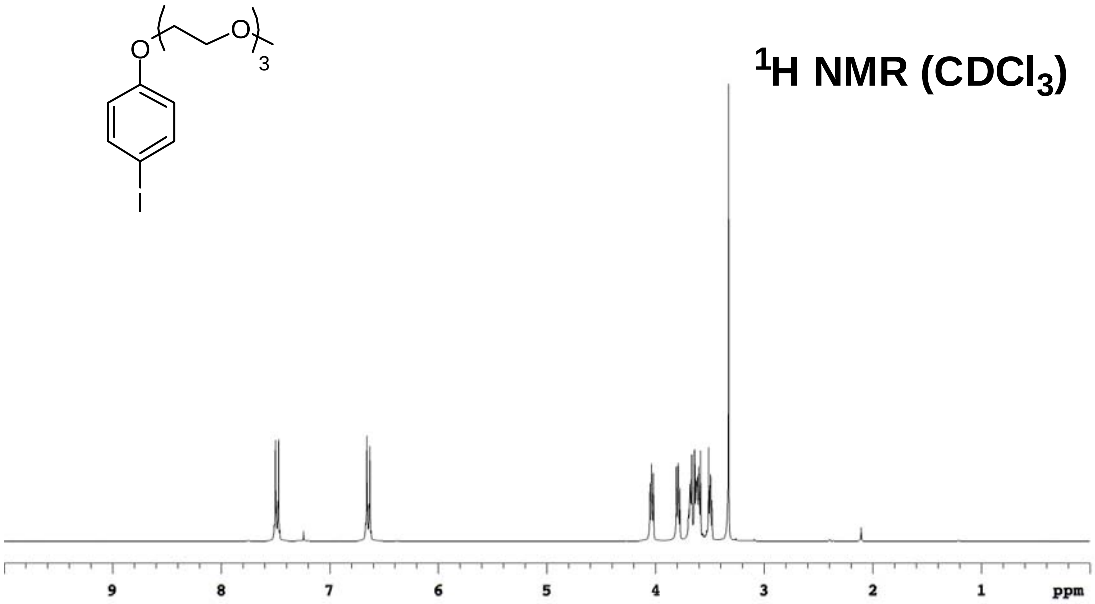
41c **38c**

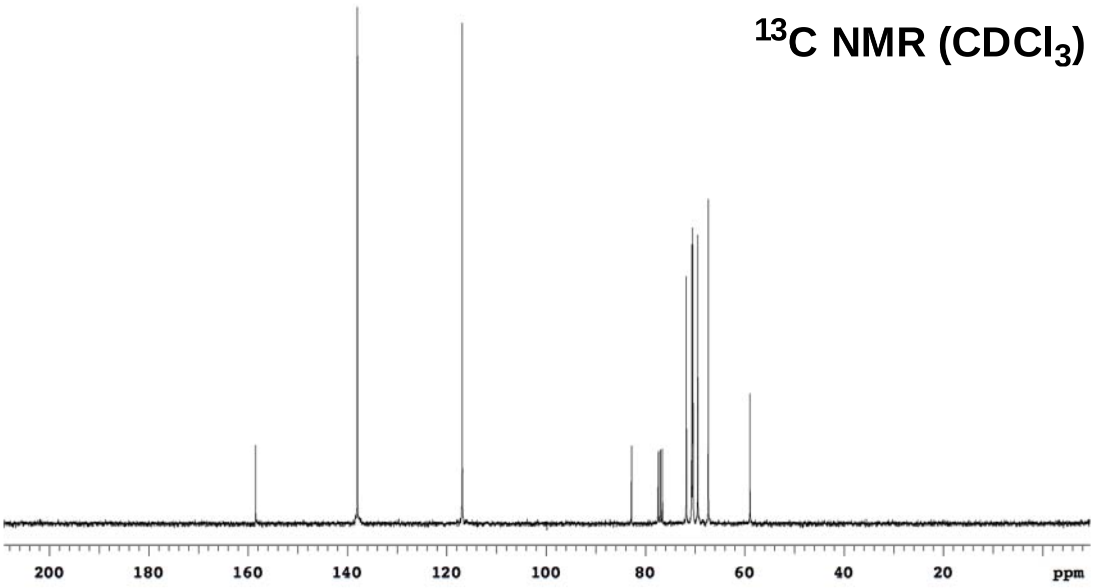
(i) A solution of 2-iodophenol (3.3 g, 15.0 mmol) and K_2CO_3 (2.7 g, 19.5 mmol) in 30 mL CH_3CN was stirred at room temperature for 10min then a solution of $TsO(CH_2CH_2O)_3CH_3$ (5.3 g, 16.5 mmol) in 20 mL CH_3CN was added dropwise over 5 min. After refluxing for 12 h, the mixture was diluted with EtOAc, washed with water and brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatograph (20 to 40 % EtOAc/Hexanes) to afford the pure product as yellow oil (4.5 g, 82 %). ^1H NMR (300 MHz, $CDCl_3$) δ 7.49 (d, 2H, J = 9.0 Hz), 6.64 (d, 2H, J = 9.0 Hz), 4.05-4.02 (m, 2H), 3.81-3.77 (m, 2H), 3.70-3.66 (m, 2H), 3.64-3.58 (m, 4H), 3.51-3.48 (m, 2H), 3.33 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 158.5, 138.0, 116.9, 82.8, 71.8, 70.7, 70.5, 70.4, 69.4, 67.3, 58.9.

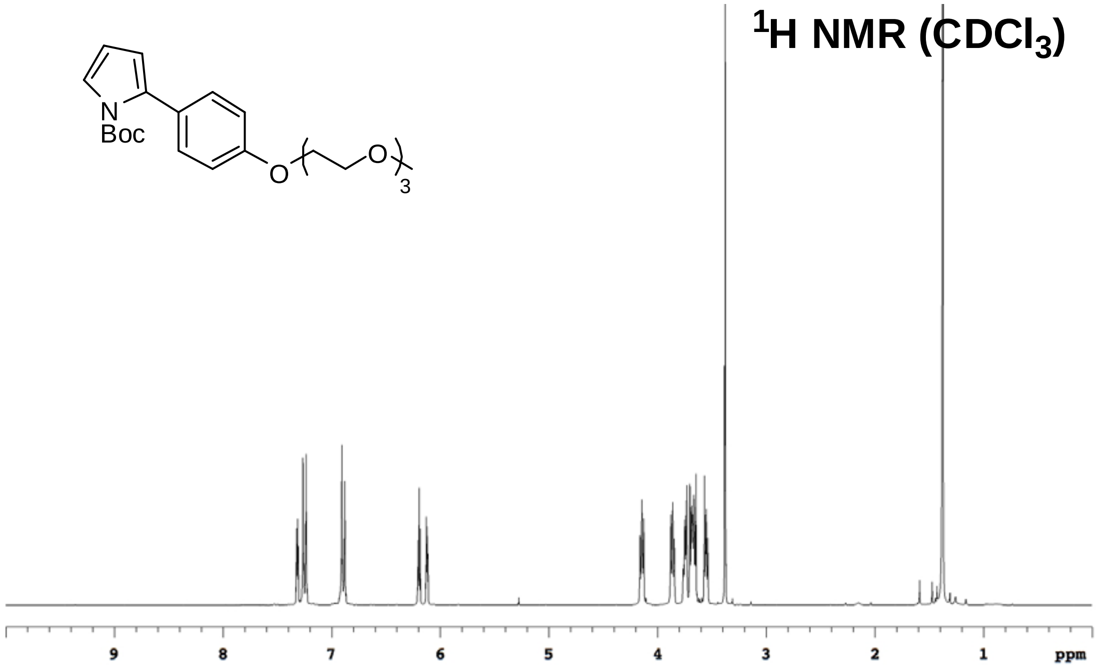
(ii) A solution of the previous product (1.83 g, 5 mmol) in 25 mL THF was cooled to 0 °C and ^{i}PrMgCl (2M in THF, 3.5 mL, 7 mmol) was added dropwise over 5 min. The solution was stirred at 0 °C for 1 h then $B(OMe)_3$ (1.2 mL, 10 mmol) was added. After stirring at room temperature for 12 h, the reaction was quenched with 2M HCl at 0 °C and stirred for 20 min. The mixture was extracted with EtOAc. The organic layer was washed with water, brine and dried over Na_2SO_4 then concentrated under reduced pressure to give the product (915 mg, 64 %) as light yellow oil which was used in the next step without further purification.

(iii) Na_2CO_3 (2M in water, 3.3 mL) was added to a solution of the pyrrole derivative (660 mg, 2.7 mmol), arylboronic acid (915 mg, 3.2 mmol) and $Pd(PPh_3)_4$ (62 mg, 0.05 mmol) in 15 mL $PhCH_3$ and 3 mL MeOH. The mixture was degassed and filled with N_2. After stirring at 80 °C for 24 h, the mixture was cooled to room temperature and extracted with EtOAc. The organic layer was washed with saturated $Na_2CO_{3(aq)}$, water, brine and dried over Na_2SO_4 then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et_3N/Hexanes) silica gel (50 % EtOAc/Hexanes) to afford the product **41c** (1.0 g, 94 %) as yellow oil. ^1H NMR (300 MHz, $CDCl_3$) δ 7.33-7.31 (m, 1H), 7.25 (d, 2H, J = 9.0 Hz), 6.89 (d, 2H, J = 9.0 Hz), 6.21-6.18 (m, 1H), 6.13-6.11 (m, 1H), 4.16-4.13 (m, 2H), 3.88-3.85 (m, 2H), 3.76-3.73 (m, 2H), 3.70-3.65 (m, 4H), 3.57-3.54 (m, 2H), 3.38 (s, 3H), 1.37 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 157.9, 149.2, 134.6, 130.2, 126.8, 122.0, 113.8, 113.5, 110.3, 83.2, 77.8, 70.7, 70.5, 70.4, 69.6, 67.2, 58.9, 27.5.

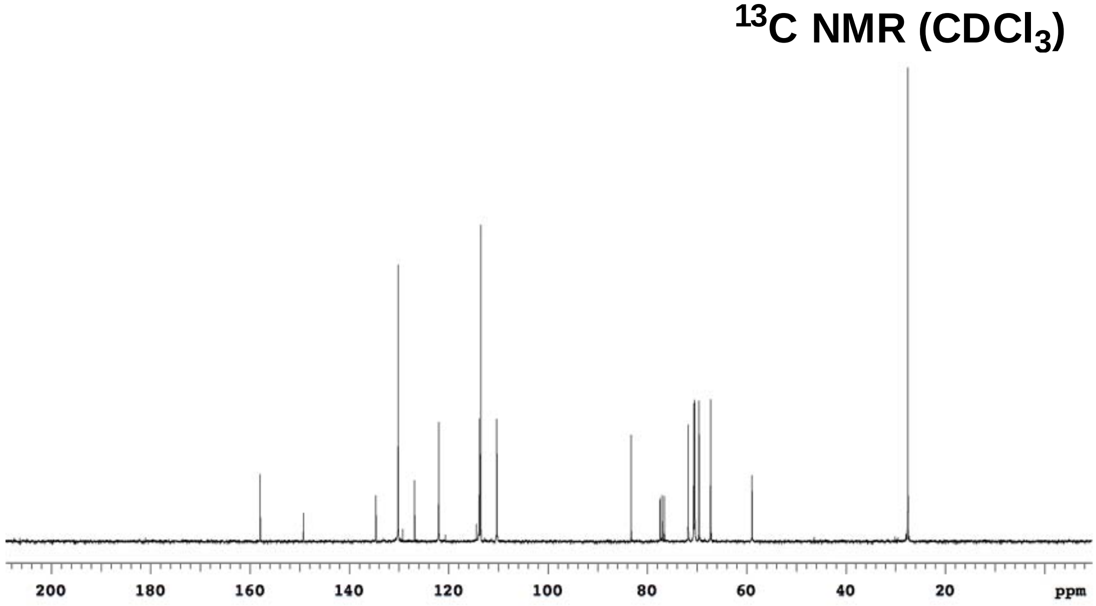
(iv) A solution of **41c** (1.48 g, 3.7 mmol) in 20 mL THF and 5 mL MeOH was cooled to 0 °C and NaOMe (0.99 g, 18.3 mmol) was added in three portions over 15 min. The reaction mixture was stirred at room temperature for 24 h then quenched with water at 0 °C. The mixture was extracted with EtOAc. The organic layer was washed with water, brine and dried over Na_2SO_4 then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et_3N/Hexanes) silica gel (40 % EtOAc/Hexanes) to afford the product **38c** (912 mg, 82 %) as yellow oil. ^1H NMR (300 MHz, $CDCl_3$) δ 9.10 (br, 1H), 7.35 (d, 2H, J = 9.0 Hz), 6.78 (d, 2H, J = 9.0 Hz), 6.74-6.72 (m, 1H), 6.40-6.38 (m, 1H), 6.24-6.21 (m, 1H), 3.98-3.95 (m, 2H), 3.75-3.71 (m, 2H), 3.68-3.59 (m, 6H), 3.51-3.48 (m, 2H), 3.31 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 156.6, 131.4, 125.7, 124.6, 118.1, 114.4, 109.0, 104.1, 71.4, 70.2, 70.1, 70.0, 69.2, 66.8, 58.4.


^{1}H NMR (CDCl$_3$)

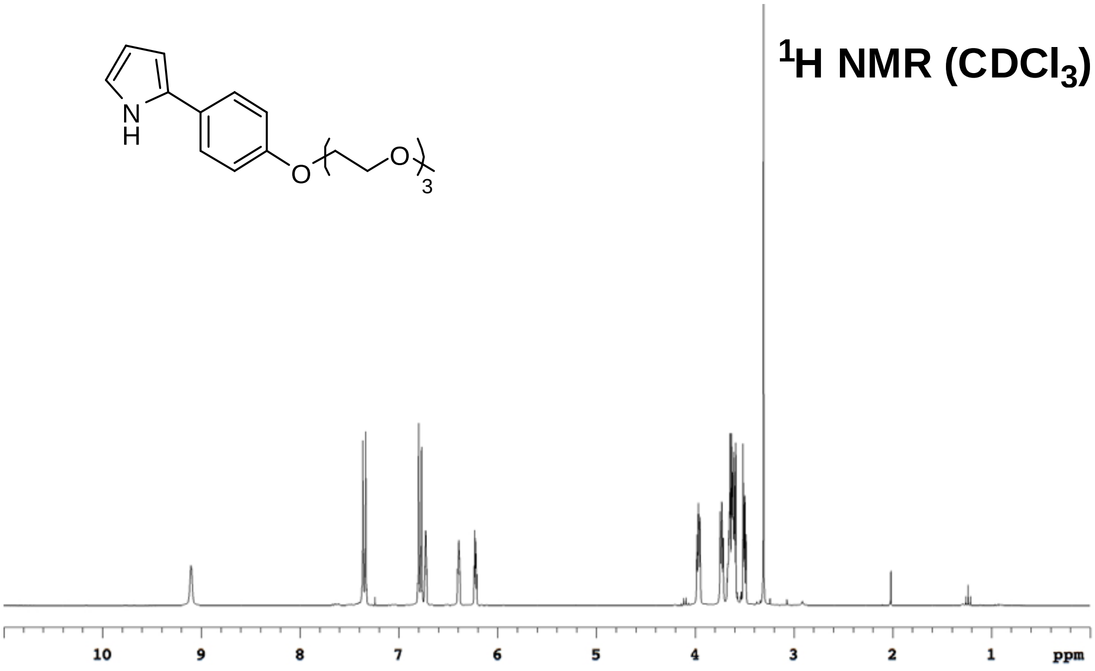

^{13}C NMR (CDCl$_3$)

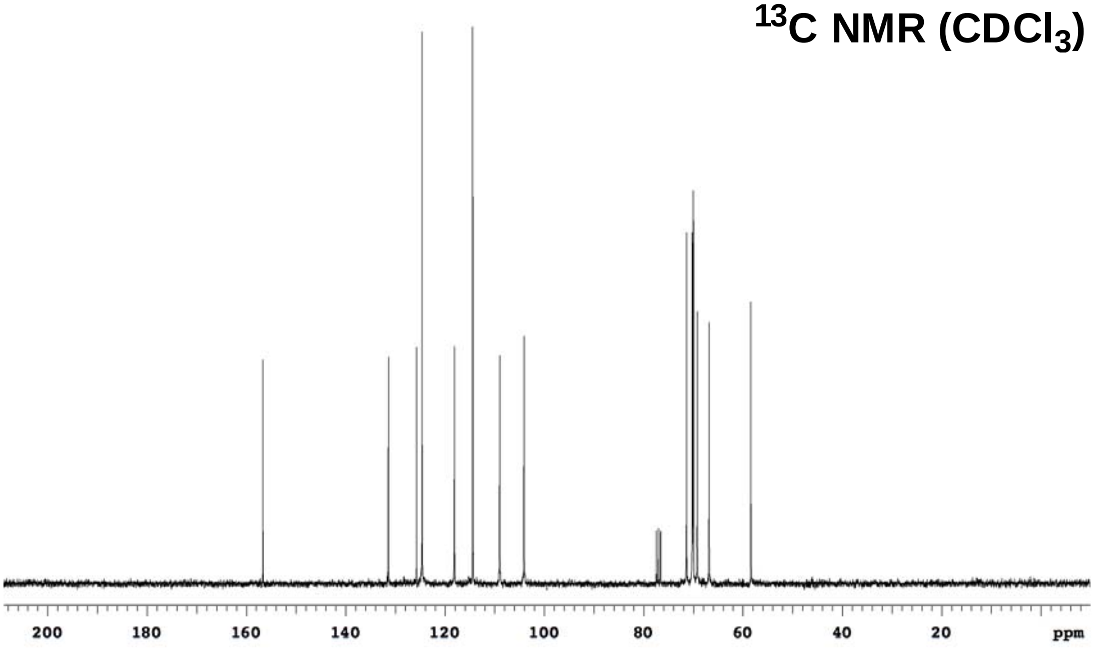


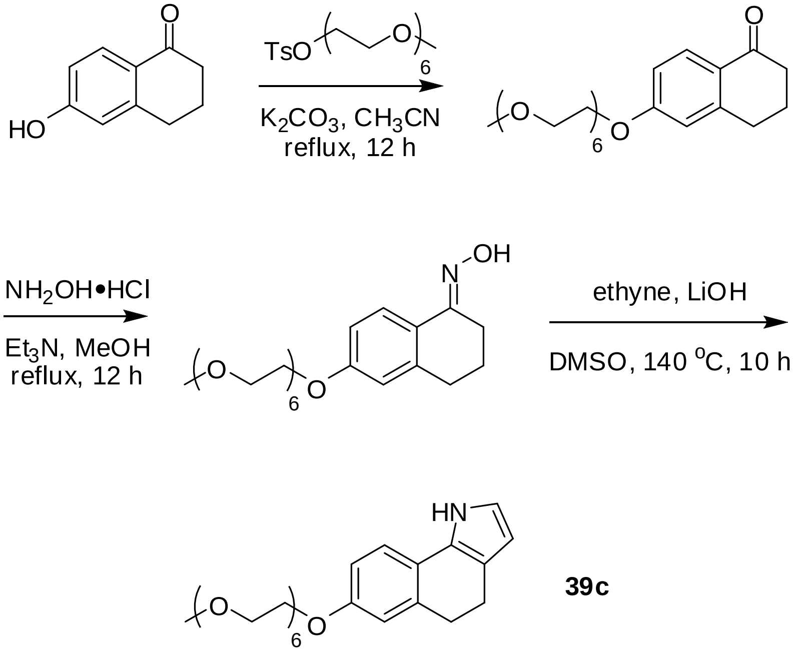
N
Boc
O
O
3
^{1}H NMR (CDCl$_3$)
9
8
7
6
5
4
3
2
1
ppm



^{13}C NMR (CDCl$_3$)
200
180
160
140
120
100
80
60
40
20
ppm


1H NMR (CDCl3)
N
H
O
O
3
10
9
8
7
6
5
4
3
2
1
ppm


13C NMR (CDCl3)
200
180
160
140
120
100
80
60
40
20
ppm

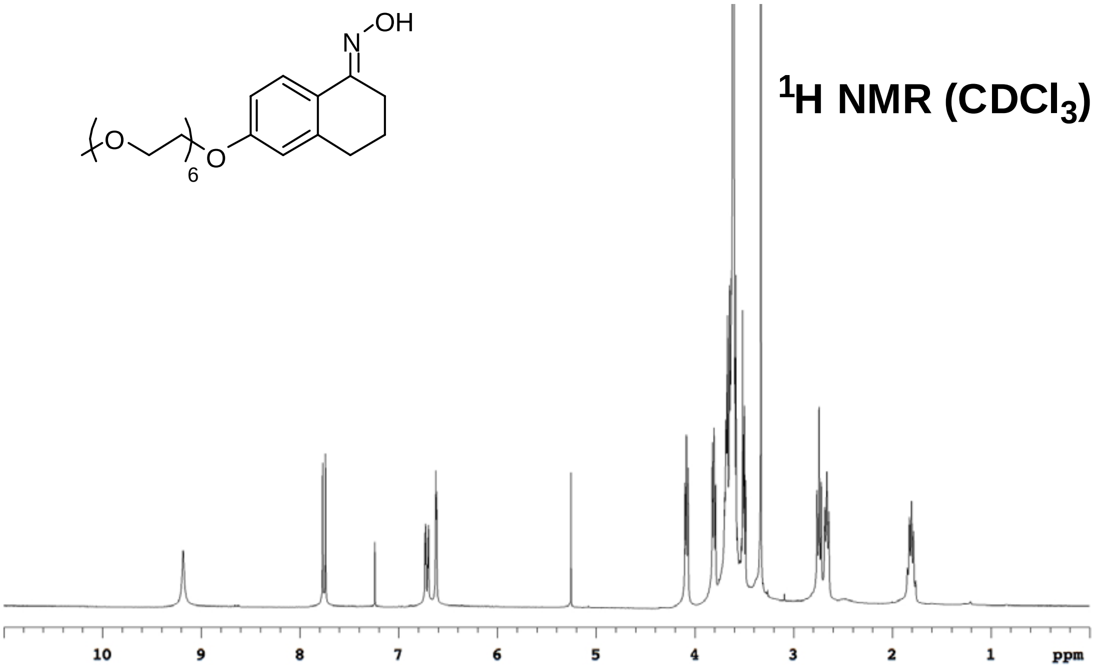


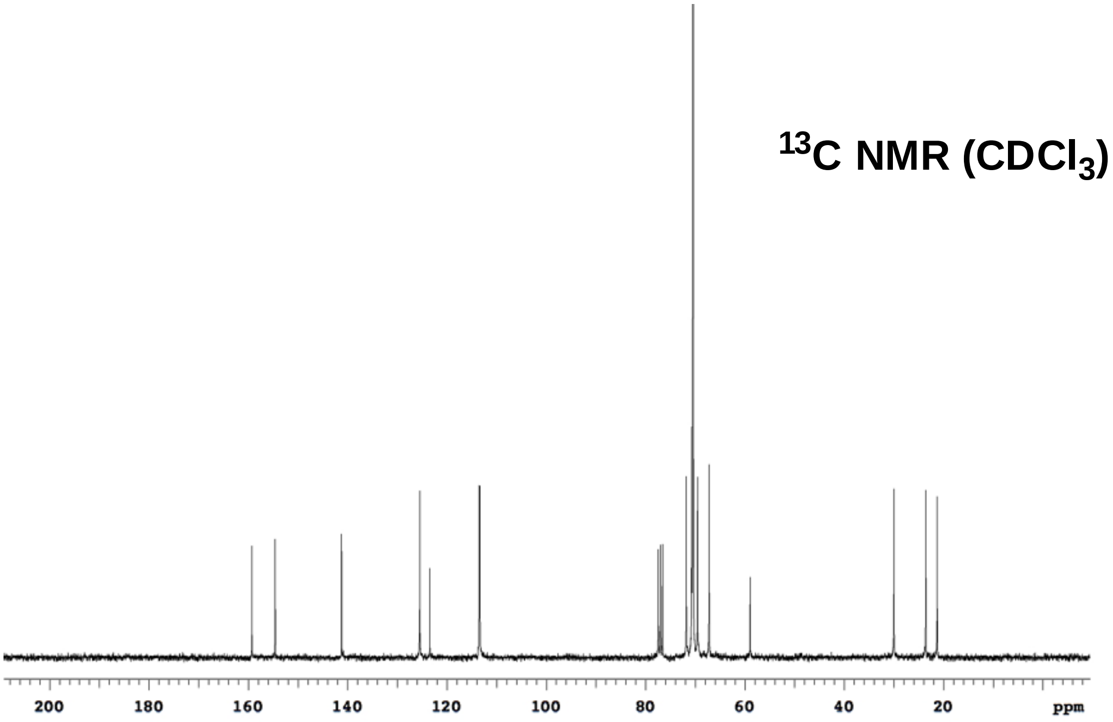
39c

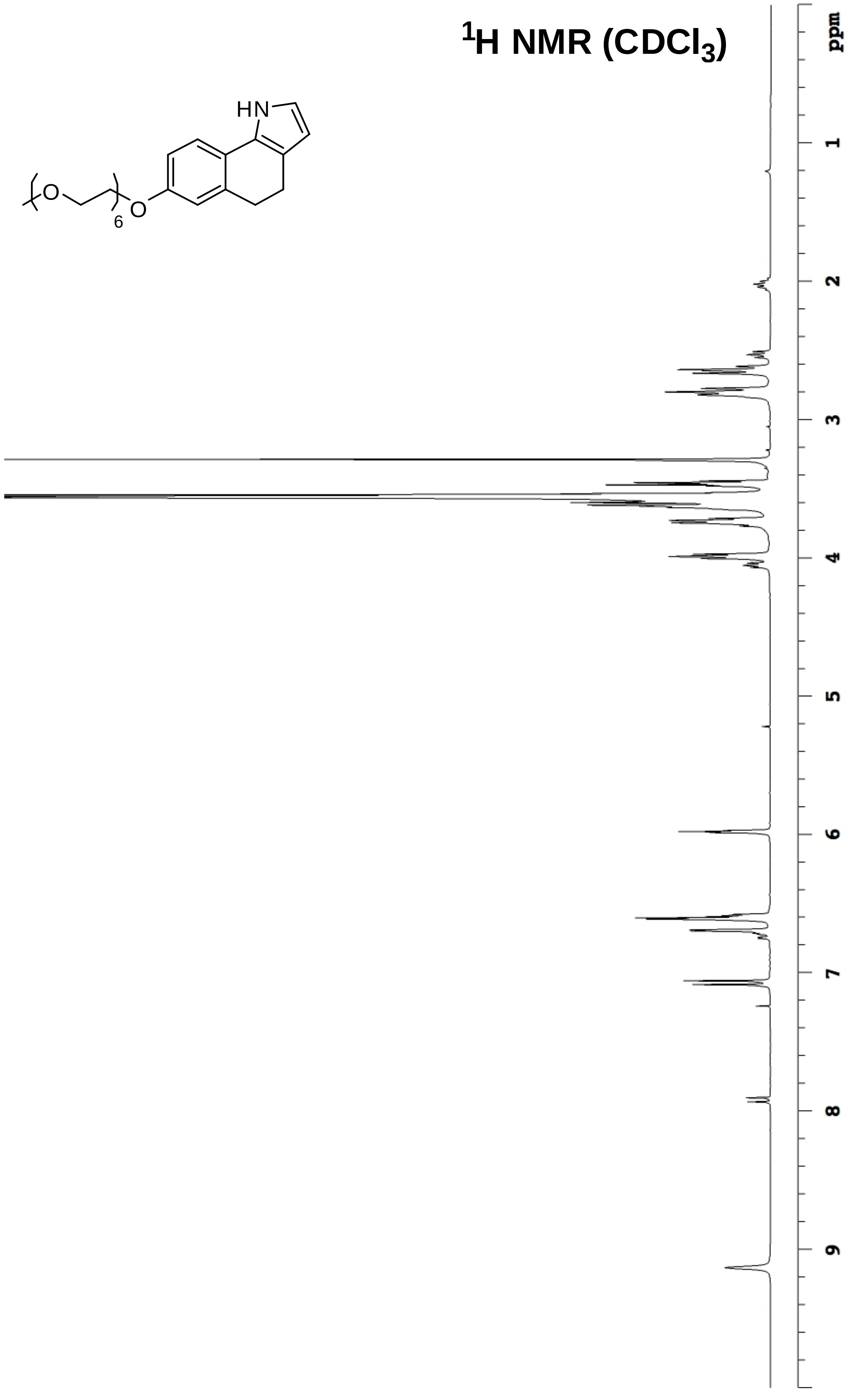
(i) 6-Hydroxy-1-tetralone (2.11 g, 13.0 mmol) and K_2CO_3 were dissolved in 30 mL CH_3CN and stirred for 10 min then a solution hexa(ethelene glycol) monomethyl ether tosylate (6.17 g, 13.7 mmol) in 20 mL CH_3CN was added. The mixture was refluxed for 12 h and cooled to room temperature. The reaction mixture was diluted with EtOAc and washed with H_2O and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH_2Cl_2) to afford the product (4.84 g, 85 %) as yellow oil which was used in the next step without further purification.

(ii) A solution of the previous product (4.77 g, 10.8 mmol), $NH_2OH{\cdot}HCl$ (2.25 g, 32.4 mmol) and Et_3N (9.0 mL, 64.8 mmol) in 50 mL MeOH was refluxed for 12 h. The reaction mixture was diluted with EtOAc and washed with H_2O and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH_2Cl_2) to afford the product (4.57 g, 93 %) as yellow oil. ^1H NMR (300 MHz, $CDCl_3$) δ 9.18 (br, 1H), 7.75 (d, 1H, J = 8.8 Hz), 6.71 (dd, 1H, J = 8.8, 2.7 Hz), 6.62 (d, 1H, J = 2.7 Hz), 4.10-4.07 (m, 2H), 3.82-3.79 (m, 2H), 3.69-3.58 (m, 18 H), 3.51-3.48 (m, 2H), 3.33 (s, 3H), 2.74 (t, 2H, *J* = 6.6 Hz), 2.66 (t, 2H, *J* = 5.9 Hz), 1.85-1.76 (m, 2H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 159.3, 154.6, 141.2, 125.5, 123.4, 113.5, 113.4, 71.8, 70.7, 70.5, 70.4 (2 peaks: 70.44, 70.36), 69.5, 67.2, 58.9, 30.0, 23.6, 21.3. (Several peaks account for more than one carbon).

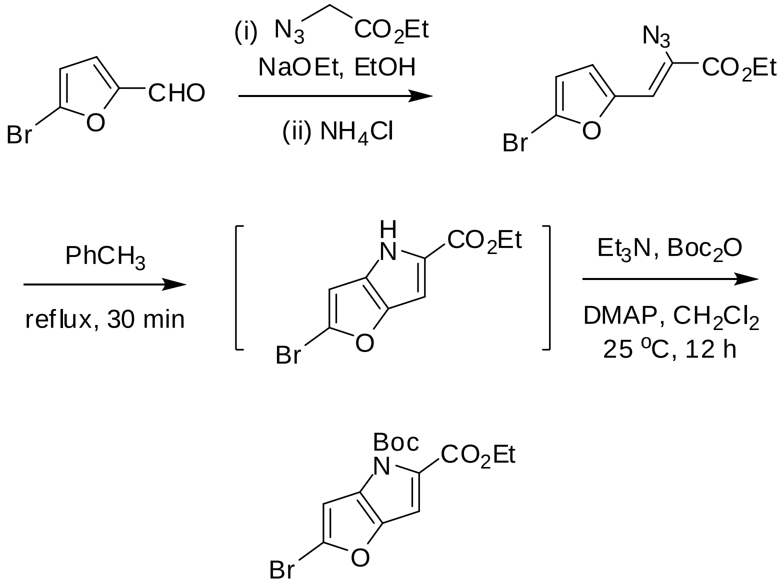
(iii) Product from (ii) (3.1 g, 6.7 mmol) and LiOH (802 mg, 33.4 mmol) were dissolved in 50 mL DMSO and ethyne was bubbled in. After stirring at 140 °C for 10 h, the mixture was cooled to room temperature and diluted with EtOAc then washed with H_2O and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH_2Cl_2) to afford the product (1.31 g, 42 %) as yellow oil (some contamination from the starting material), which was used in the next step without further purification. ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.13 (br, 1H), 7.07 (d, 1H, *J* = 8.3 Hz), 6.69 (d, 1H, *J* = 2.4 Hz), 6.61-6.58 (m, 2H), 5.99-5.97 (m, 1H), 4.00-3.97 (m, 2H), 3.75-3.71 (m, 2H), 3.63-3.54 (m, 18H), 3.47-3.44 (m, 2H), 3.29 (s, 3H), 2.82-2.77 (m, 2H), 2.66-2.62 (m, 2H).


N
OH
O
O
6
^{1}H NMR (CDCl$_3$)
10
9
8
7
6
5
4
3
2
1
ppm


^{13}C NMR (CDCl$_3$)
200
180
160
140
120
100
80
60
40
20
ppm



¹H NMR ($CDCl_3$)

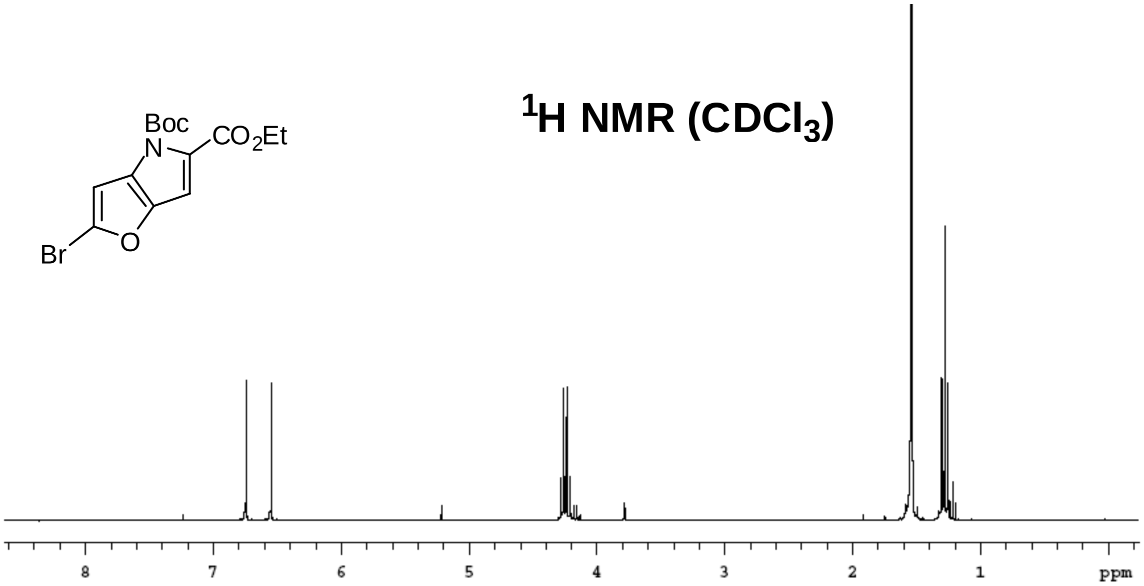


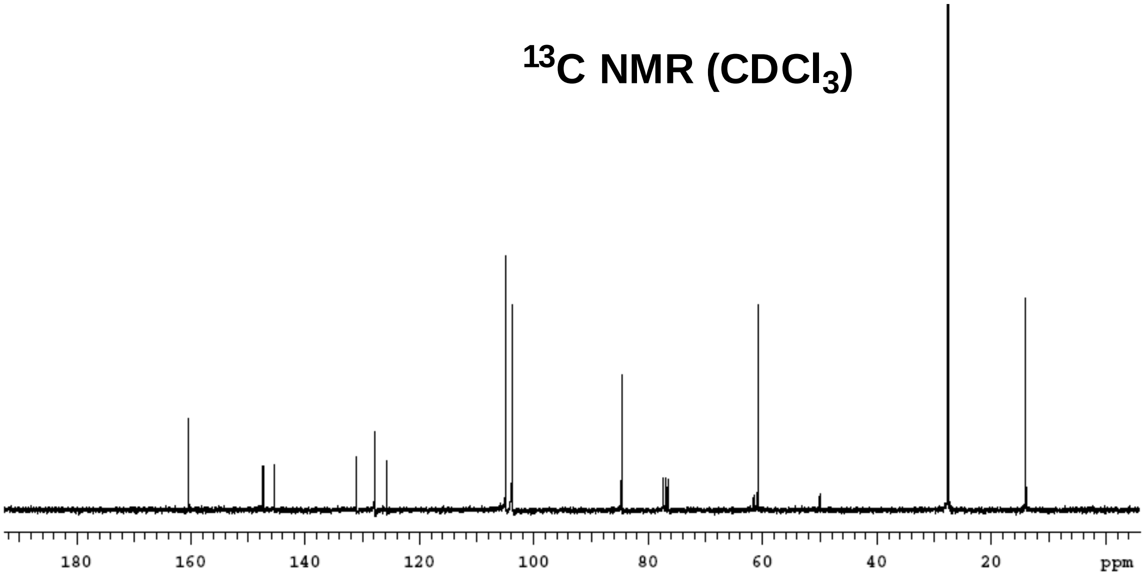
42 73 % (3 steps)

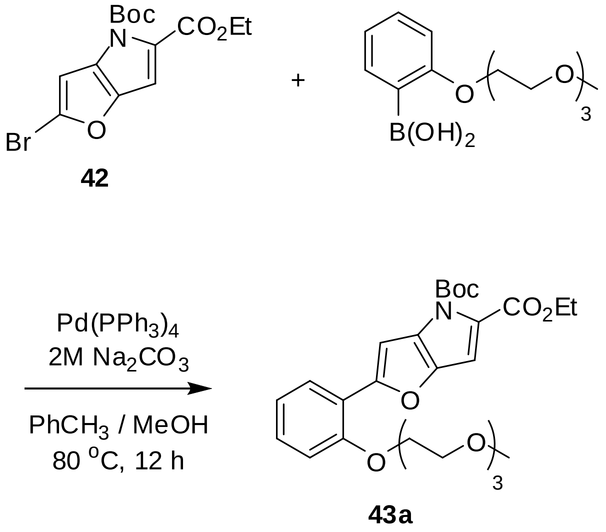
(1) A solution of 5-bromo-2-furaldehyde (1.75 g, 10 mmol) and ethyl azidoacetate (10.3 g, 80 mmol) in 100 mL EtOH was cooled to 0 °C then sodium ethoxide (15 mL, 20 % in EtOH) was added dropwise over 5 min. The reaction mixture was stirred at 10 °C for 1 h then quenched with sat'd $NH_4Cl_{(aq)}$ at -78 °C. The mixture was diluted with water and extracted with Et_2O (3 times). The organic layer was washed with brine, dried over Na_2SO_4 and evaporated under reduced pressure. The residue was passed through a short pad of silica gel and eluted with 50 % CH_2Cl_2/hexanes. The product was used for the next step without further purification.

(2) A solution of the above product in 30 mL $PhCH_3$ was heated to reflux and stirred for 20 min then cooled to 25 °C. Et_3N (1.7 mL, 12 mmol) and DMAP (122 mg, 1 mmol) were added to the reaction mixture followed by a solution of Boc_2O (2.62 g, 12 mmol) in 20 mL CH_2Cl_2. The mixture was stirred at 25 °C for 12 h then washed with water and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness under reduced pressure. The residue was purified by flash chromatography (5 % Ethyl acetate/Hexanes) to afford the pure product **42** (2.6 g, 73 %) as a yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 6.74 (d, 1H, *J* = 0.8 Hz), 6.55 (d, 1H, *J* = 0.8 Hz), 4.25 (q, 2H, *J* =

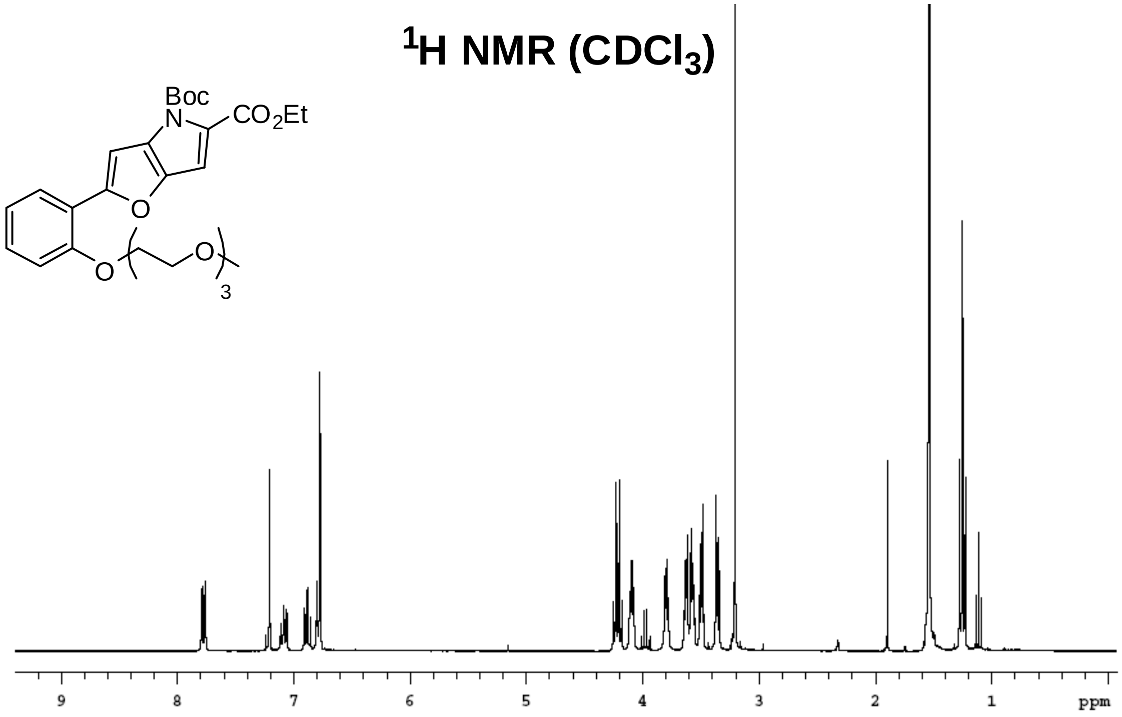
7.2 Hz), 1.54 (s, 9H), 1.28 (t, 3H, *J* = 7.2 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 160.4, 147.4, 145.4, 131.0, 127.9, 125.8, 104.9, 103.9, 84.7, 60.8, 27.6, 14.0.



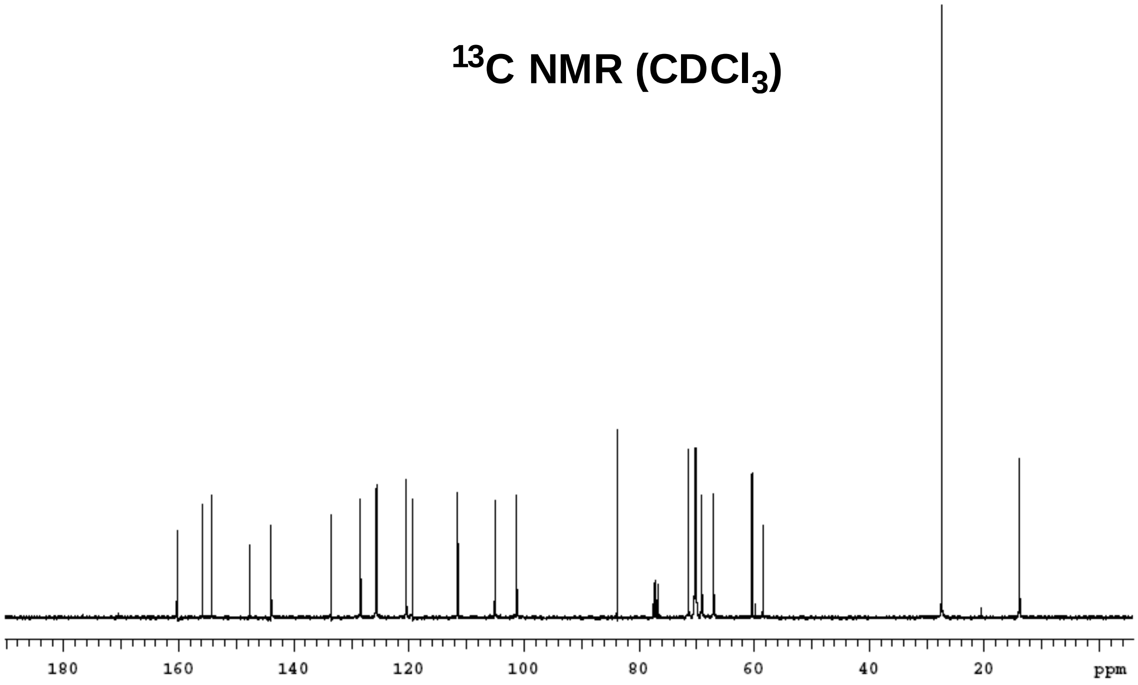




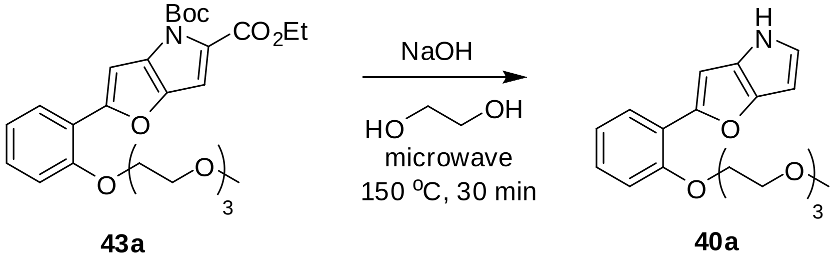
To a solution of pyrrole **42** (1.31 g, 3.7 mmol), arylboronic acid (1.26 g, 4.4 mmol) and $Pd(PPh_3)_4$ (85 mg, 0.07 mmol) in 20 mL $PhCH_3$ and 4 mL MeOH was added Na_2CO_3 (2M in water, 4.4 mL). The mixture was degassed and filled with N_2. After stirring at 80 ^{o}C for 24 h, the mixture was cooled to room temperature and extracted with EtOAc. The organic layer was washed with saturated $Na_2CO_{3\ (aq)}$, water, brine and dried over Na_2SO_4 then concentrated under reduced pressure. The residue was purified by flash chromatography (40 % EtOAc/Hexanes) to afford the product **43a** (1.33 g, 70 %) as a yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.77 (dd, 1H, J = 7.8, 1.7 Hz), 7.21 (d, 1H, J = 0.9 Hz), 7.11-7.06 (m, 1H), 6.91-6.85 (m, 1H), 6.80-6.77 (m, 2H), 4.21 (q, 2H, J = 7.1 Hz), 4.11-4.08 (m, 2H), 3.81-3.78 (m, 2H), 3.65-3.61 (m, 2H), 3.58-3.55 (m, 2H), 3.51-3.48 (m, 2H), 3.37-3.34 (m, 2H), 3.20 (s, 3H), 1.54 (s, 9H), 1.25 (t, 3H, J = 7.1 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ 160.3, 155.9, 154.2, 147.7, 143.9, 133.5, 128.4, 125.7, 125.6, 120.5, 119.4, 111.5, 105.0, 101.2, 83.8, 71.4, 70.3, 70.2, 70.0, 69.1, 67.0, 60.3, 58.4, 27.4, 13.9. MS (ESI) m/z calcd for $(M+H)^+$ $C_{27}H_{36}NO_9$ 518.24; found 518.24.



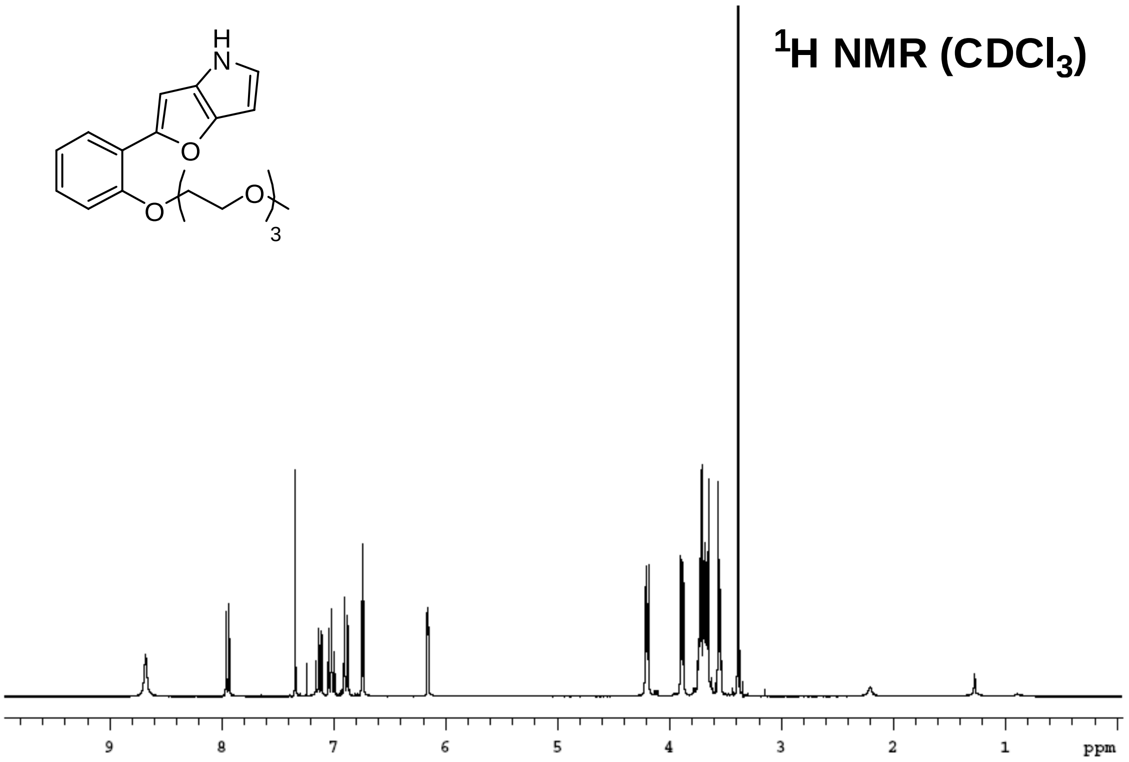
^{1}H NMR (CDCl$_3$)
Boc
N
CO$_2$Et
O
O
O
3
9
8
7
6
5
4
3
2
1
ppm



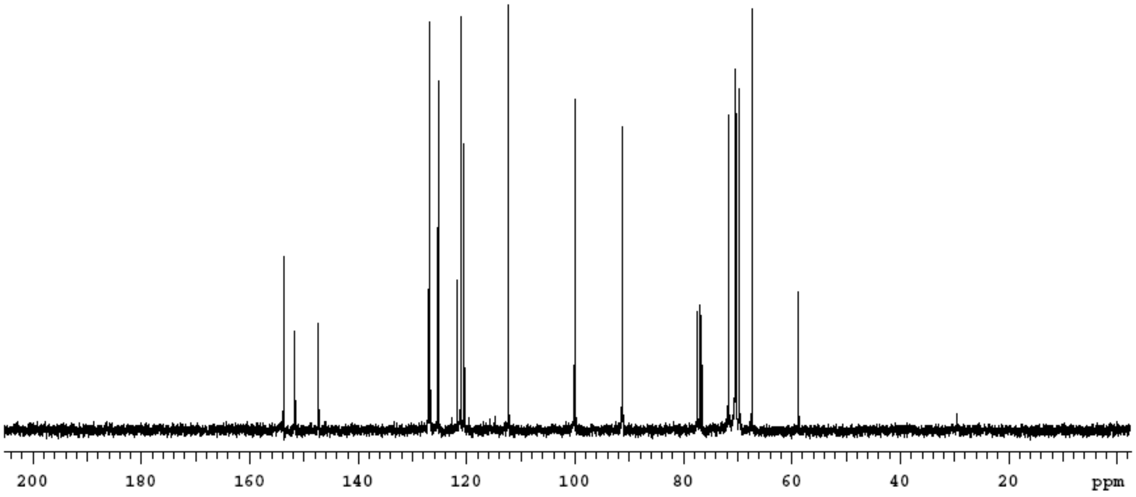
^{13}C NMR (CDCl$_3$)
180
160
140
120
100
80
60
40
20
ppm



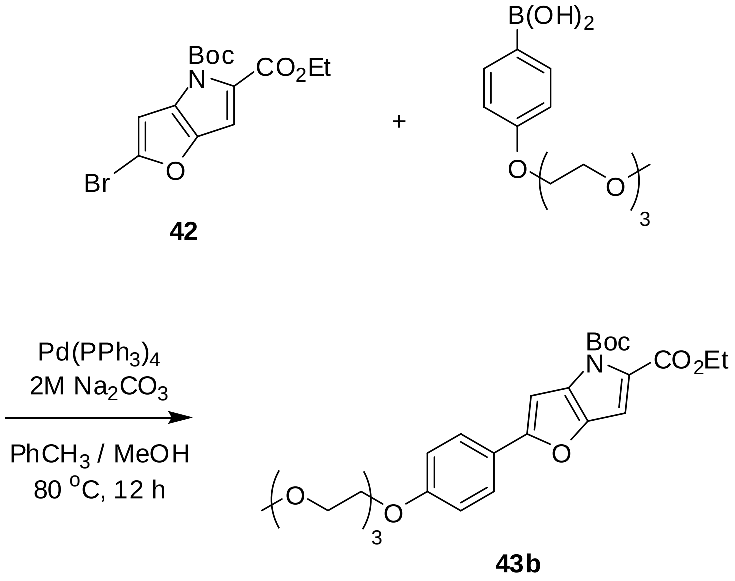
A mixture of pyrrole **43a** (211 mg, 0.41 mmol) and NaOH (163 mg, 4.1 mmol) in 3 mL ethylene glycol was heated to 150 °C in the microwave under N_2 and stirred for 30 min. The reaction mixture was cooled to 25 °C and diluted with water then extracted with EtOAc. The organic layer was washed with brine, dried over Na_2SO_4 and evaporated to dryness under reduced pressure. The residue was purified on neutralized silica gel (washed with 5 % Et_3N/Hexanes) eluted with 60 % EtOAc/Hexanes to afford the pure product **40a** (103 mg, 73 %) as brown oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.68 (br, 1H), 7.95 (dd, 1H, *J* = 7.8, 1.7 Hz), 7.34 (s, 1H), 7.16-7.10 (m, 1H), 7.04-6.99 (m, 1H), 6.91-6.87 (m, 1H), 6.75-6.73 (m, 1H), 6.17-6.15 (m, 1H), 4.22-4.18 (m, 2H), 3.90-3.87 (m, 2H), 3.75-3.64 (m, 6H), 3.57-3.54 (m, 2H), 3.38 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 153.8, 151.6, 147.3, 127.0, 126.7, 125.3, 121.7, 121.0, 120.4, 112.2, 100.1, 91.3, 71.7, 70.5, 70.4, 70.2, 69.6, 67.3, 58.7.



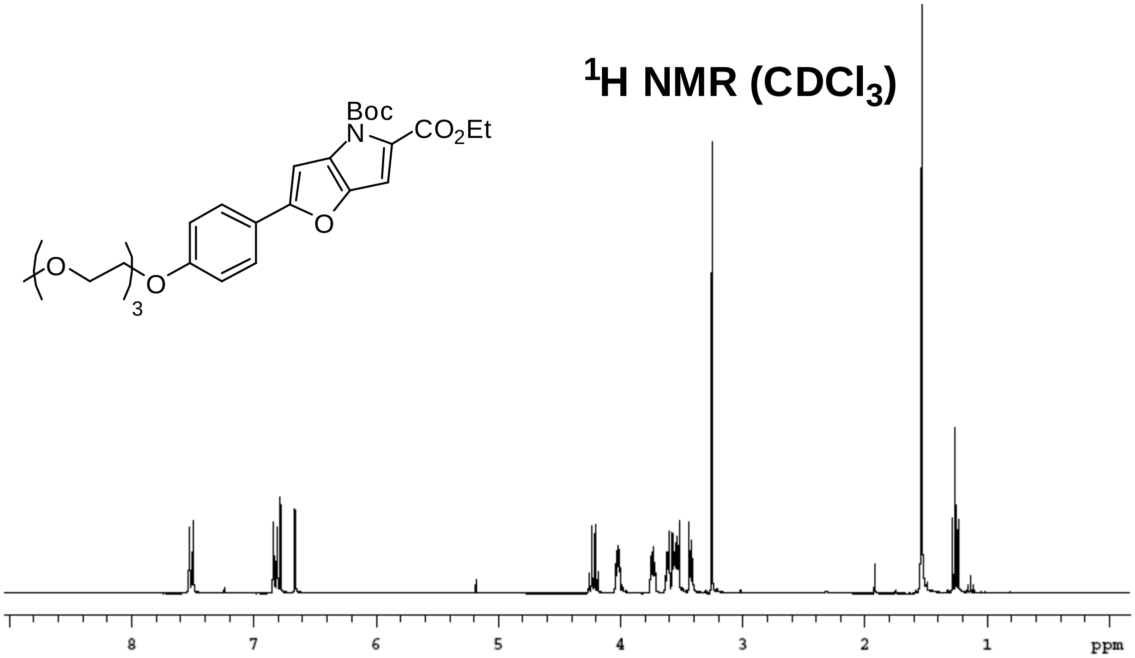
^{1}H NMR (CDCl$_3$)

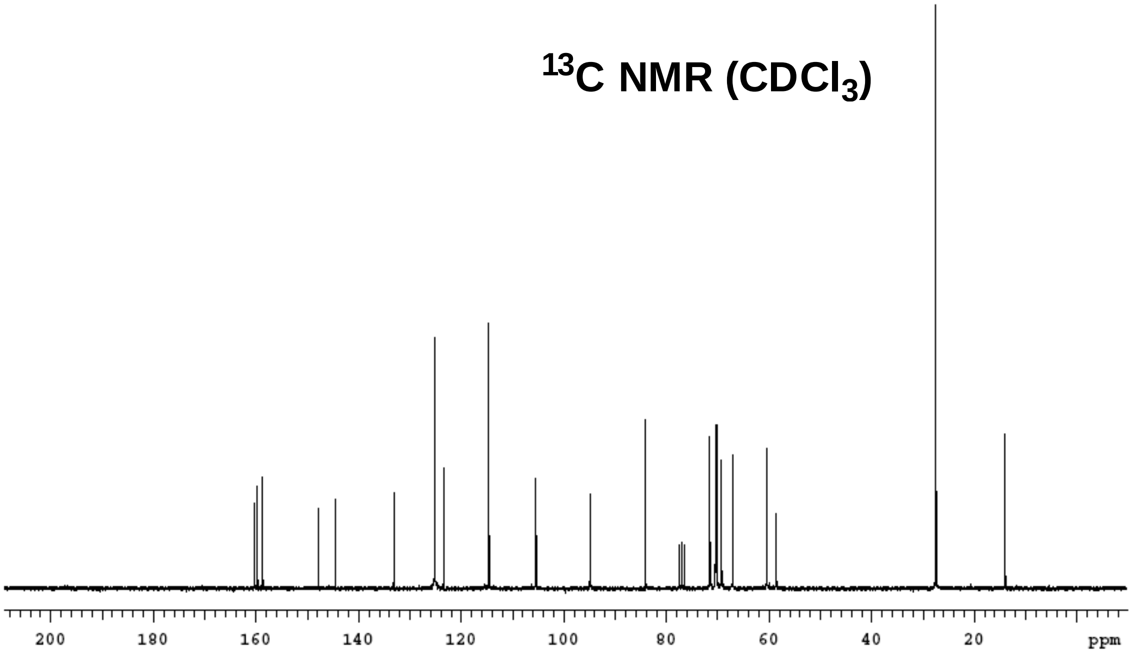


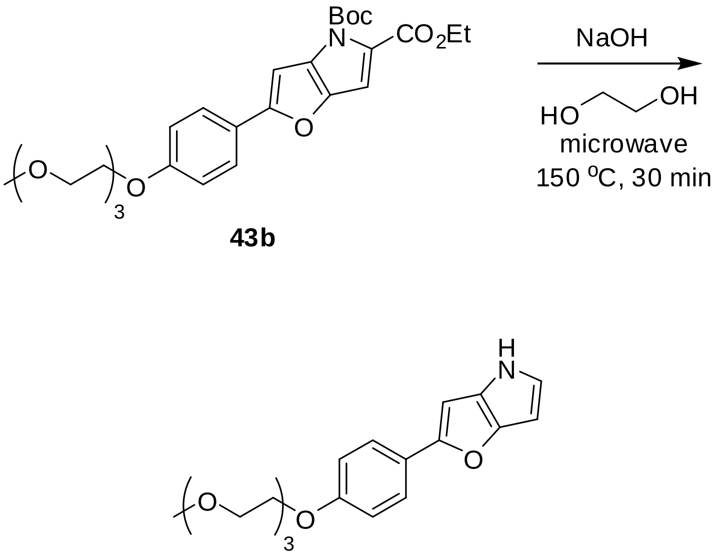
^{13}C NMR (CDCl$_3$)



To a solution of compound **42** (2.59 g, 7.23 mmol), arylboronic acid (3.07 g, 10.8 mmol) and Pd(PPh$_3$)$_4$ (162 mg, 0.14 mmol) in 40 mL PhCH$_3$ and 8 mL MeOH was added Na$_2$CO$_3$ (2M in water, 10.8 mL). The mixture was degassed and filled with N$_2$. After stirring at 80 °C for 24 h, the mixture was cooled to room temperature and extracted with EtOAc. The organic layer was washed with saturated Na$_2$CO$_3$ $_{(aq)}$, water, brine and dried over Na$_2$SO$_4$ then concentrated under reduced pressure. The residue was purified by flash chromatography (40 % EtOAc/Hexanes) to afford the product **43b** (4.42 g, 96 %) as a yellow oil. ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.51 (d, 2H, *J* = 9.0 Hz), 6.82 (d, 2H, *J* = 9.0 Hz), 6.78 (s, 1H), 6.66 (s, 1H), 4.22 (q, 2H, *J* = 7.0 Hz), 4.04-4.01 (m, 2H), 3.75-3.72 (m, 2H), 3.63-3.60 (m, 2H), 3.58-3.52 (m, 4H), 3.44-3.41 (m, 2H), 3.25 (s, 3H), 1.54 (s, 9H), 1.26 (t, 3H, *J* = 7.0 Hz); ^{13}C NMR (75 MHz, CDCl$_3$) δ 160.3, 159.7, 158.7, 147.8, 144.5, 133.1, 125.2, 125.1, 123.3, 114.6, 105.4, 94.9, 84.1, 71.5, 70.4, 70.2 (2 peaks: 70.24, 70.16), 69.3, 67.1, 60.4, 58.6, 27.5, 14.0.


^{1}H NMR (CDCl$_3$)
Boc
N
CO$_2$Et
O
O
O
3
8
7
6
5
4
3
2
1
ppm

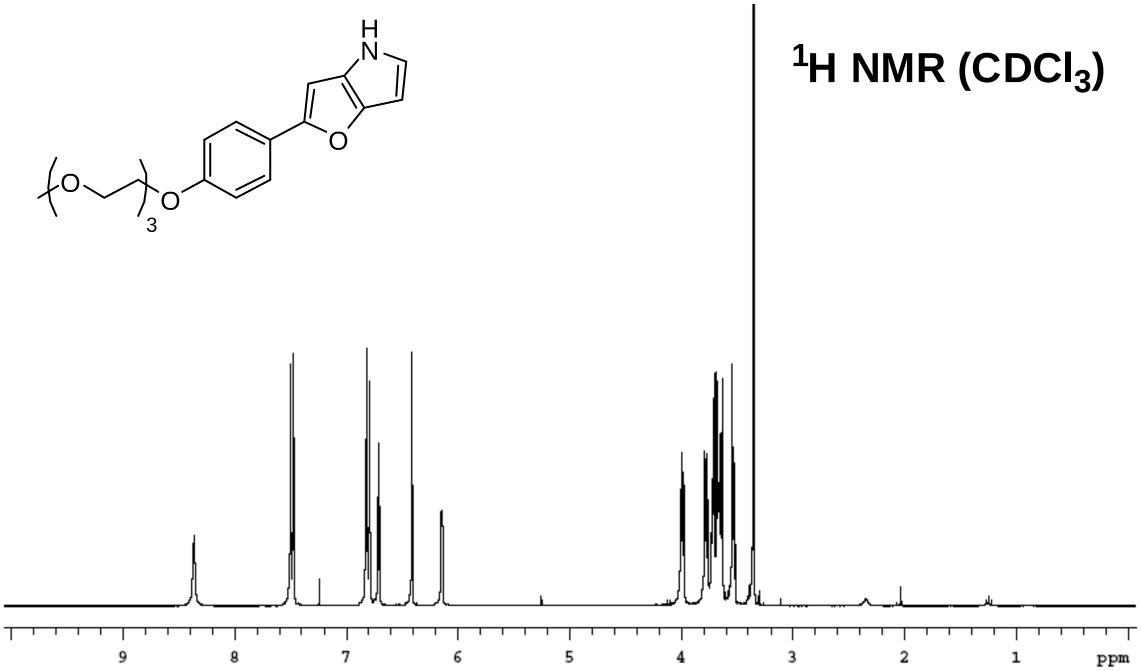

^{13}C NMR (CDCl$_3$)
200
180
160
140
120
100
80
60
40
20
ppm



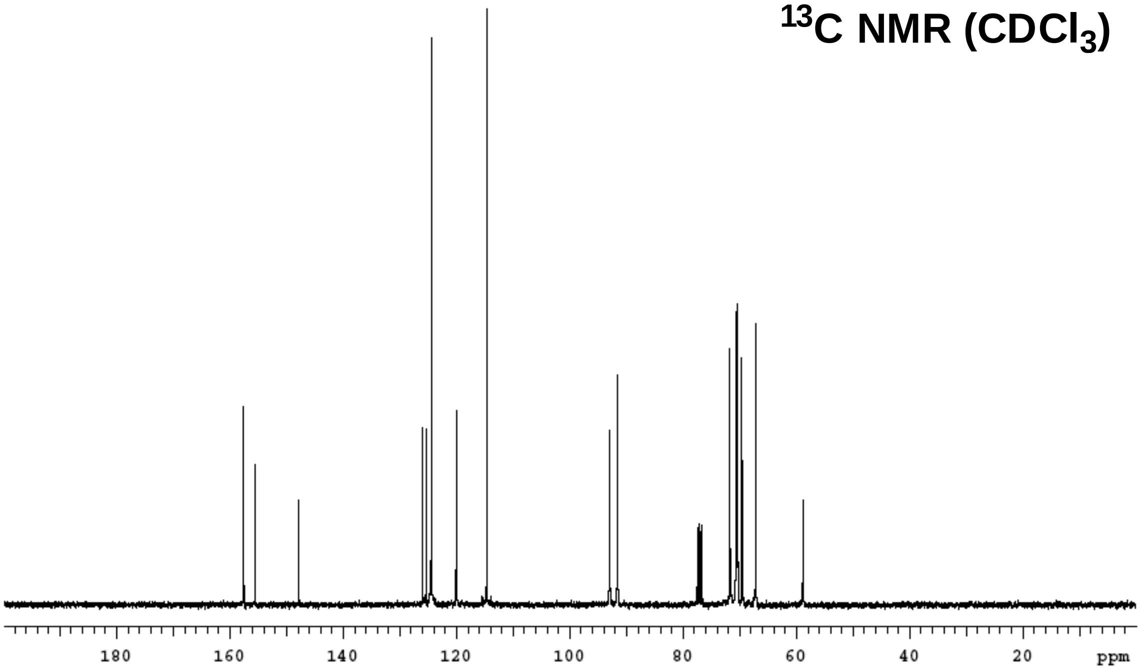
43b

40b

A mixture of pyrrole (2.38 g, 4.6 mmol) and NaOH (1.84 g, 46 mmol) in 50 mL ethylene glycol was heated to 150 °C in the microwave under N_2 and stirred for 30 min. The reaction mixture was cooled to 25 °C and diluted with water then extracted with EtOAc. The organic layer was washed with brine, dried over Na_2SO_4 and evaporated to dryness under reduced pressure. The residue was purified on neutralized silica gel (washed with 5 % Et_3N/Hexanes) eluted with 60 % EtOAc/Hexanes to afford the pure product (1.28 g, 81 %) as a gray solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.36 (br, 1H), 7.49 (d, 2H, *J* = 8.8 Hz), 6.80 (d, 2H, *J* = 8.8 Hz), 6.72-6.70 (m, 1H), 6.41 (s, 1H), 6.15-6.13 (m, 1H), 4.01-3.98 (m, 2H), 3.79-3.76 (m, 2H), 3.74-3.63 (m, 6H), 3.55-3.52 (m, 2H), 3.35 (m, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 157.5, 155.4, 147.8, 125.9, 125.3, 124.4, 120.0, 114.6, 92.9, 91.5, 71.7, 70.6, 70.4, 70.3, 69.6, 67.1, 58.8.

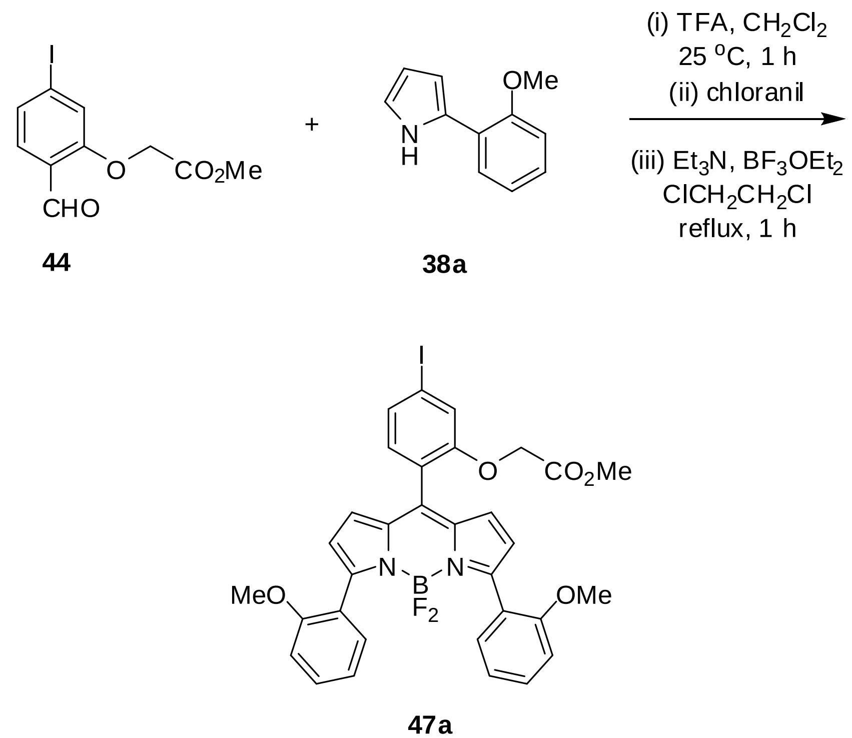


^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

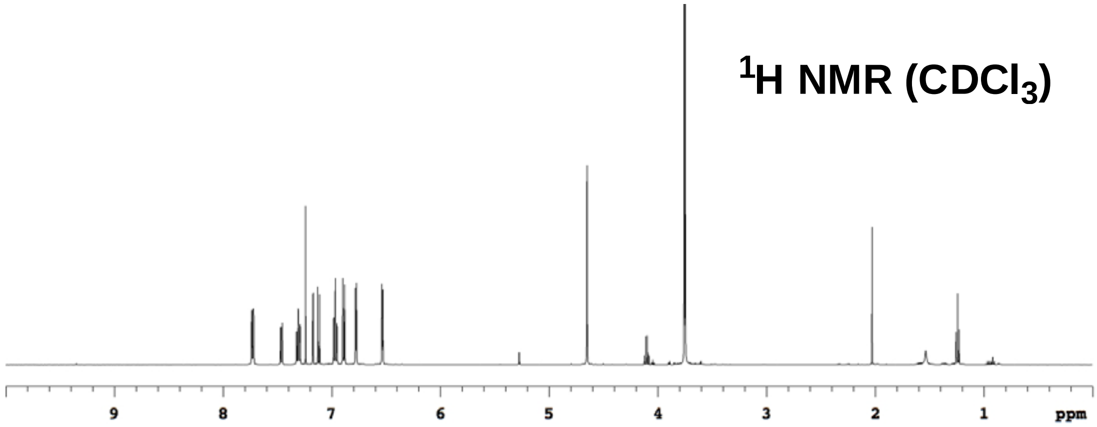
(3) Syntheses of BODIPYs 47

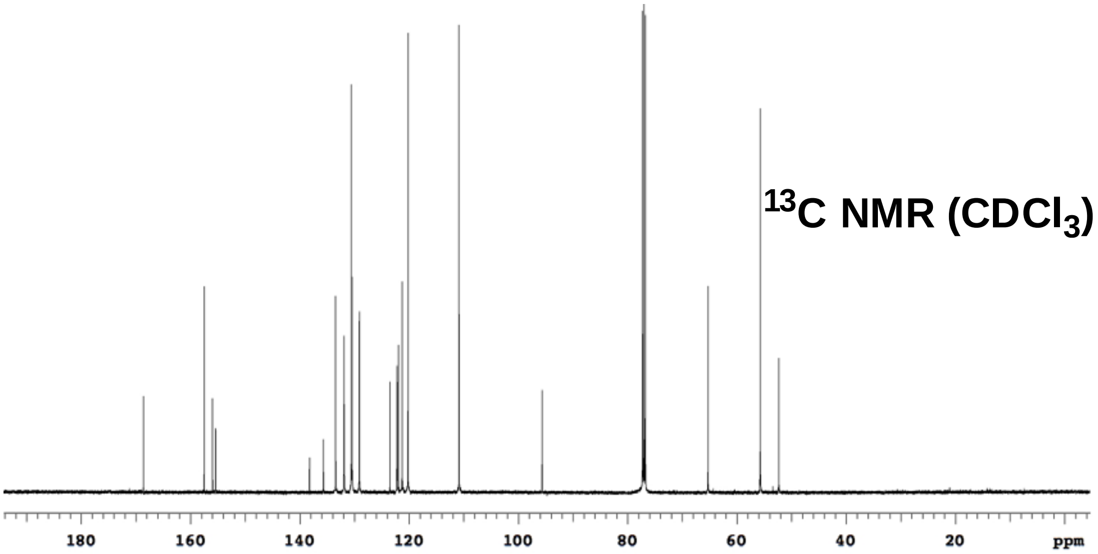


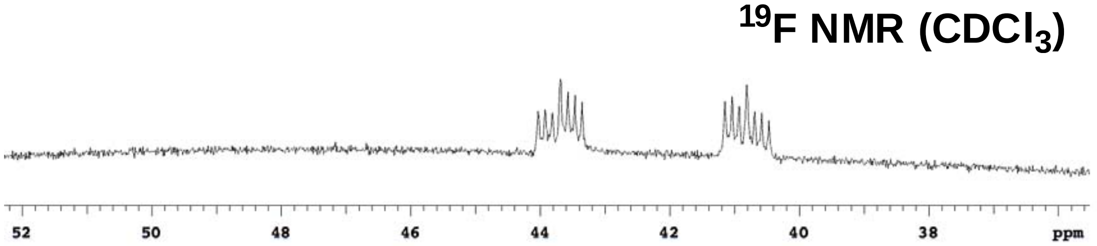
44 **38a**

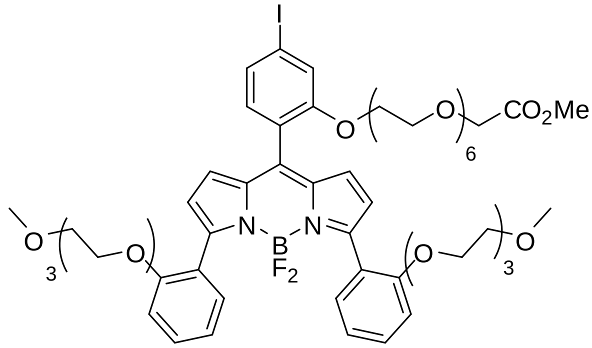
47a

Compound **44** (320 mg, 1 mmol) and pyrrole **38a** (346 mg, 2 mmol) were dissolved in 10 mL CH_2Cl_2 and one drop of TFA was added. The solution was stirred at room temperature until TLC showed the complete consumption of the aldehyde (around 1 h). Chloranil (246 mg, 1 mmol) was added and stirring was continued for 30 min. The mixture was passed through a short pad of basic alumina eluting with CH_2Cl_2. The solvents were removed under reduced pressure. The residue was dissolved in 20 mL 1,2-dichloroethane and Et_3N (0.42 mL, 3 mmol) was added followed by BF_3OEt_2 (0.63 mL, 5 mmol) after 10 min. The solution was refluxed for 1 h then cooled to room temperature. The reaction mixture was washed with water and brine. The organic layer was dried over Na_2SO_4 and evaporated to dryness. The residue was purified by flash chromatography (20 to 40 % EtOAc/Hexanes) to afford the product (480 mg, 69 %) as red solid. ^1H NMR (500 MHz, $CDCl_3$) δ 7.74-7.72 (m, 2H), 7.46 (dd, 1H, J = 7.9, 1.3 Hz), 7.32-7.29 (m, 2H), 7.17 (d, 1H, J = 1.3 Hz), 7.12 (d, 1H, J = 7.9 Hz), 6.98-6.95 (m, 2H), 6.89 (d, 2H, J = 8.3 Hz), 6.78 (d, 2H, J = 4.3 Hz), 6.53 (d, 2H, J = 4.3 Hz), 4.65 (s, 2H), 3.76 (s, 3H), 3.75 (s, 6H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 168.6, 157.5, 156.0, 155.4, 138.2, 135.7, 133.4, 131.9 (t, *J* = 5.3 Hz), 130.6, 130.4, 129.1, 123.5, 122.2, 121.9, 121.3, 120.2, 110.8, 95.7, 65.3, 55.7, 52.3; ^{19}F NMR (282 MHz, CDCl$_3$) δ 43.69 (dq, 1F, *J* = 97.2, 30.8 Hz), 40.81 (dq, 1F, *J* = 97.2, 30.8 Hz); MS (ESI) m/z calcd for (M+H)$^+$ $C_{32}H_{27}BF_2IN_2O_5$ 695.10; found 695.11.





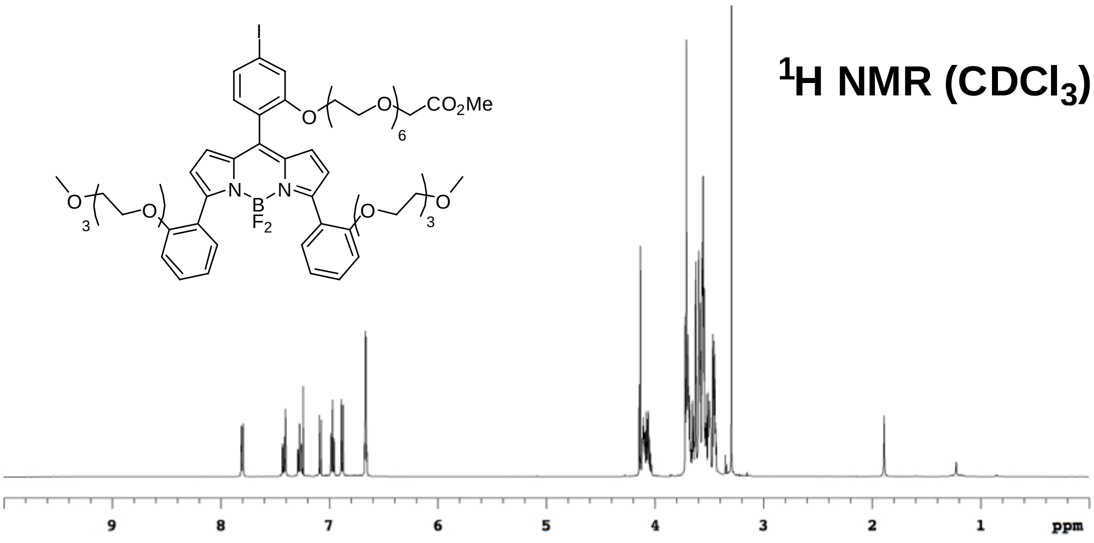




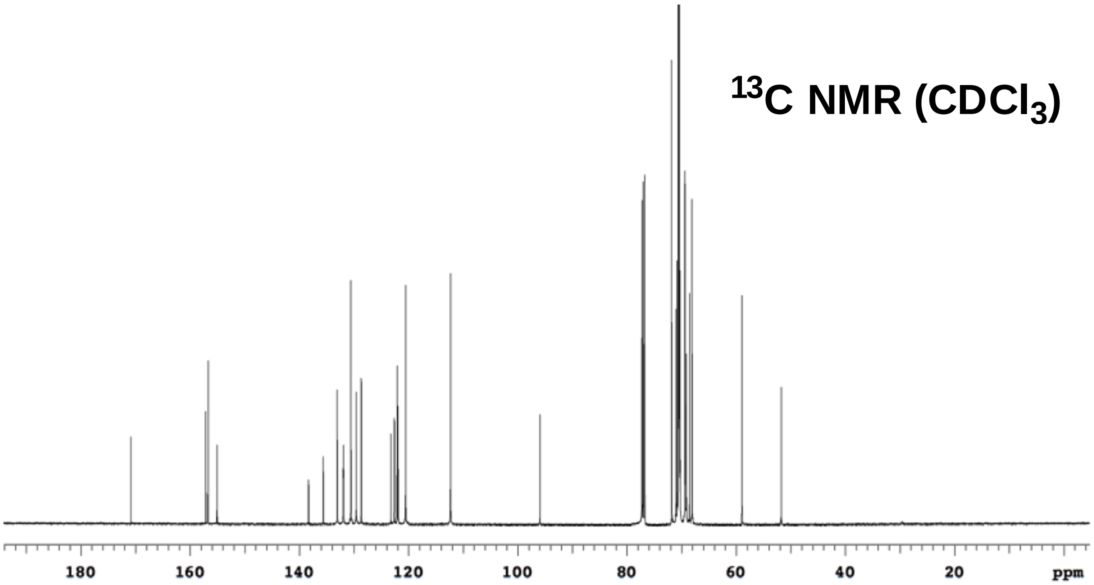
47b

Purple sticky solid (352 mg, 15 %), prepared from benzaldehyde **45** (1.22 g, 2.1 mmol) and pyrrole **38b** (1.31 g, 4.3 mmol) using similar conditions as described for **47a**.

^{1}H NMR (500 MHz, CDCl$_3$) δ 7.80 (d, 2H, *J* = 7.8 Hz), 7.43-7.40 (m, 2H), 7.29-7.26 (m, 2H), 7.08 (d, 1H, *J* = 7.8 Hz), 6.99-6.96 (m, 2H), 6.88 (d, 2H, *J* = 8.3 Hz), 6.68-6.65 (m, 4H), 4.15-4.03 (m, 8H), 3.72-3.44 (m, 45H), 3.30 (s, 6H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 170.8, 157.1, 156.7, 155.1, 138.3, 135.6, 133.0, 131.9 (t, *J* = 6.1 Hz), 130.5, 129.6, 128.7, 123.2, 122.6, 122.1, 120.5, 112.3, 95.9, 71.8, 71.0, 70.8 (2 peaks: 70.83, 70.82), 70.6, 70.5 (5 peaks: 70.55, 70.54, 70.52, 70.49, 70.46), 70.4 (2 peaks: 70.43, 70.38), 70.3 (2 peaks: 70.30, 70.26), 69.4, 69.3, 69.1, 68.5, 68.1, 58.9, 51.7, several peaks account for more than one carbon; ^{19}F NMR (282 MHz, CDCl$_3$) δ 49.00-48.31 (m, 1F), 38.57-37.91 (m, 1F); MS (MALDI) m/z calcd for (M+Na)$^+$ $C_{56}H_{74}BF_2IN_2NaO_{17}$ 1245.40; found 1245.34.

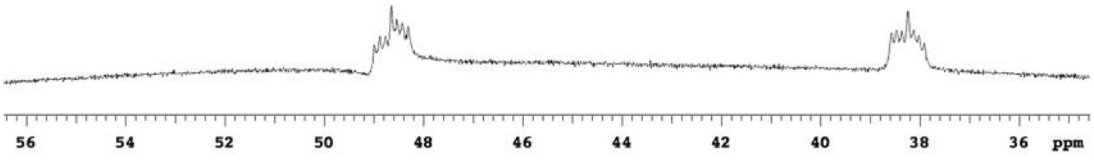


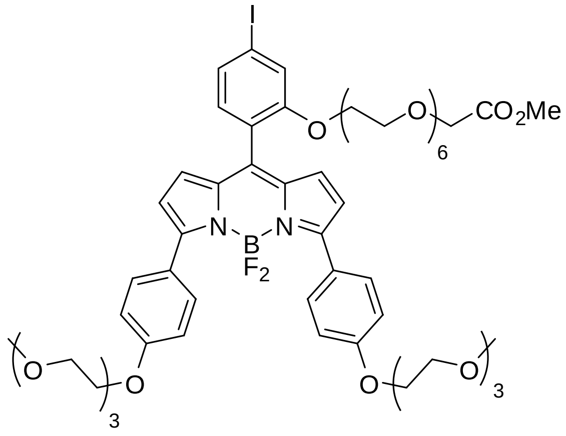
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

^{19}F NMR (CDCl$_3$)

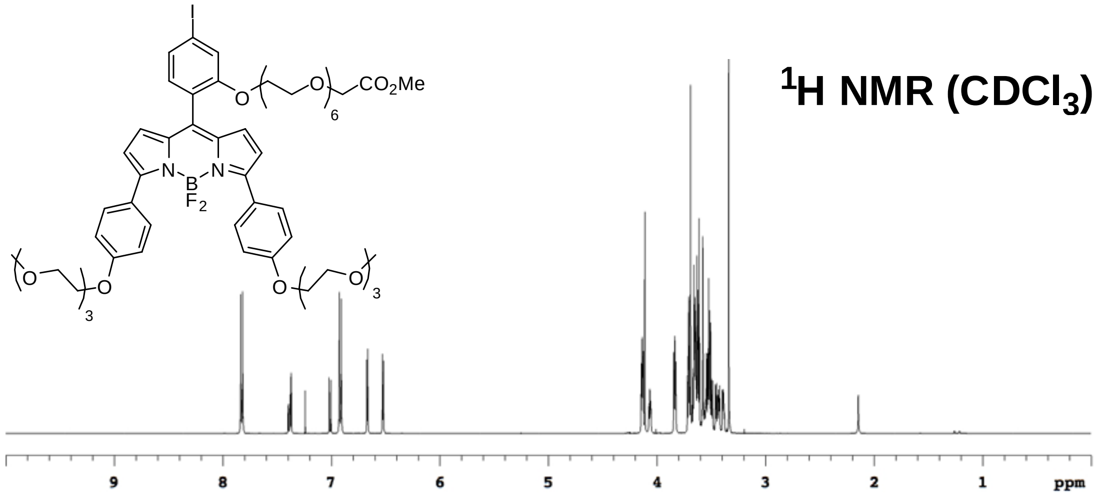




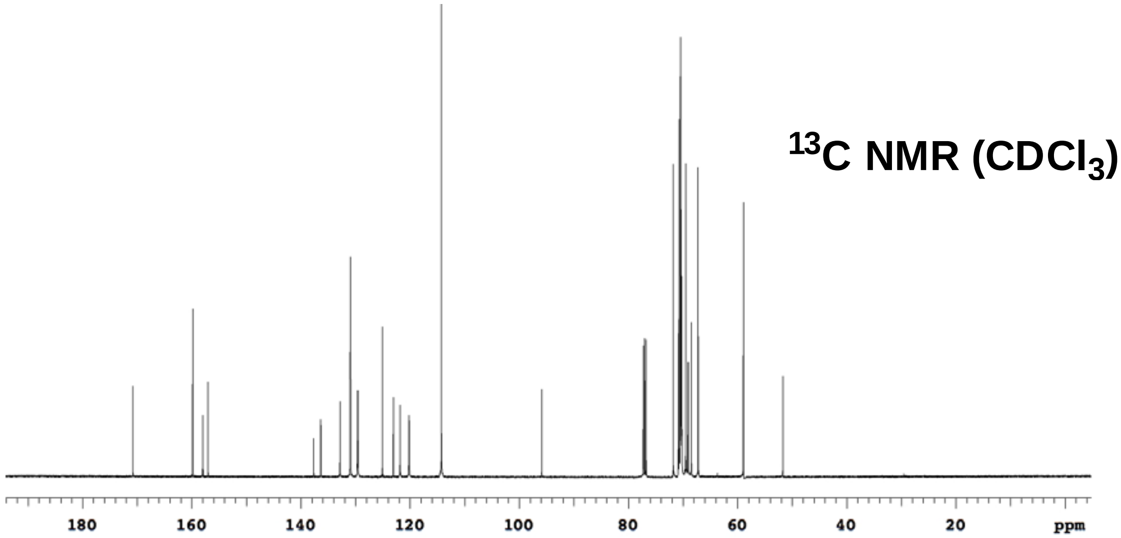
47c

Blue sticky solid (915 mg, 64 %), prepared from benzaldehyde **45** (683 mg, 1.17 mmol) and pyrrole **38c** (715 mg, 2.34 mmol) using similar conditions as described for **47a**.

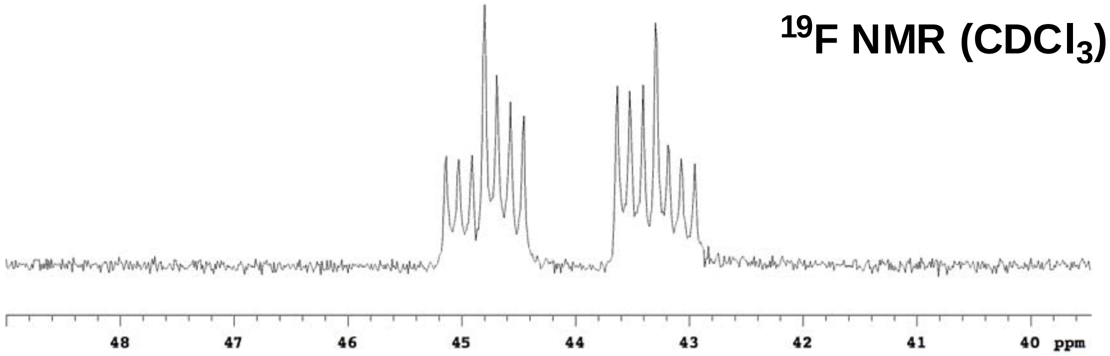
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.82 (d, 4H, *J* = 8.8 Hz), 7.40-7.37 (m, 2H), 7.01 (d, 1H, *J* = 7.8 Hz), 6.92 (d, 4H, *J* = 8.8 Hz), 6.67 (d, 2H, *J* = 4.4 Hz), 6.52 (d, 2H, *J* = 4.4 Hz), 4.15-4.13 (m, 4H), 4.11 (s, 2H), 4.07-4.06 (m, 2H), 3.84-3.83 (m, 4H), 3.72-3.38 (m, 41H), 3.34 (s, 6H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 170.8, 159.8, 158.0, 157.0, 137.6, 136.4, 132.8, 130.9 (t, *J* = 4.2 Hz), 129.6, 129.5, 125.1, 123.1, 121.8, 120.2, 114.3, 95.9, 71.8, 70.8, 70.7 (2 peaks: 70.73, 70.68), 70.5 (3 peaks: 70.51, 70.48, 70.45), 70.4 (5 peaks: 70.43, 70.40, 70.39. 70.37, 70.35), 70.3, 70.2 (2 peaks: 70.24, 70.20), 69.5, 69.1, 69.0, 68.4, 67.2, 58.9, 51.7, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 44.80 (dq, 1F, *J* = 94.8, 32.1 Hz), 43.30 (dq, 1F, *J* = 94.8, 32.1 Hz); MS (MALDI) m/z calcd for M$^+$ $C_{56}H_{74}BF_2IN_2O_{17}$ 1222.41; found 1222.36.



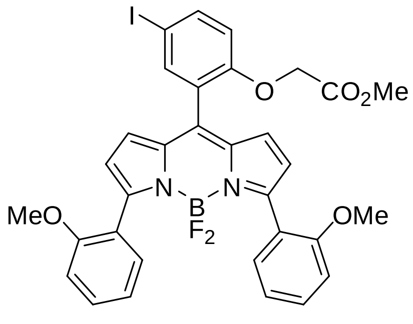
^1H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)



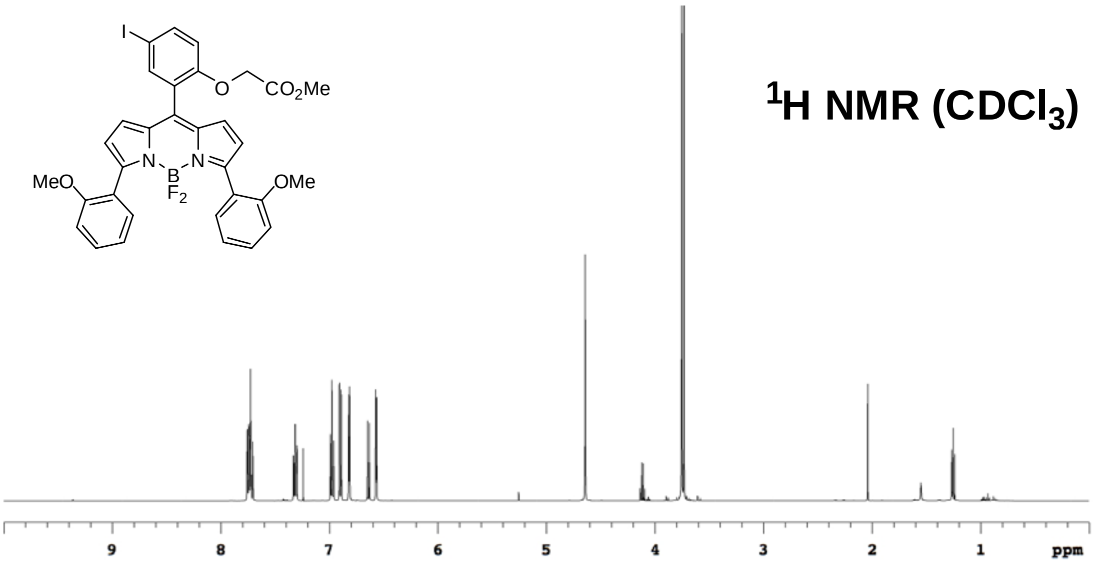
^{19}F NMR ($CDCl_3$)



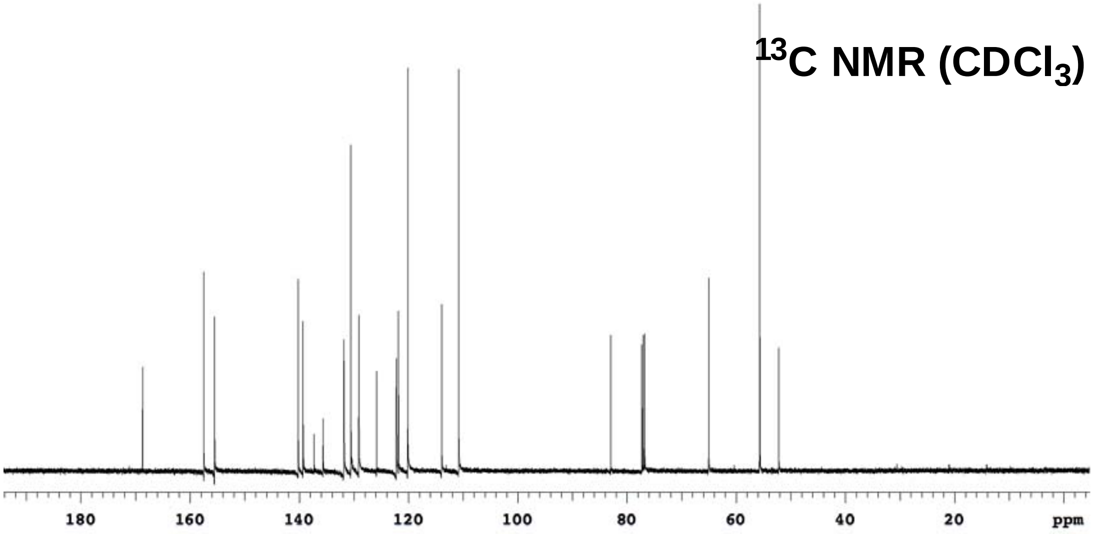
47d

Red solid (345 mg, 50 %), prepared from benzaldehyde **46** (320 mg, 1.0 mmol) and pyrrole **38a** (346 mg, 2.0 mmol) using similar conditions as described for **47a**.

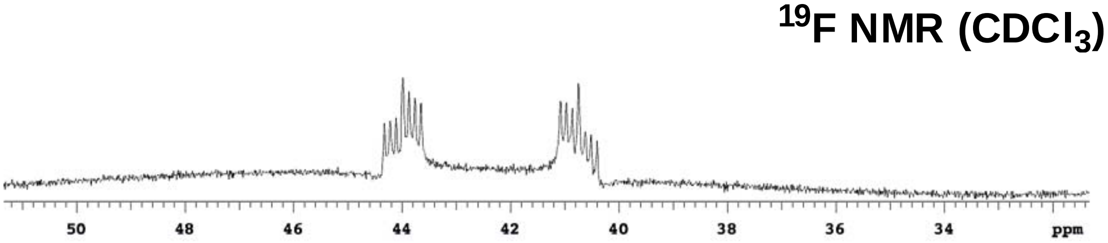
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.76-7.70 (m, 4H), 7.33-7.30 (m, 2H), 6.99-6.96 (m, 2H), 6.90 (d, 2H, *J* = 8.3 Hz), 6.82 (d, 2H, *J* = 4.3 Hz), 6.64 (d, 1H, *J* = 8.5 Hz), 6.57 (d, 2H, *J* = 4.3 Hz), 4.64 (s, 2H), 3.75 (s, 6H), 3.73 (s, 3H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 169.0, 157.5, 155.5, 140.2, 139.3, 137.3, 135.6, 131.8 (t, *J* = 5.0 Hz), 130.6, 129.1, 125.9, 122.2, 121.8, 120.1, 113.9, 110.8, 83.0, 65.0, 55.7, 52.2, several peaks account for more than one carbon; ^{19}F NMR (282 MHz, CDCl$_3$) δ 43.98 (dq, 1F, *J* = 97.2, 30.6 Hz), 40.75 (dq, 1F, *J* = 97.2, 30.6 Hz); MS (MALDI) m/z calcd for M$^+$ $C_{32}H_{26}BF_2IN_2O_5$ 694.09; found 694.30.



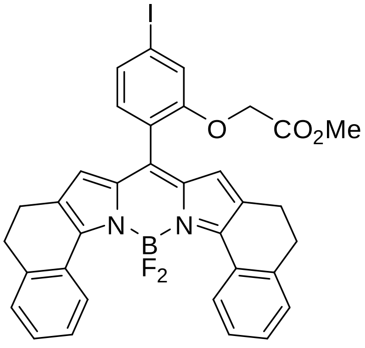
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)



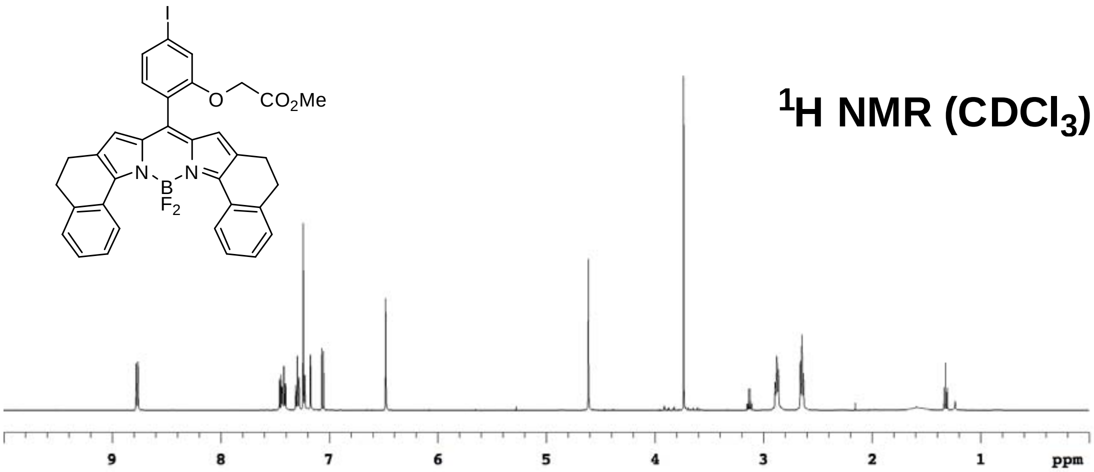
^{19}F NMR (CDCl$_3$)



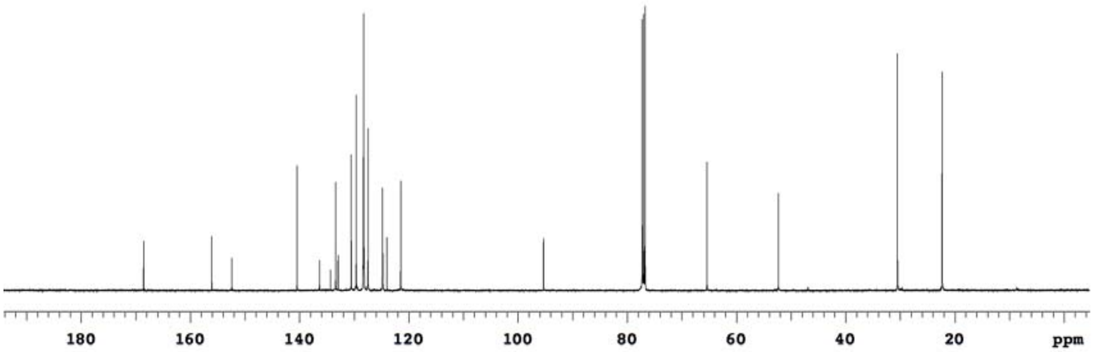
47e

Blue solid (604 mg, 88 %), prepared from benzaldehyde **44** (320 mg, 1.0 mmol) and pyrrole **39a** (338 mg, 2.0 mmol) using similar conditions as described for **47a**.

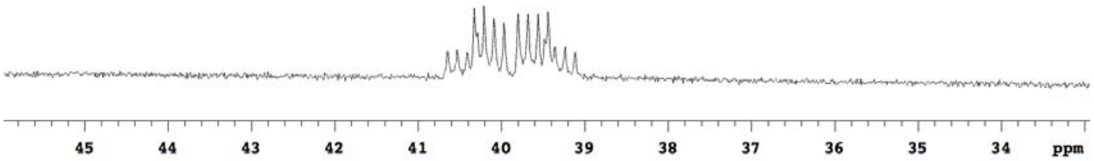
^{1}H NMR (500 MHz, CDCl$_3$) δ 8.77 (d, 2H, *J* = 8.0 Hz), 7.45 (dd, 1H, *J* = 7.8, 1.5 Hz), 7.44-7.40 (m, 2H), 7.31-7.28 (m, 2H), 7.24-7.23 (m, 2H), 7.17 (d, 1H, *J* = 1.5 Hz), 7.06 (d, 1H, *J* = 7.8 Hz), 6.48 (s, 2H), 4.61 (s. 2H), 3.74 (s, 3H), 2.88 (t, 4H, *J* = 7.2 Hz), 2.65 (t, 4H, *J* = 7.2 Hz); ^{13}C NMR (125 MHz, CDCl$_3$) δ 168.5, 156.1, 152.4, 140.5, 136.3, 134.3, 133.4, 132.9, 130.5, 129.6, 128.3, 128.2 (2 peaks: 128.245, 128.172), 127.5, 124.8, 124.0, 121.5, 95.3, 65.4, 52.3, 30.5, 22.3; ^{19}F NMR (282 MHz, CDCl$_3$) δ 40.65-39.12 (m); MS (MALDI) m/z calcd for M$^+$ $C_{34}H_{26}BF_2IN_2O_3$ 686.10; found 686.02.



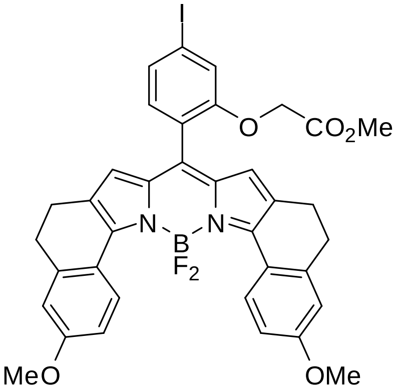
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)



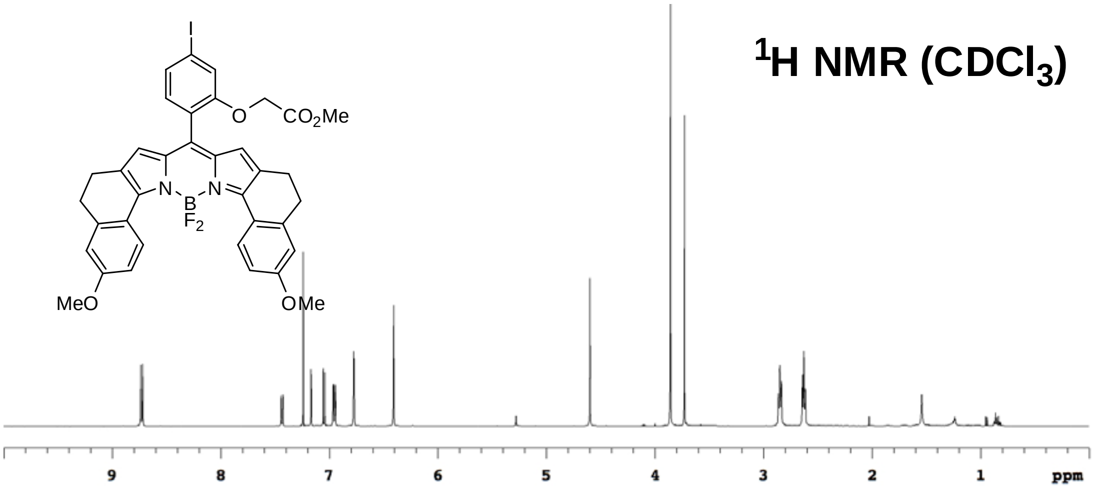
^{19}F NMR (CDCl$_3$)



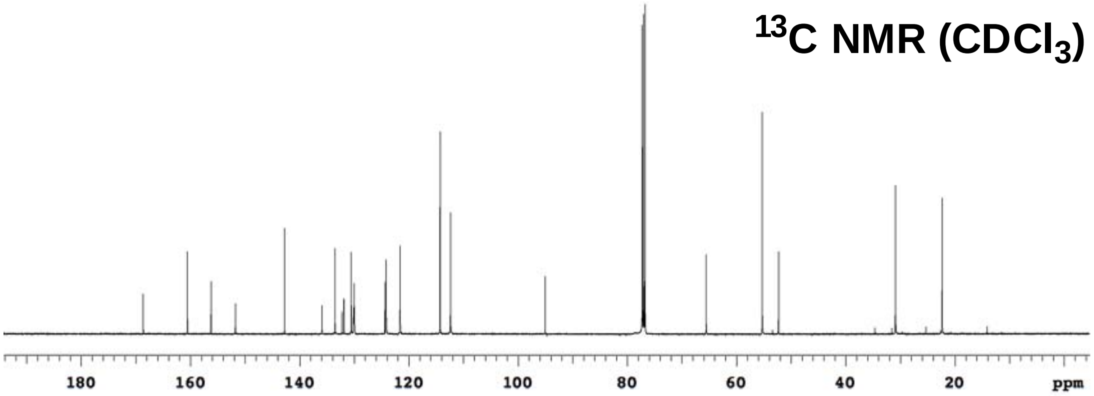
47f

Green solid (330 mg, 80 %), prepared from benzaldehyde **44** (160 mg, 0.5 mmol) and pyrrole **39b** (205 mg, 1.0 mmol) using similar conditions as described for **47a**.

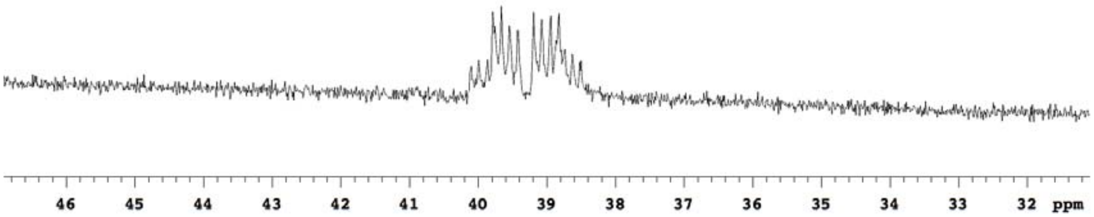
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.73 (d, 2H, *J* = 8.9 Hz), 7.43 (dd, 1H, *J* = 7.9, 1.3 Hz), 7.17 (d, 1H, *J* = 1.3 Hz), 7.05 (d, 1H, *J* = 7.9 Hz), 6.95 (dd, 2H, *J* = 8.9, 2.7 Hz), 6.77 (d, 2H, *J* = 2.7 Hz), 6.41 (s, 2H), 4.60 (s, 2H), 3.86 (s, 6H), 3.73 (s, 3H), 2.85 (t, 4H, *J* = 7.1 Hz), 2.63 (t, 4H, *J* = 7.1 Hz); ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.6, 160.5, 156.2, 151.7, 142.8, 135.9, 133.5, 132.2, 131.9, 130.5, 130.0 (t, *J* = 10.9 Hz), 124.4, 124.2, 121.6 (2 peaks: 121.61, 121.58), 114.3, 112.4, 95.0, 65.5, 55.3, 52.3, 30.9, 22.3; ^{19}F NMR (282 MHz, $CDCl_3$) δ 40.11-39.42 (m, 1F), 39.19-38.50 (m, 1F); MS (MALDI) m/z calcd for M^+ $C_{36}H_{30}BF_2IN_2O_5$ 746.13; found 746.03.



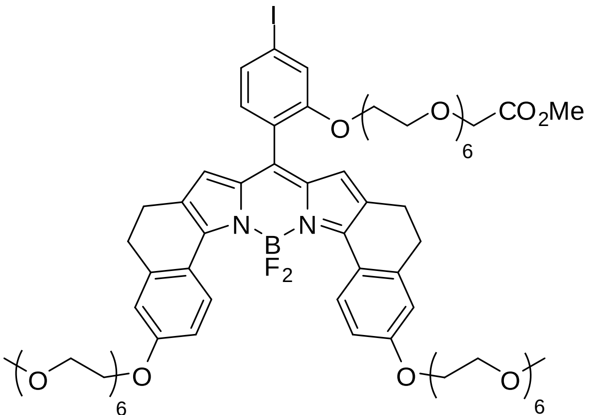
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)



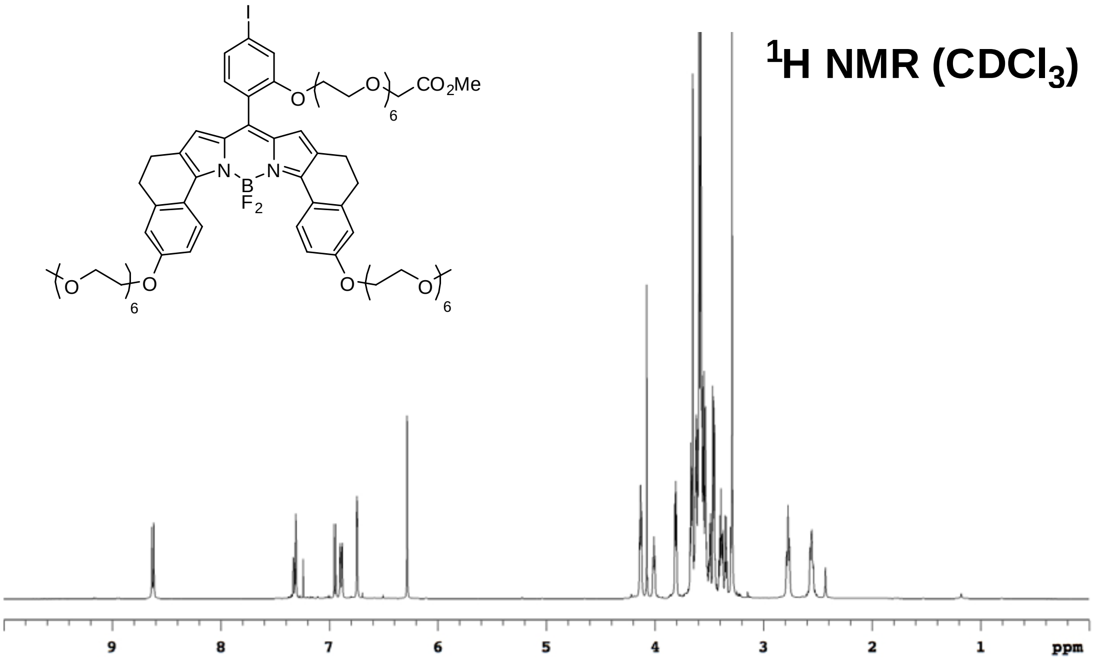
^{19}F NMR (CDCl$_3$)

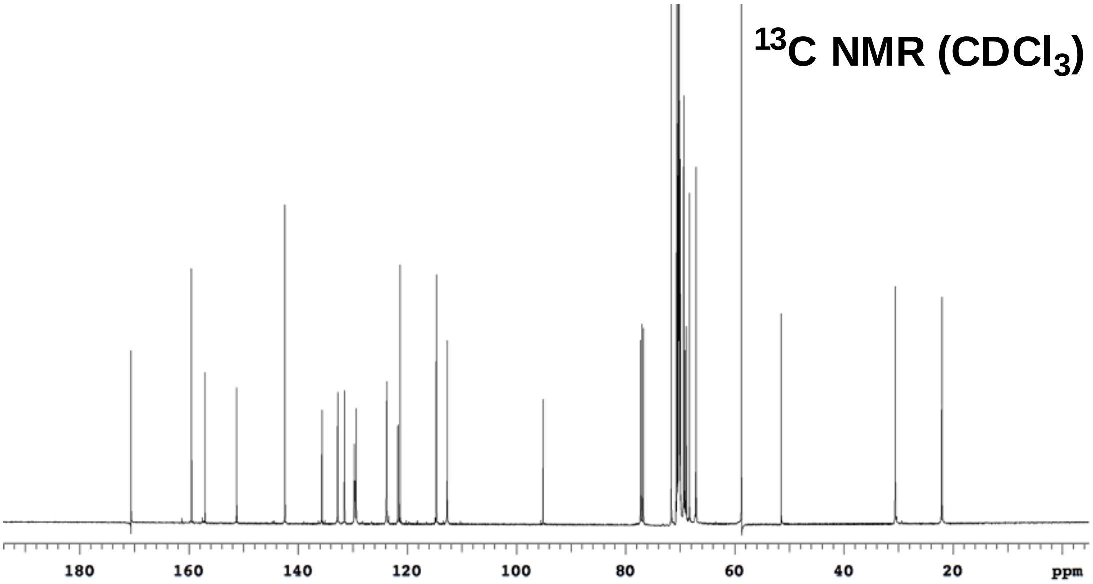


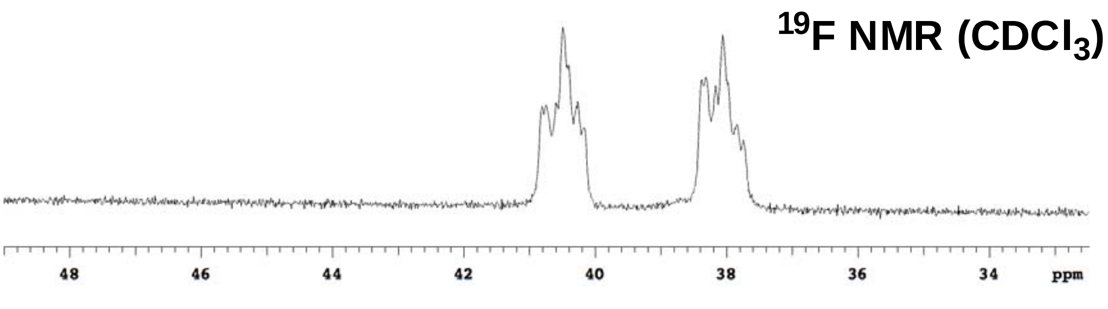
47g

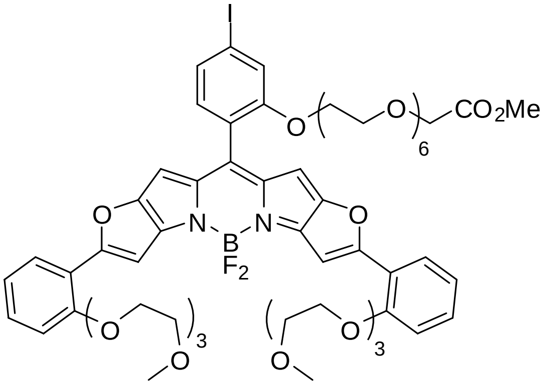
Green sticky solid (984 mg, 47 %), prepared from benzaldehyde **45** (800 mg, 1.4 mmol) and pyrrole **39c** (1.3 g, 2.8 mmol) using similar conditions as described for **47a**.

^{1}H NMR (500 MHz, CDCl$_3$) δ 8.63 (d, 2H, *J* = 9.0 Hz), 7.33-7.30 (m, 2H), 6.95 (d, 1H, *J* = 7.8 Hz), 6.89 (dd, 2H, *J* = 9.0, 2.8 Hz), 6.74 (d, 2H, *J* = 2.8 Hz), 6.28 (s, 2H), 4.14-4.12 (m, 4H), 4.08 (s, 2H), 4.02-4.00 (m, 2H), 3.82-3.80 (m, 4H), 3.67-3.34 (m, 65H), 3.29 (s, 6H), 2.79-2.76 (m, 4H), 2.57-2.54 (m, 4H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 170.6, 159.5, 157.0, 151.2, 142.4, 135.6, 132.7 (2 peaks: 132.74, 132.72), 131.5, 130.0 (t, *J* = 10.6 Hz), 129.3, 123.8, 123.7, 121.7, 121.3, 114.7, 112.7, 95.1, 71.6, 70.7, 70.6 (2 peak: 70.59, 70.56), 70.3 (4 peaks: 70.35, 70.32, 70.28, 70.26), 70.2 (2 peaks: 70.21, 70.17), 70.1, 70.0 (2 peaks: 69.99, 69.95), 69.3, 69.1, 68.9, 68.3, 67.1, 58.8, 51.5, 30.6, 22.1, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 40.80-40.16 (m, 1F), 38.38-37.75 (m, 1F); MS (MALDI) m/z calcd for M$^+$ $C_{72}H_{102}BF_2IN_2O_{23}$ 1538.60; found 1538.53.


I
O
O
6
CO$_2$Me
N
B
N
F$_2$
O
O
6
O
O
6
^{1}H NMR (CDCl$_3$)
9
8
7
6
5
4
3
2
1
ppm


^{13}C NMR (CDCl$_3$)
180
160
140
120
100
80
60
40
20
ppm

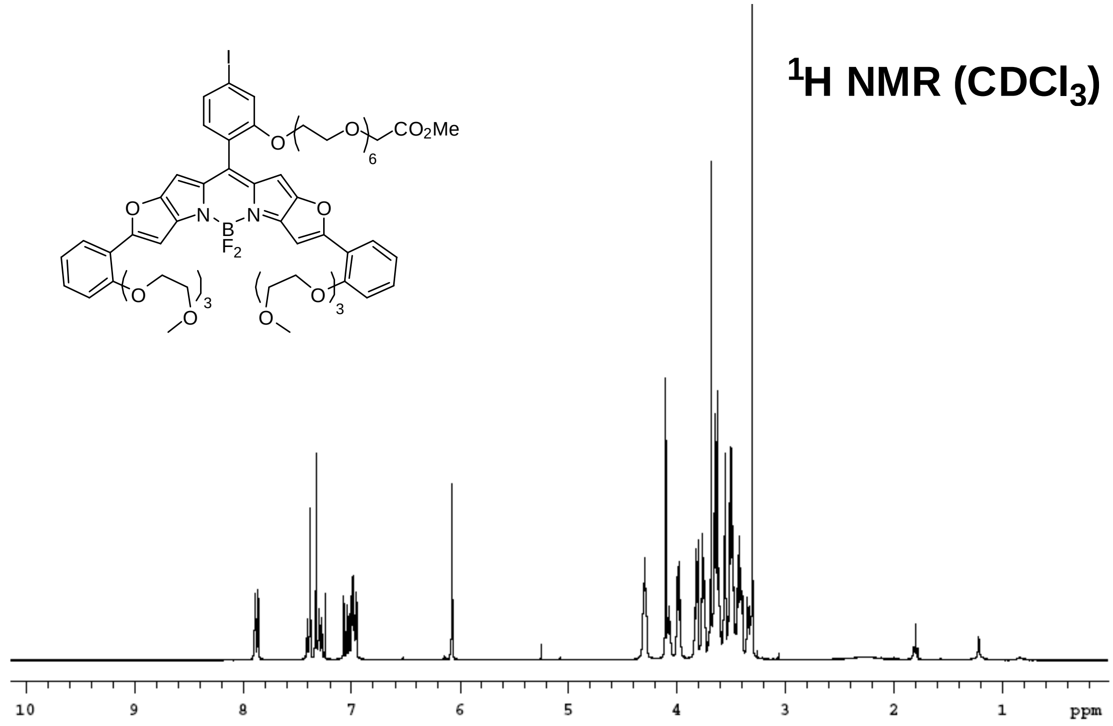

^{19}F NMR (CDCl$_3$)
48
46
44
42
40
38
36
34
ppm



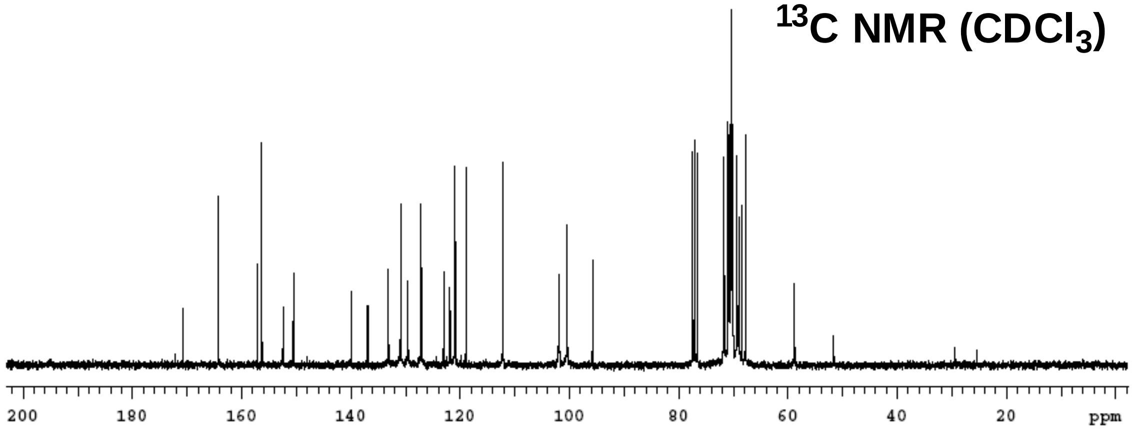
47h

Green sticky solid (391 mg, 75 %), prepared from benzaldehyde **45** (233 mg, 0.4 mmol) and pyrrole **40a** (294 mg, 0.85 mmol) using similar conditions as described for **47a**.

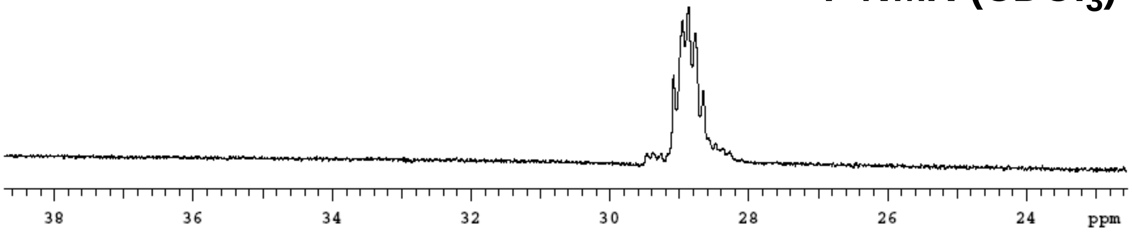
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.87 (dd, 2H, *J* = 7.8, 1.7 Hz), 7.41-7.38 (m, 2H), 7.33-7.27 (m, 4H), 7.06 (d, 1H, *J* = 7.8 Hz), 7.02-6.95 (m, 4H), 6.07 (s, 2H), 4.31-4.28 (m, 4H), 4.10 (s, 2H), 4.09-4.06 (m, 2H), 4.00-3.96 (m, 4H), 3.84-3.81 (m, 4H), 3.77-3.74 (m, 4H), 3.68 (s, 3H), 3.66-3.32 (m, 30H), 3.30 (s, 6H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.8, 164.3, 157.1, 156.3, 152.4, 150.5, 139.9, 136.9, 133.1, 130.9, 129.6, 127.1, 123.0, 121.9, 120.9, 119.0, 112.2, 101.9, 100.4, 95.7, 71.7, 70.9, 70.8, 70.7 (2 peaks: 70.72, 70.68), 70.4, 70.3 (2 peaks: 70.32, 70.29), 70.2, 70.1, 69.4, 69.0, 68.4, 67.8, 58.8, 51.6, several peaks account for more than one carbon; ^{19}F NMR (282 MHz, $CDCl_3$) δ 29.09-28.65 (m); MS (MALDI) m/z calcd for M^+ $C_{60}H_{74}BF_2IN_2O_{19}$ 1302.40; found 1302.16.



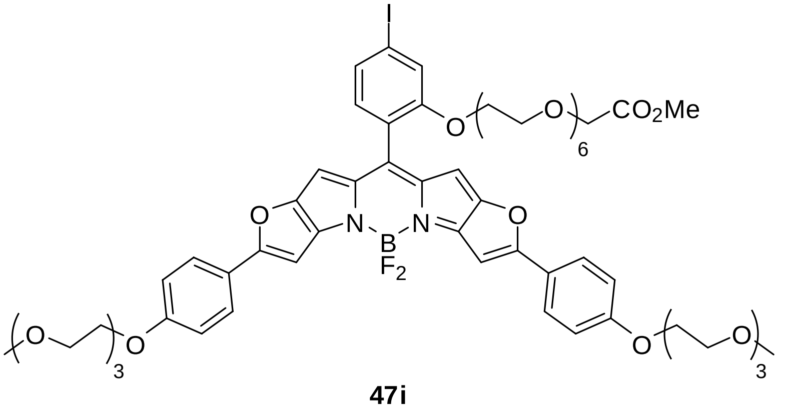
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)



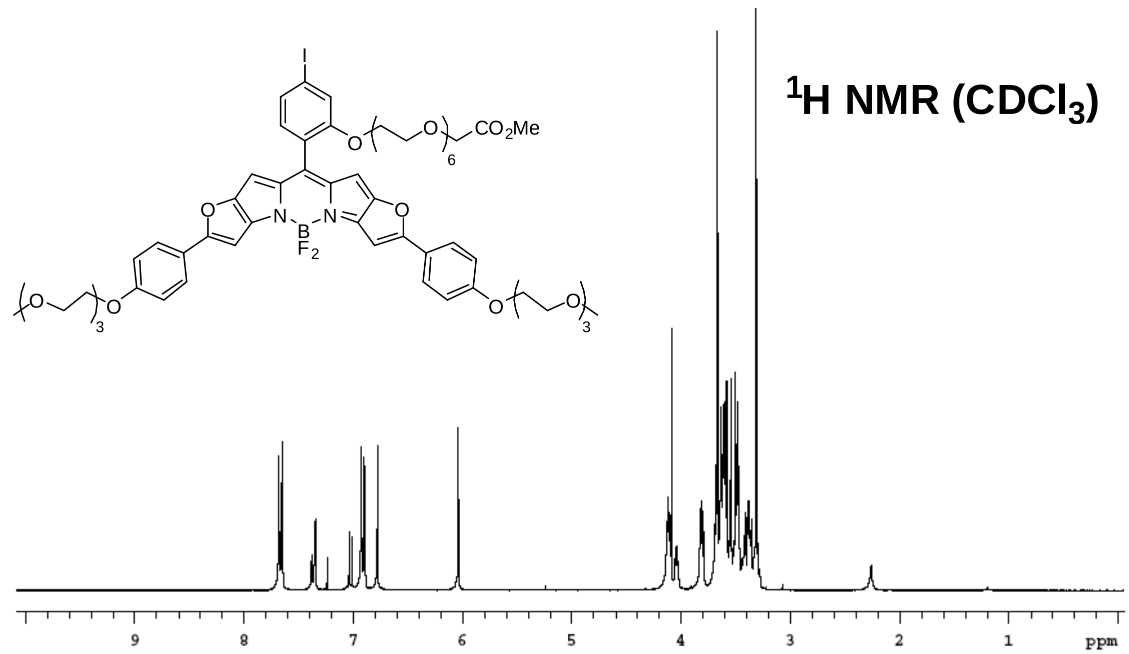
^{19}F NMR (CDCl$_3$)

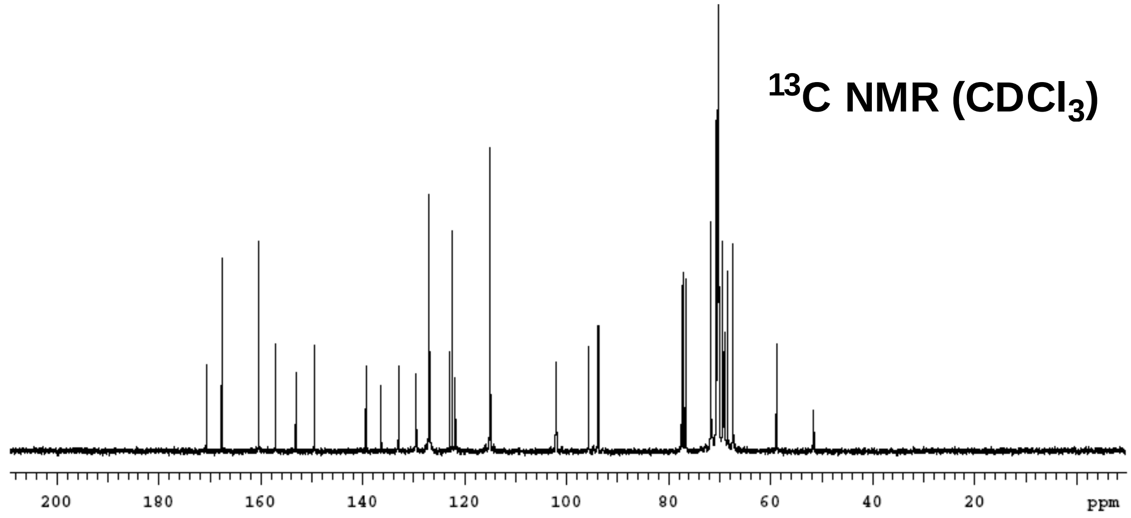


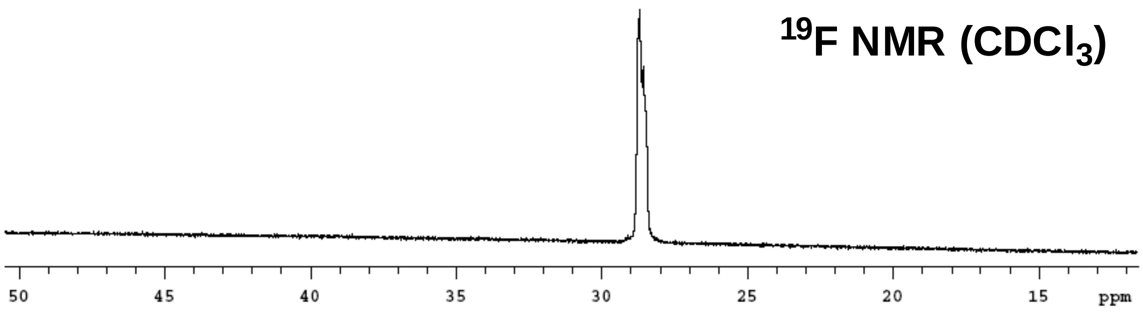
47i

Green sticky solid (1.18 g, 41 %), prepared from benzaldehyde **45** (1.3 g, 2.23 mmol) and pyrrole **40b** (1.58 g, 4.57 mmol) using similar conditions as described for **47a**.

^{1}H NMR (300 MHz, CDCl$_3$) δ 7.67 (d, 4H, *J* = 8.8 Hz), 7.38-7.34 (m, 2H), 7.02 (d, 1H, *J* = 7.8 Hz), 6.91 (d, 4H, *J* = 8.8 Hz), 6.78 (s, 2H), 6.04 (s, 2H), 4.13-4.10 (m, 4H), 4.09 (s, 2H), 4.05-4.02 (m, 2H), 3.83-3.79 (m, 4H), 3.70-3.36 (m, 41H), 3.31 (s, 6H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 170.7, 167.7, 160.4, 157.0, 153.1, 149.5, 139.3, 136.4, 133.0, 129.5, 126.9, 122.9, 122.4, 121.8, 115.0, 102.1, 95.7, 93.8, 71.7, 70.7 (2 peaks: 70.72, 70.65), 70.4 (2 peaks: 70.42, 70.35), 70.3, 70.2, 70.1 (2 peaks: 70.11, 70.08), 69.4, 69.0, 68.9, 68.4, 67.4, 58.8, 51.6, several peaks account for more than one carbon; ^{19}F NMR (282 MHz, CDCl$_3$) δ 28.75-28.48 (m); MS (MALDI) m/z calcd for (M-F)$^+$ $C_{60}H_{74}BFIN_2O_{19}$ 1283.40, found 1283.30; calcd for M$^+$ $C_{60}H_{74}BF_2IN_2O_{19}$ 1302.40, found 1302.29; calcd for (M+Na)$^+$ $C_{60}H_{74}BF_2INaN_2O_{19}$ 1325.39, found 1325.28.

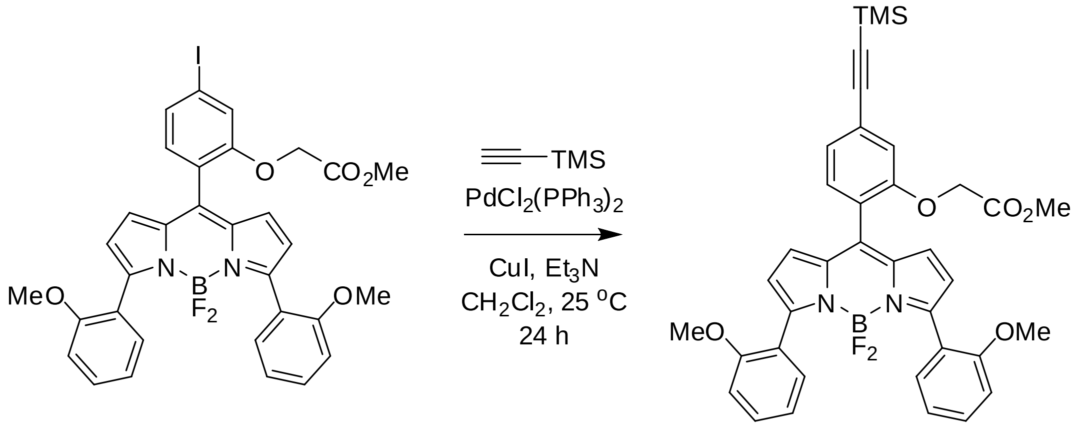

1H NMR (CDCl3)
I
O
CO2Me
6
O
N
B
F2
N
O
O
3
O
O
3
9
8
7
6
5
4
3
2
1
ppm


13C NMR (CDCl3)
200
180
160
140
120
100
80
60
40
20
ppm


19F NMR (CDCl3)
50
45
40
35
30
25
20
15
ppm

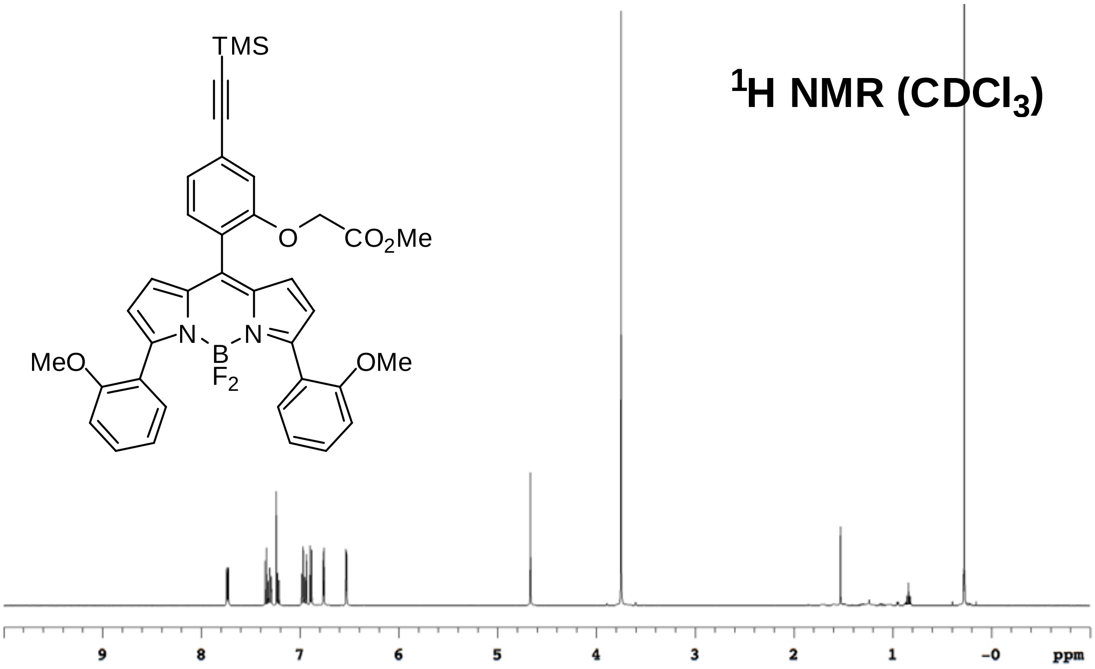
(4) Syntheses of BODIPYs 48

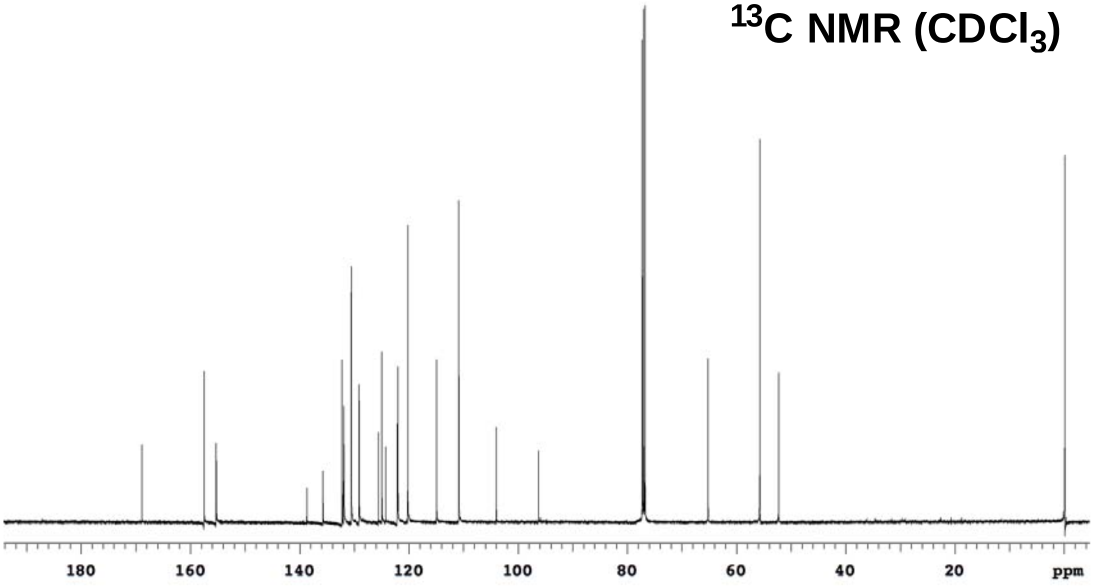
Typical procedures:

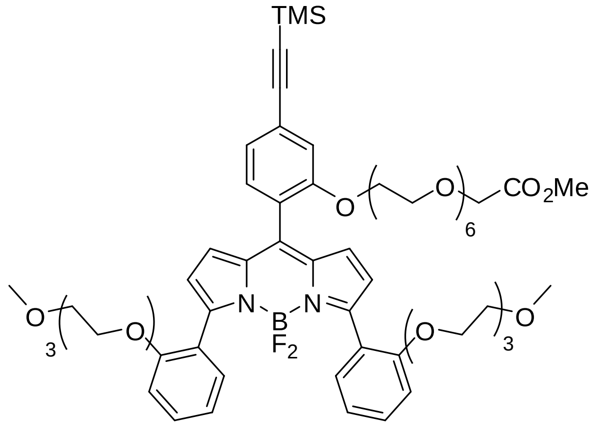


47a **48a**

BODIPY **47a** (208 mg, 0.3 mmol), trimethylsilylacetylene (0.4 mL, 3.0 mmol), $PdCl_2(PPh_3)_2$ (11 mg, 0.02 mmol), CuI (3 mg, 0.02 mmol) were dissolved in 20 mL CH_2Cl_2 then Et_3N (0.4 mL, 3.0 mmol) was added. After stirring at room temperature for 24 h, the reaction mixture was evaporated to dryness under reduced pressure. The residue was purified by flash chromatography (20 to 30 % EtOAc/Hexanes) to afford the product **48a** (184 mg, 92 %) as red solid. ^1H NMR (500 MHz, $CDCl_3$) δ 7.74-7.72 (m, 2H), 7.34 (d, 1H, J = 7.8 Hz), 7.32-7.29 (m, 2H), 7.22 (dd, 1H, J = 7.8, 1.3 Hz), 6.98-6.95 (m, 2H), 6.93 (d, 1H, J = 1.3 Hz), 6.89 (d, 2H, J = 8.3 Hz), 6.76 (d, 2H, J = 4.3 Hz), 6.53 (d, 2H, J = 4.3 Hz), 4.67 (s, 2H), 3.75 (s, 9H), 0.28 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.9, 157.5, 155.3 (2 peaks: 155.33, 155.26), 138.7, 135.7, 132.2, 131.9 (t, J = 5.4 Hz), 130.5, 129.1, 125.6, 124.9, 124.2, 122.1, 122.0, 120.2, 114.9, 110.8, 104.0, 96.2, 65.2, 55.7, 52.3, -0.1; MS (MALDI) m/z calcd for M^+ $C_{37}H_{35}BF_2N_2O_5Si$ 664.24; found 664.31.

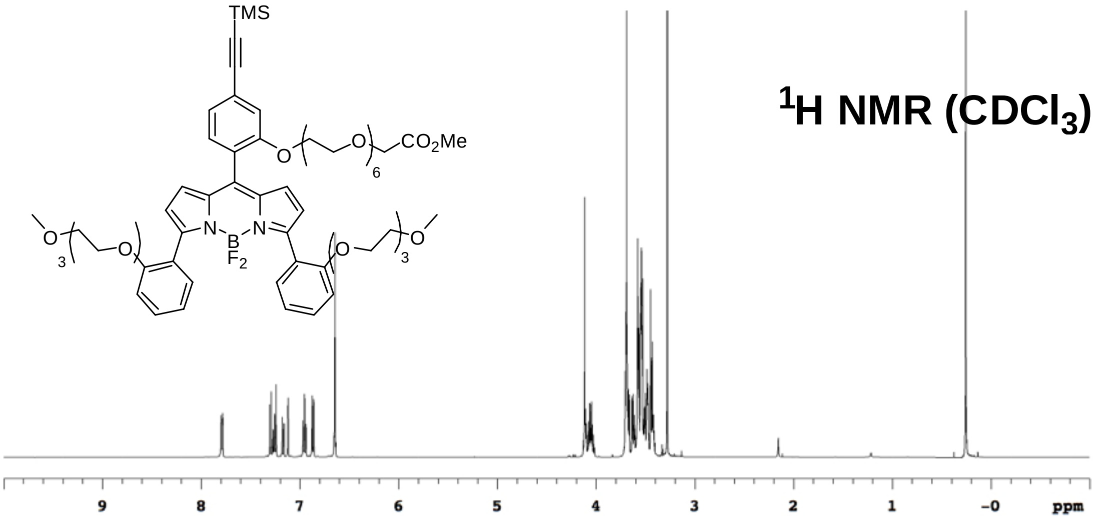

TMS
O
CO2Me
MeO
OMe
N
B
N
F2
1H NMR (CDCl3)
9
8
7
6
5
4
3
2
1
-0
ppm

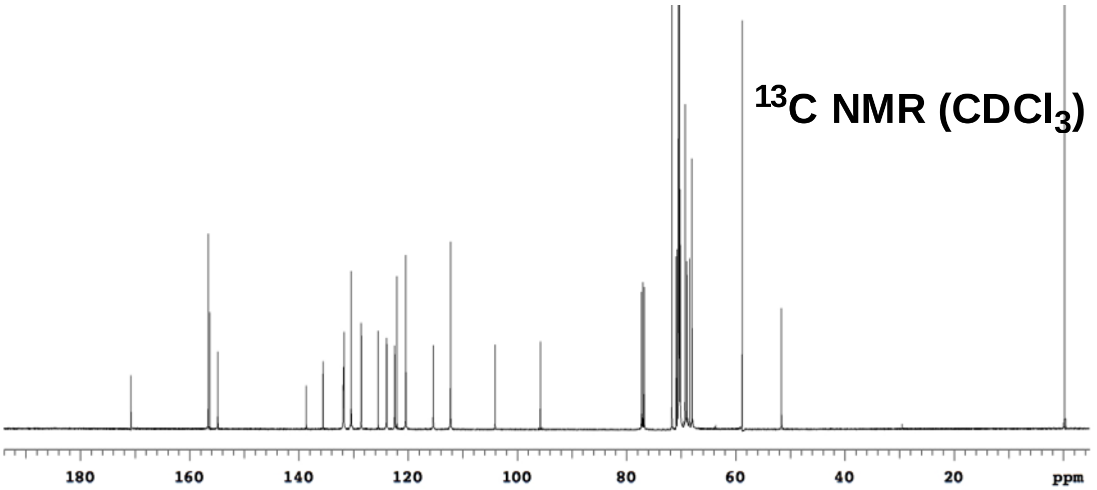

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm

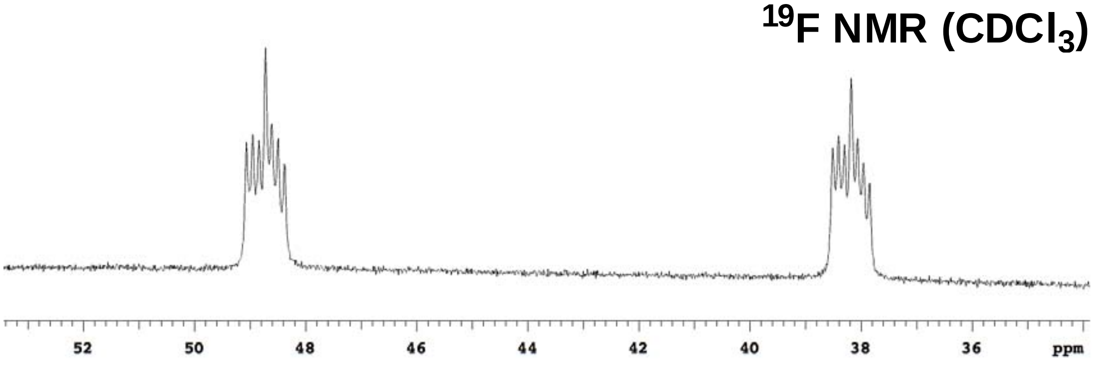


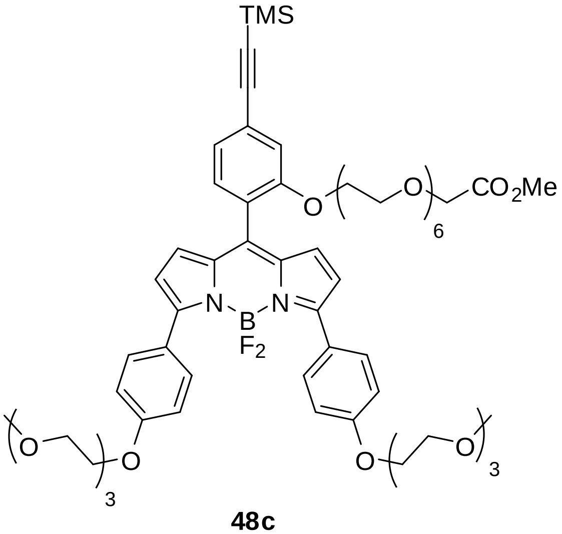
48b

Purple semi-solid (271 mg, 86 %), prepared from **47b** (322 mg, 0.26 mmol). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80-7.78 (m, 2H), 7.30 (d, 1H, *J* = 7.7 Hz), 7.27-7.24 (m, 2H), 7.17 (dd, 1H, *J* = 7.7, 1.5 Hz), 7.12 (d, 1H, *J* = 1.5 Hz), 6.97-6.94 (m, 2H), 6.88-6.86 (m, 2H), 6.65-6.63 (m, 4H), 4.12-4.03 (m, 8H), 3.70-3.42 (m, 45 H), 3.28 (s, 6H), 0.26 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.7, 156.6, 156.3, 154.8, 138.6, 135.6, 131.8 (t, *J* = 5.6 Hz), 131.7, 130.4, 128.6, 125.5, 123.9 (2 peaks: 123.93, 123.88), 122.4, 122.0, 120.4, 115.4, 112.2, 104.1, 95.8, 71.7, 70.9, 70.7, 70.5, 70.4 (4 peaks: 70.45, 70.44, 70.38, 70.37), 70.3 (3 peaks: 70.34, 70.32, 70.28), 70.2, 70.1, 69.3, 69.0, 68.9, 68.4, 68.0, 58.8, 51.6, -0.2, several peaks account for more than one carbon; ^{19}F NMR (282 MHz, $CDCl_3$) δ 48.72 (dq, 1F, *J* = 96.9, 31.7 Hz), 38.18 (dq, 1F, *J* = 96.9, 31.7 Hz); MS (MALDI) m/z calcd for $(M+Na)^+$ $C_{61}H_{83}BF_2N_2NaO_{17}Si$ 1215.54; found 1215.47.


TMS
O
O
6
CO2Me
O
3
O
N
B
N
F2
O
O
3
1H NMR (CDCl3)
9
8
7
6
5
4
3
2
1
–0
ppm

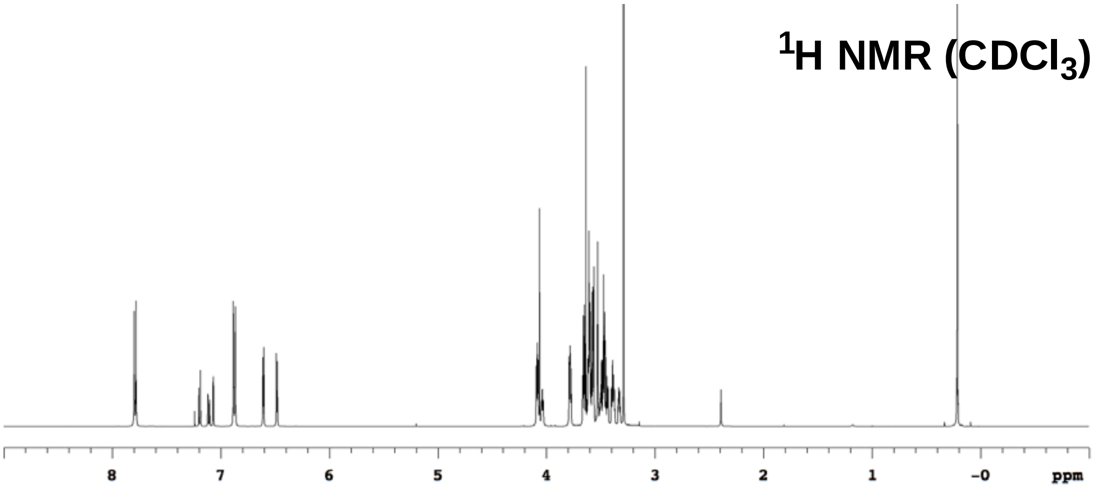

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm


19F NMR (CDCl3)
52
50
48
46
44
42
40
38
36
ppm

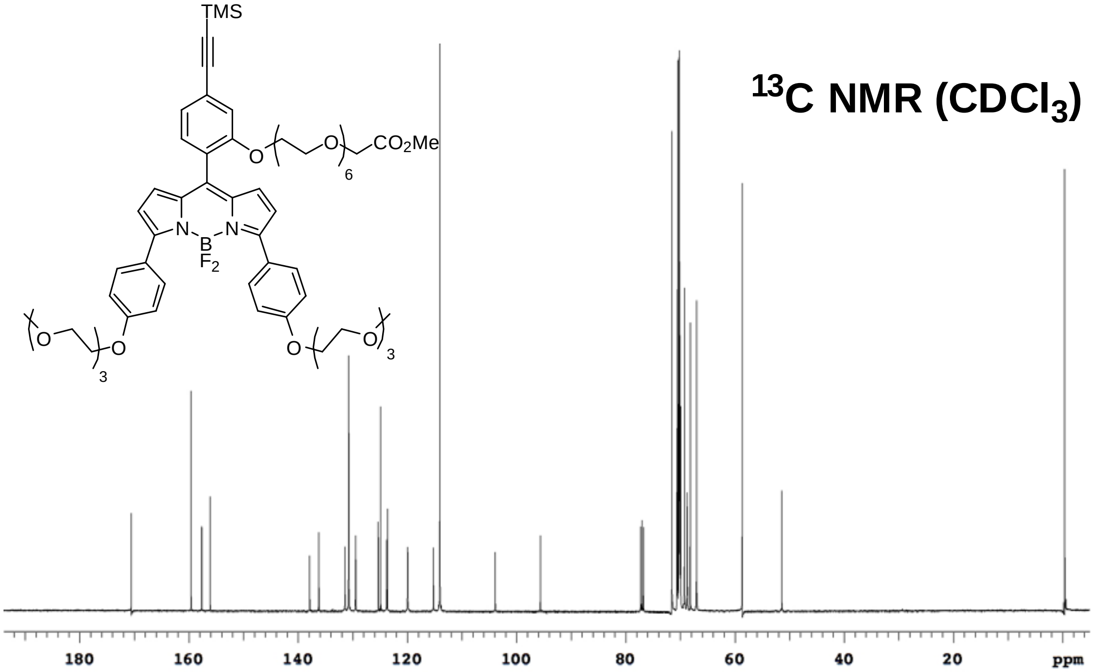


48c

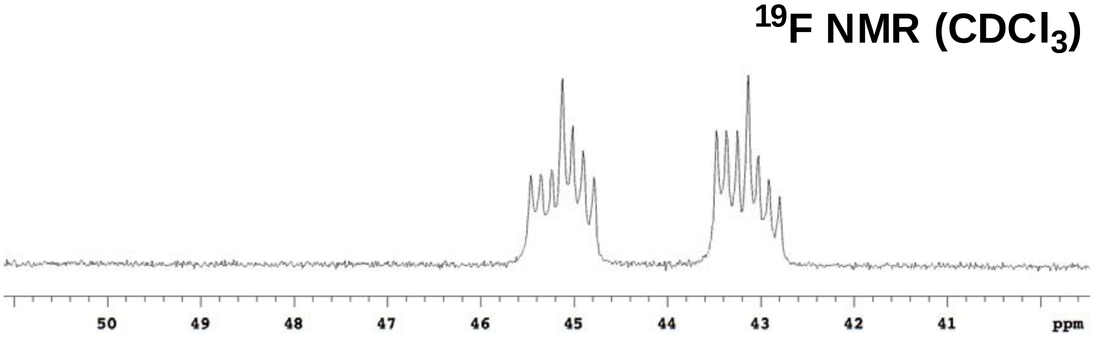
Blue semi-solid (819 mg, *quant.*), prepared from **47c** (840 mg, 0.69 mmol). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (d, 4H, *J* = 9.0 Hz), 7.19 (d, 1H, *J* = 7.8 Hz), 7.11 (dd, 1H, *J* = 7.8, 1.2 Hz), 7.07 (d, 1H, *J* = 1.2 Hz), 6.88 (d, 4H, *J* = 9.0 Hz), 6.61 (d, 2H, *J* = 4.3 Hz), 6.48 (d, 2H, *J* = 4.3 Hz), 4.09-4.07 (m, 4H), 4.06 (s, 2H), 4.05-4.03 (m, 2H), 3.79-3.77 (m, 4H), 3.67-3.43 (m, 35H), 3.40-3.37 (m, 4H), 3.34-3.32 (m, 2H), 3.29 (s, 6H), 0.21 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.5, 159.6, 157.6, 156.1, 137.9, 136.2, 131.4, 130.7 (t, *J* = 4.2 Hz), 129.4, 125.3, 124.9, 123.8, 123.6, 119.9, 115.2, 114.0, 103.9, 95.6, 71.6, 70.6, 70.5 (2 peaks: 70.50, 70.46), 70.3, 70.2 (4 peaks: 70.23, 70.20, 70.17, 70.16), 70.1 (3 peaks: 70.14, 70.12, 70.07), 70.0, 69.9, 69.3, 68.8, 68.7, 68.2, 67.0, 58.7, 51.4, -0.4, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, $CDCl_3$) δ 45.12 (dq, 1F, *J* = 96.3, 31.7 Hz), 43.13 (dq, 1F, *J* = 96.3, 31.7 Hz); MS (MALDI) m/z calcd for $(M-F)^+$ $C_{61}H_{83}BFN_2O_{17}Si$ 1173.55; found 1173.55.



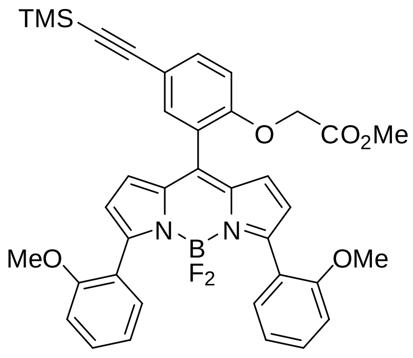
^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

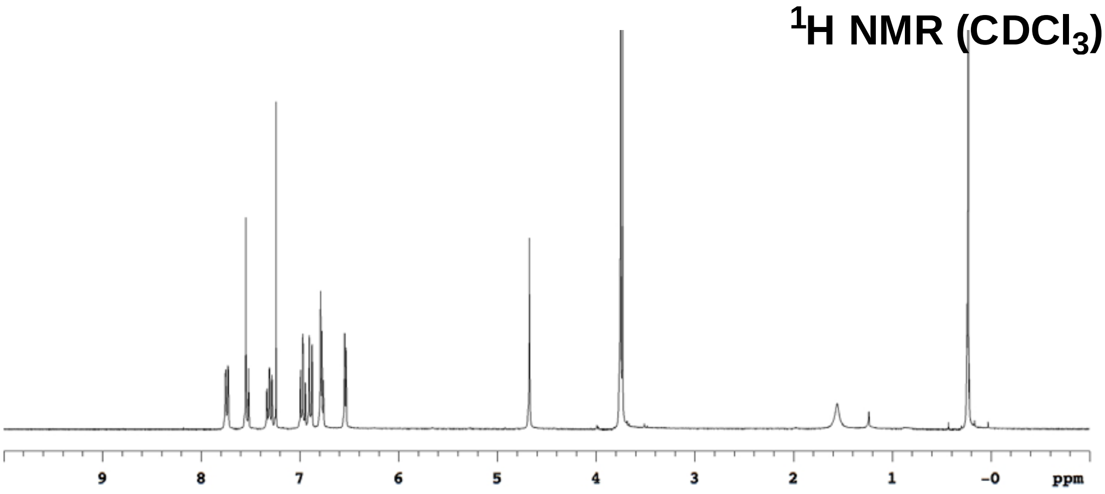


^{19}F NMR ($CDCl_3$)

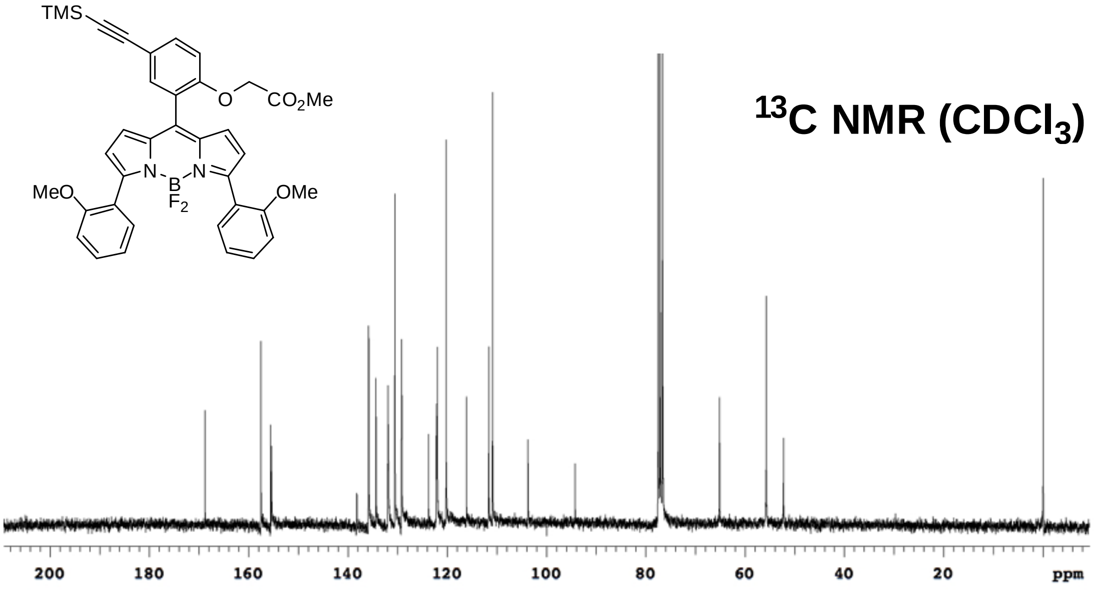


48d

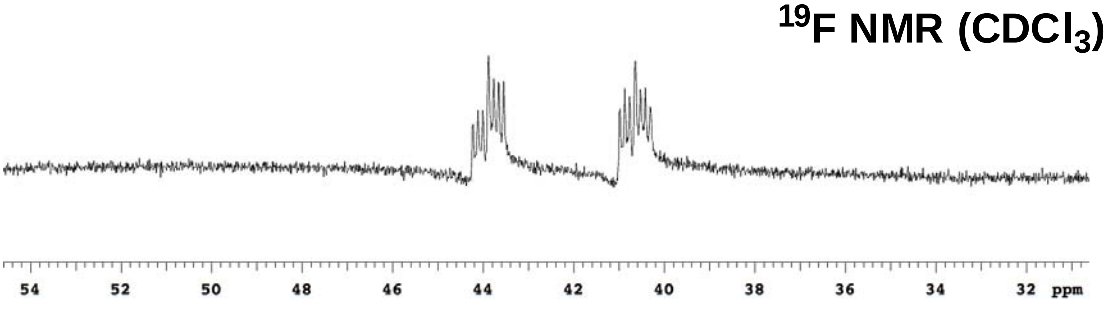
Red solid (194 mg, 83 %), prepared from **47d** (245 mg, 0.35 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.75-7.72 (m, 2H), 7.55-7.51 (m, 2H), 7.34-7.28 (m, 2H), 6.99-6.94 (m, 2H), 6.91-6.87 (m, 2H), 6.79-6.76 (m, 3H), 6.54 (d, 2H, *J* = 4.3 Hz), 4.67 (s, 2H), 3.75 (s, 6H), 3.73 (s, 3H), 0.23 (s, 9H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 168.8, 157.5, 155.6, 155.4, 138.2, 135.8, 134.3, 131.9, 130.5, 129.2, 123.7, 122.1, 122.0, 120.2, 116.1, 111.6, 110.9, 103.7, 94.2, 65.1, 55.7, 52.2, -0.1, (one carbon is ambiguous, some peak accounts for two carbons); ^{19}F NMR (282 MHz, CDCl$_3$) δ 43.88 (dq, 1F, *J* = 96.8, 30.6 Hz), 40.64 (dq, 1F, *J* = 96.8, 30.6 Hz); MS (MALDI) m/z calcd for M$^+$ $C_{37}H_{35}BF_2N_2O_5Si$ 664.24; found 664.12.



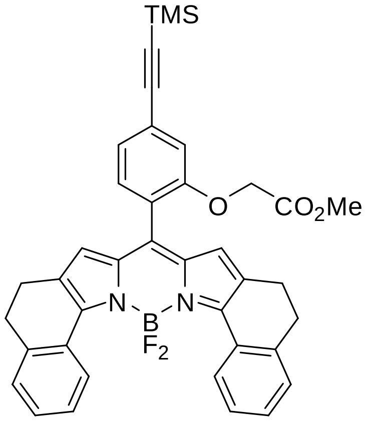
^{1}H NMR (CDCl$_3$)
9
8
7
6
5
4
3
2
1
-0
ppm



TMS
O
CO$_2$Me
N
B
N
F$_2$
MeO
OMe
^{13}C NMR (CDCl$_3$)
200
180
160
140
120
100
80
60
40
20
ppm

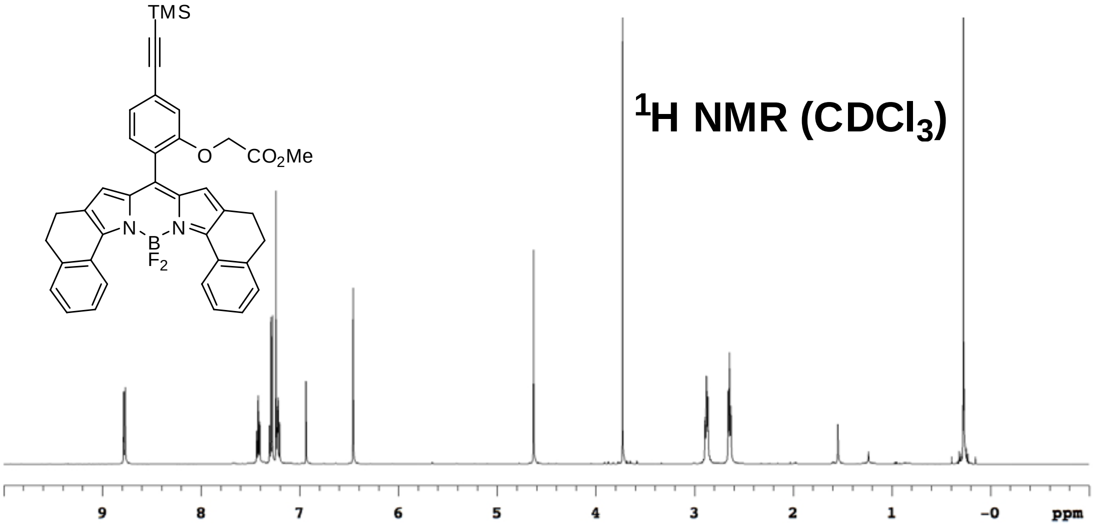


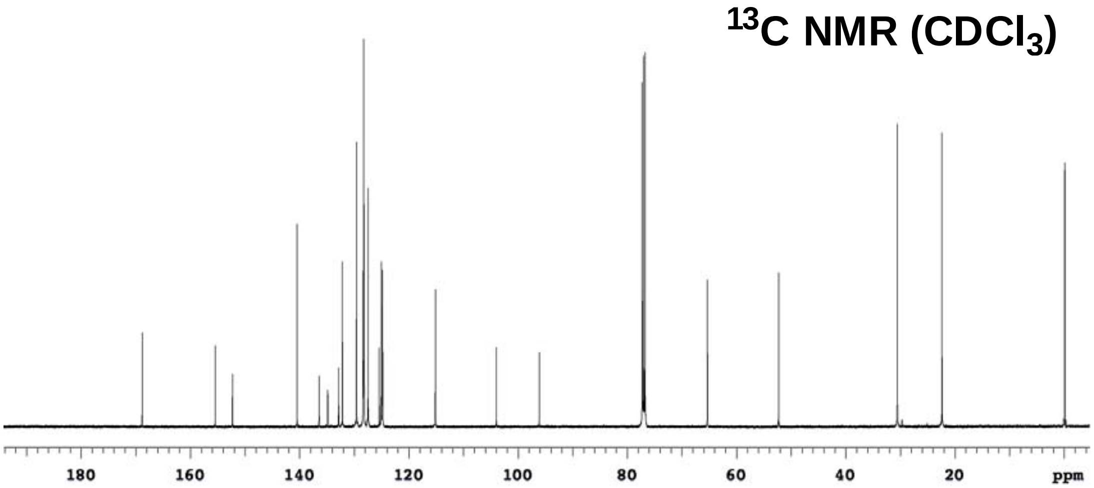
^{19}F NMR (CDCl$_3$)
54
52
50
48
46
44
42
40
38
36
34
32
ppm

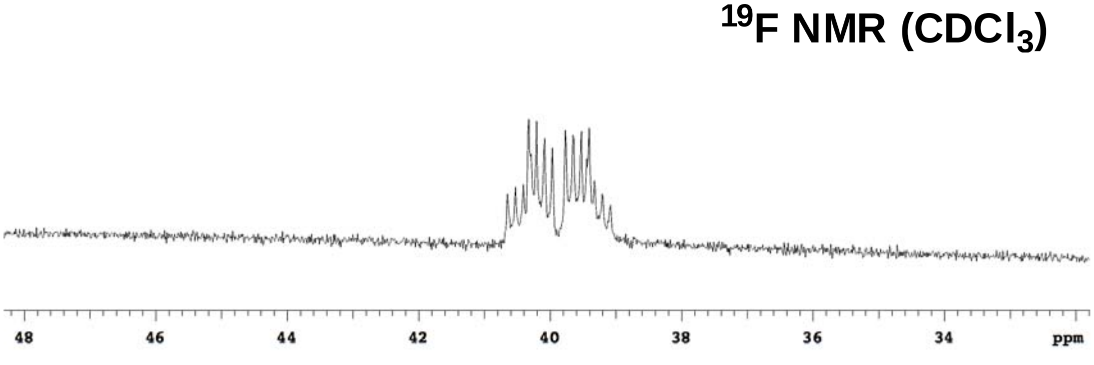


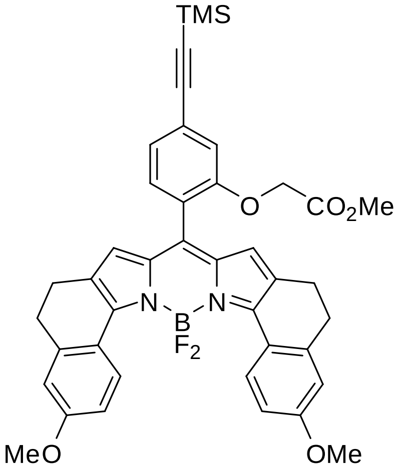
48e

Blue solid (174 mg, 88 %), prepared from **47e** (206 mg, 0.3 mmol). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.78 (d, 2H, *J* = 8.0 Hz), 7.44-7.40 (m, 2H), 7.31-7.28 (m, 3H), 7.24-7.20 (m, 3H), 6.94 (d, 1H, *J* = 1.3 Hz), 6.46 (s, 2H), 4.63 (s, 2H), 3.73 (s, 3H), 2.88 (t, 4H, *J* = 7.2 Hz), 2.64 (t, 4H, *J* = 7.2 Hz), 0.27 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.8, 155.4, 152.3, 140.5, 136.4, 134.8, 132.8, 132.1, 129.6, 128.4, 128.3, 128.2, 127.5, 125.4, 125.0, 124.9, 124.8, 115.1, 104.0, 96.1, 65.3, 52.3, 30.5, 22.4, -0.1; ^{19}F NMR (282 MHz, $CDCl_3$) δ 40.67-39.08 (m); MS (MALDI) m/z calcd for M^+ $C_{39}H_{35}BF_2N_2O_3Si$ 656.25; found 656.16.


^{1}H NMR (CDCl$_3$)

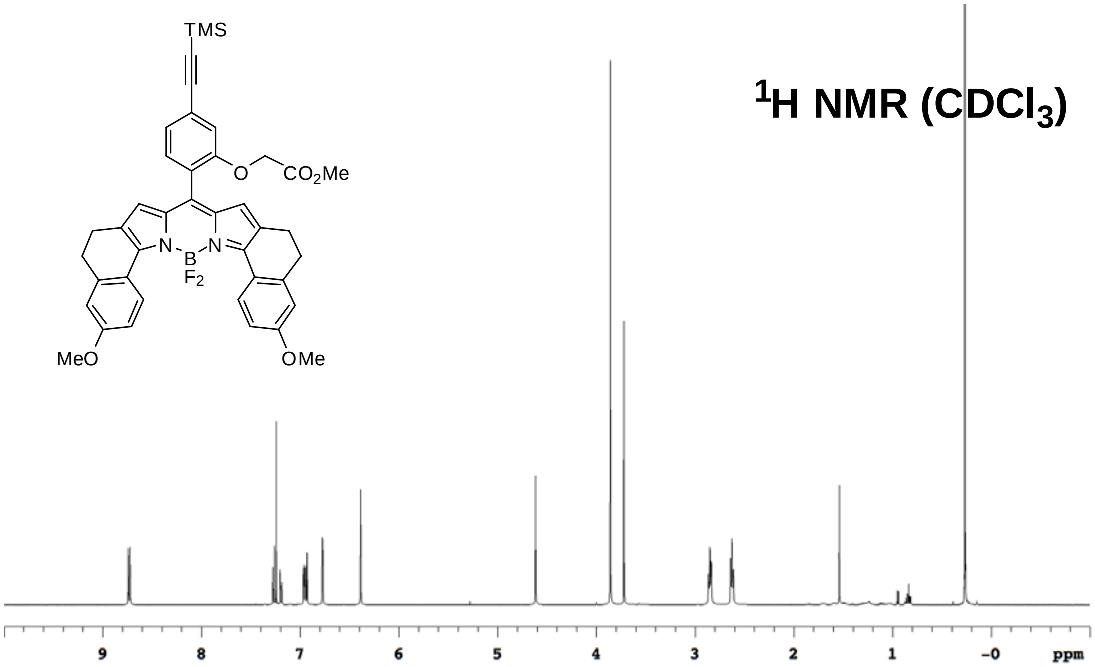

^{13}C NMR (CDCl$_3$)


^{19}F NMR (CDCl$_3$)

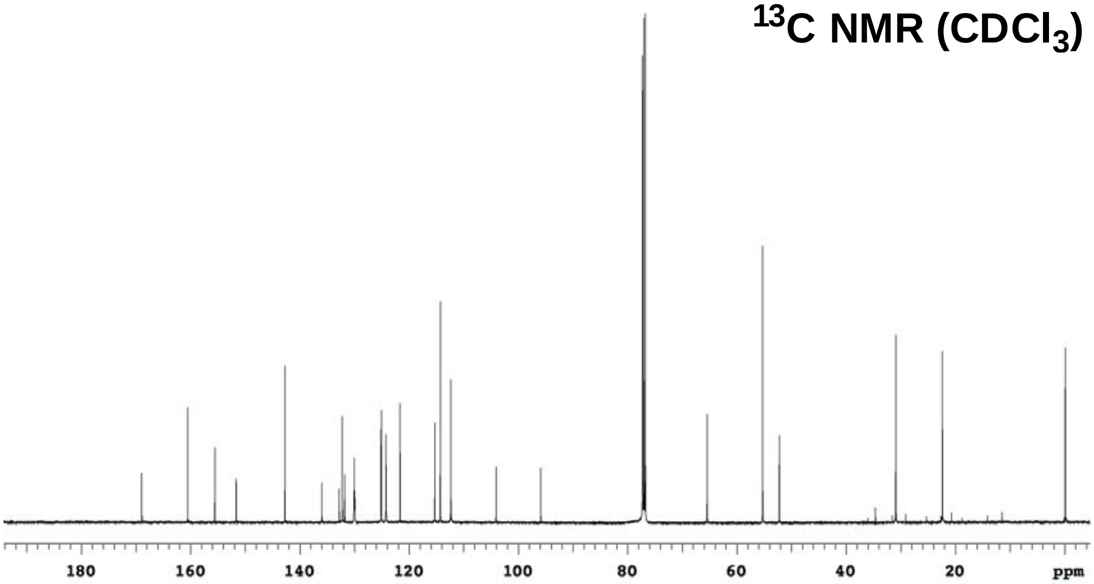


48f

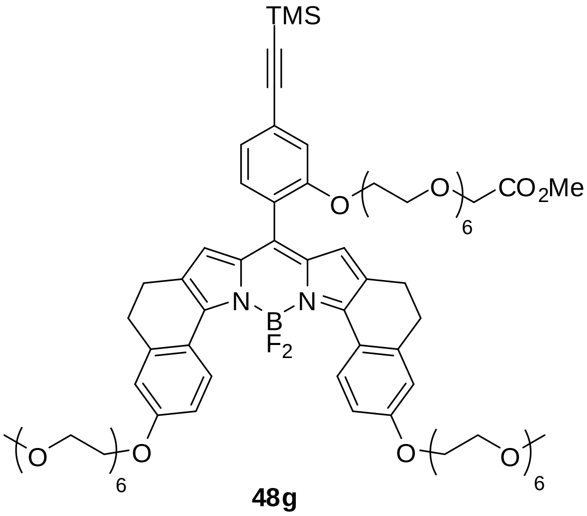
Blue solid (124 mg, 87 %), prepared from **47f** (149 mg, 0.2 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.73 (d, 2H, *J* = 9.0 Hz), 7.27 (d, 1H, *J* = 7.6 Hz), 7.19 (dd, 1H, *J* = 7.6, 1.2 Hz), 6.95 (dd, 2H, *J* = 9.0, 2.7 Hz), 6.93 (d, 1H, *J* = 1.2 Hz), 6.77 (d, 2H, *J* = 2.7 Hz), 6.39 (s, 2H), 4.62 (s, 2H), 3.86 (s, 6H), 3.72 (s, 3H), 2.85 (t, 4H, *J* = 7.1 Hz), 2.63 (t, 4H, *J* = 7.1 Hz), 0.27 (s, 9H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 168.9, 160.5, 155.6, 151.6, 142.7, 136.0, 132.8, 132.2, 131.8, 130.0 (t, *J* = 11.2 Hz), 125.2, 125.1, 124.2, 121.6, 115.3, 114.3, 112.3, 104.1, 95.9, 65.4, 55.3, 52.2, 30.9, 22.4, -0.1, (one carbon is ambiguous, some peak accounts for two carbons); MS (MALDI) m/z calcd for M$^+$ $C_{41}H_{39}BF_2N_2O_5Si$ 716.27; found 716.13.



^{1}H NMR (CDCl$_3$)

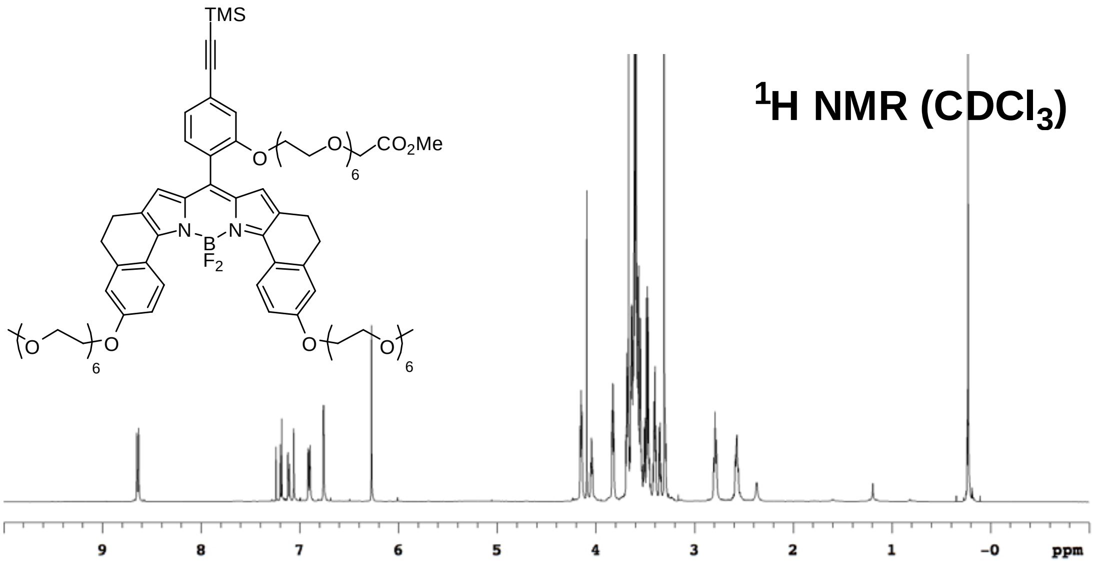


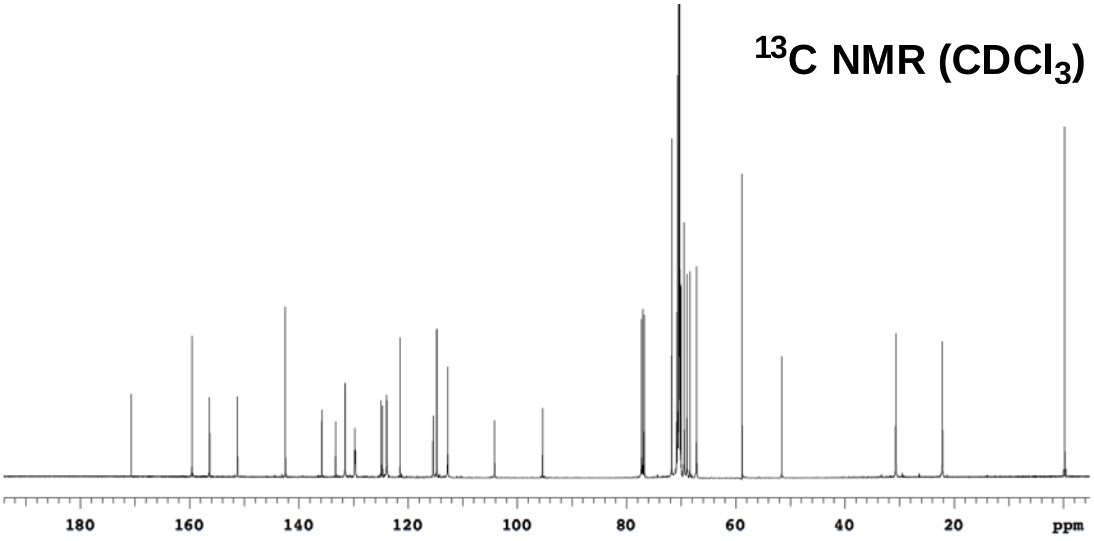
^{13}C NMR (CDCl$_3$)

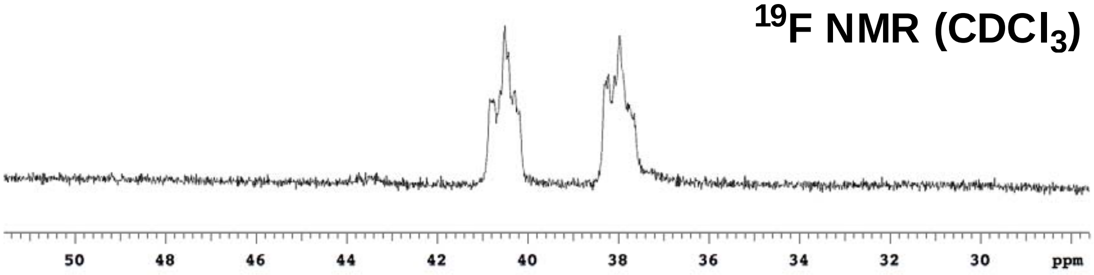


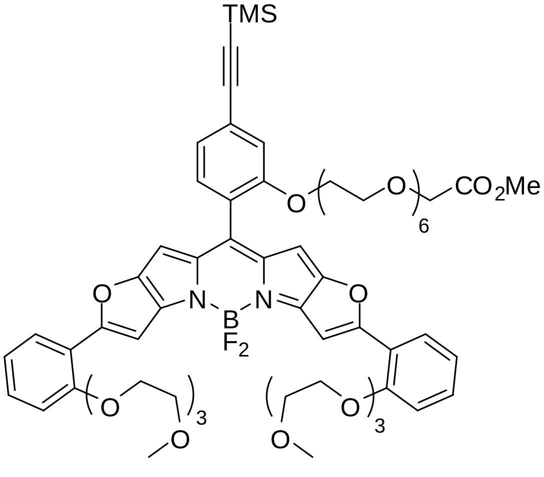
48g

Green sticky solid (566 mg, 80 %), prepared from **47f** (724 mg, 0.47 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.64 (d, 2H, *J* = 9.0 Hz), 7.19 (d, 1H, *J* = 7.7 Hz), 7.11 (dd, 1H, *J* = 7.7, 1.3 Hz), 7.06 (d, 1H, *J* = 1.3 Hz), 6.90 (dd, 2H, *J* = 9.0, 2.6 Hz), 6.76 (d, 2H, *J* = 2.6 Hz), 6.27 (s, 2H), 4.16-4.14 (m, 4H), 4.09 (s, 2H), 4.05-4.03 (m, 2H), 3.84-3.84 (m, 4H), 3.69-3.35 (m, 65H), 3.31 (s, 6H), 2.81-2.78 (m, 4H), 2.59-2.56 (m, 4H), 0.23 (s, 9H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 170.7, 159.6, 156.3, 151.2, 142.5, 135.8, 133.3, 131.5 (2 peaks: 131.54, 131.52), 129.7 (t, *J* = 10.7 Hz), 124.9, 124.7, 123.9 (2 peaks: 123.92, 123.88), 121.4, 115.4, 114.7, 112.7, 104.2, 95.4, 71.7, 70.8, 70.7, 70.6, 70.4 (3 peaks: 70.42, 70.39, 70.35), 70.3 (3 peaks: 70.34, 70.28, 70.27), 70.2 (2 peaks: 70.24, 70.18), 70.1, 70.0, 69.4, 68.9, 68.4, 67.2, 58.8, 51.6, 30.7, 22.1, -0.2, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 40.84-40.19 (m, 1F), 38.31-37.65 (m, 1F); MS (MALDI) m/z calcd for M$^+$ $C_{77}H_{111}BF_2N_2O_{23}Si$ 1508.74; found 1508.68.


^{1}H NMR (CDCl$_3$)

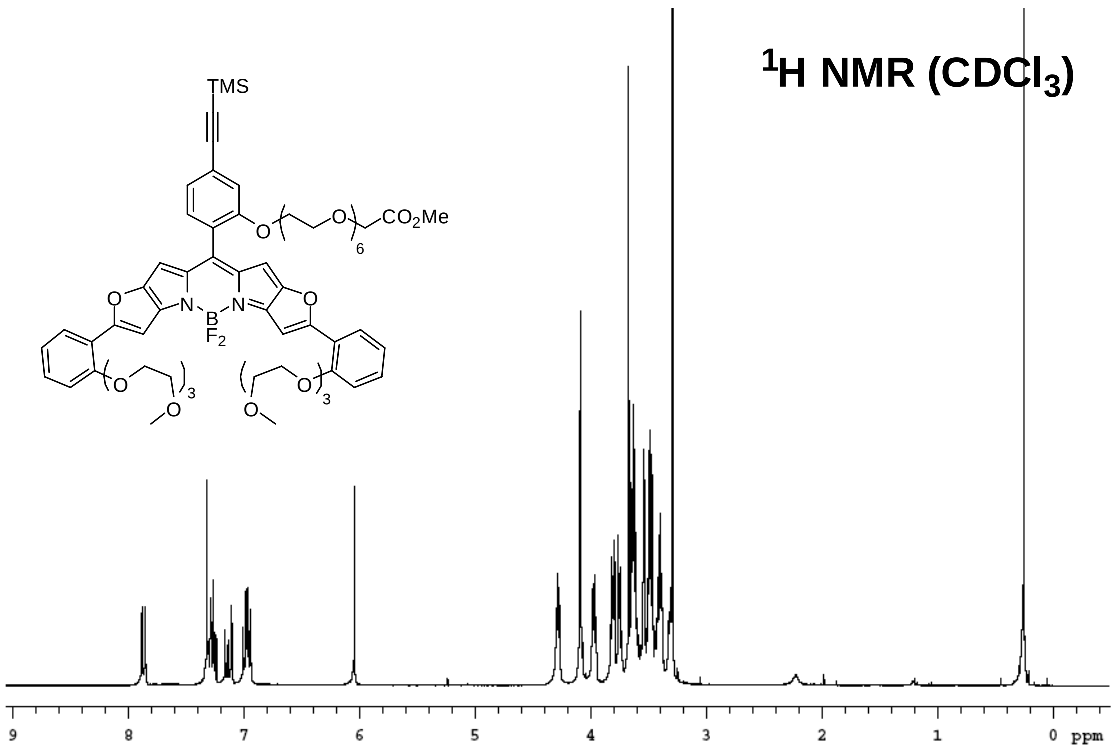

^{13}C NMR (CDCl$_3$)


^{19}F NMR (CDCl$_3$)

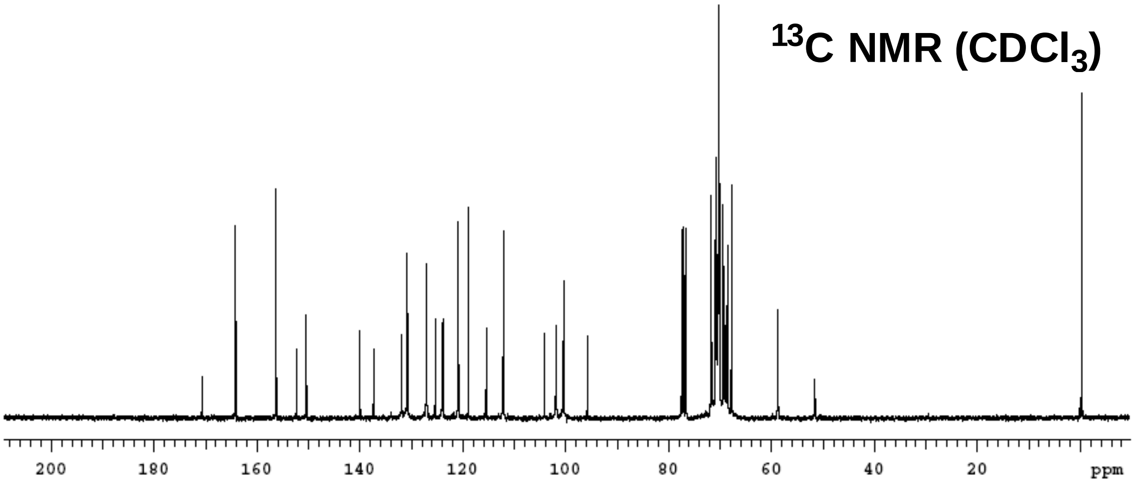


48h

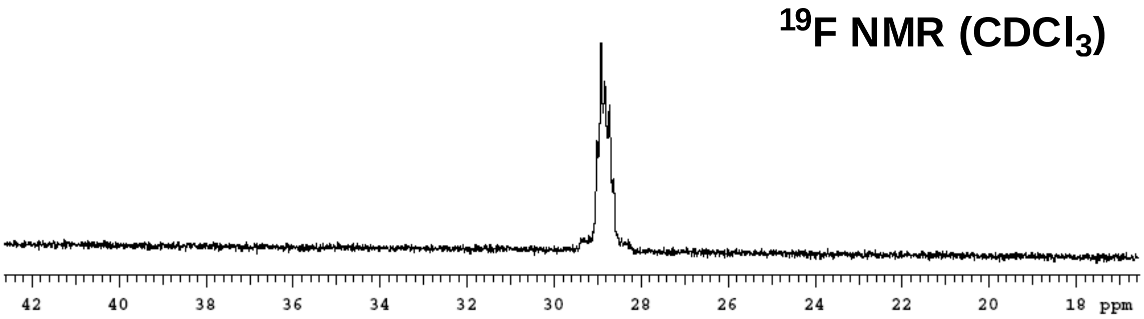
Green semi-solid (443 mg, *quant.*), prepared from **47h** (458 mg, 0.35 mmol). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.87 (dd, 2H, *J* = 7.9, 1.7 Hz), 7.32-7.26 (m, 5H), 7.17-7.11 (m, 2H), 7.01-6.94 (m, 4H), 6.05 (s, 2H), 4.30-4.27 (m, 4H), 4.10-4.07 (m, 4H), 3.99-3.96 (m, 4H), 3.83-3.80 (m, 4H), 3.76-3.73 (m, 4H), 3.67 (s, 3H), 3.65-3.31 (m, 30H), 3.29 (s, 6H), 0.25 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.7, 164.2, 156.3, 156.2, 152.3, 150.4, 140.0, 137.3, 131.9, 130.8, 127.1, 125.4, 124.0, 123.8, 120.9, 119.0, 115.5, 112.1, 104.1, 101.9, 100.4, 95.8, 71.7, 70.9, 70.7 (3 peak: 70.74, 70.69, 70.65), 70.4, 70.3 (2 peaks: 70.33, 70.28), 70.2, 70.1, 69.4, 68.9, 68.7, 68.4, 67.7, 58.8, 51.6, -0.2, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, $CDCl_3$) δ 29.01-28.63 (m); MS (MALDI) m/z calcd for $(M-F)^+$ $C_{65}H_{83}BFN_2O_{19}Si$ 1253.54, found 1253.49; calcd for $(M+Na)^+$ $C_{65}H_{83}BF_2NaN_2O_{19}Si$ 1295.53, found 1295.47.



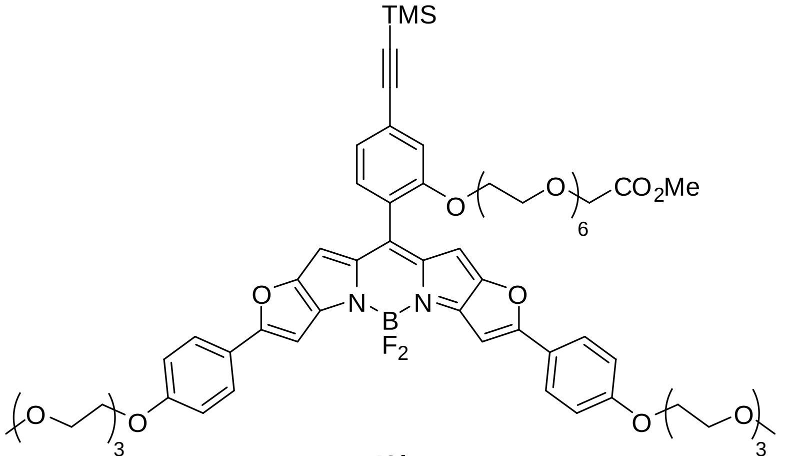
^{1}H NMR (CDCl$_3$)



^{13}C NMR (CDCl$_3$)

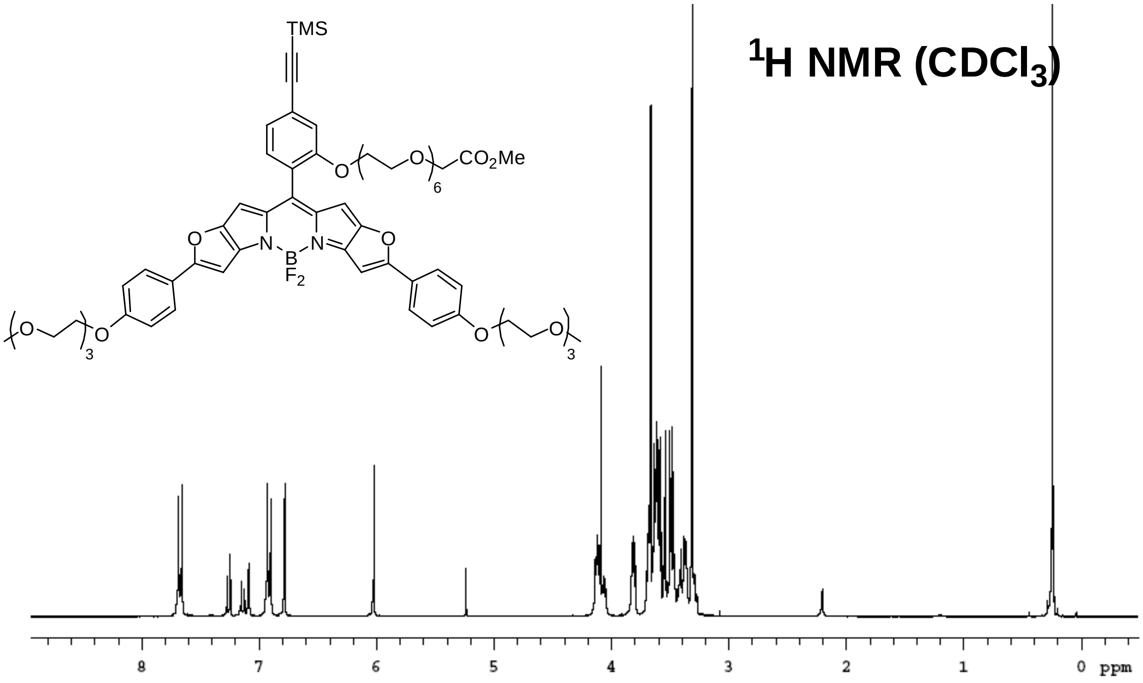


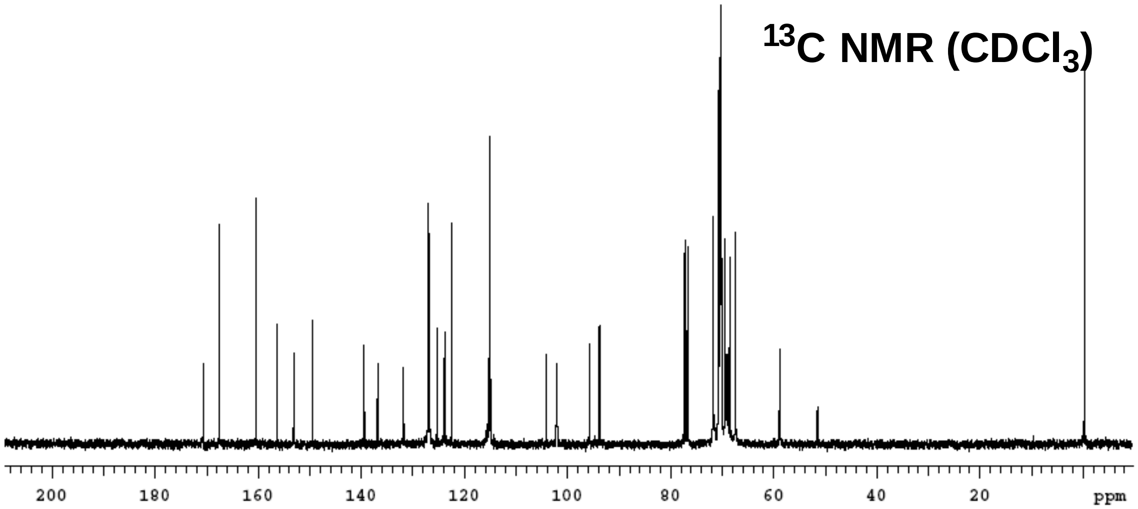
^{19}F NMR (CDCl$_3$)

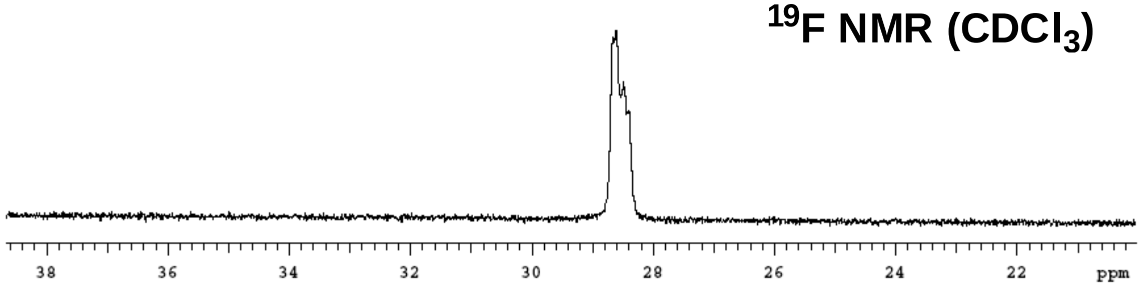


48i

Green semi-solid (1.08 g, 94 %), prepared from **47i** (1.18 g, 0.91 mmol). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.67 (d, 4H, *J* = 8.9 Hz), 7.26 (d, 1H, *J* = 7.7 Hz), 7.15-7.12 (m, 1H), 7.09-7.08 (m, 1H), 6.92 (d, 4H, *J* = 8.9 Hz), 6.78 (s, 2H), 6.02 (s, 2H), 4.14-4.10 (m, 4H), 4.09 (s, 2H), 4.08-4.05 (m, 2H), 3.83-3.79 (m, 4H), 3.70-3.27 (m, 47H), 0.24 (s, 9H); ^{13}C NMR (75 MHz, CDCl$_3$) δ 170.7, 167.6, 160.4, 156.3, 153.1, 149.4, 139.4, 136.9, 131.8, 126.9, 125.4, 124.0, 123.7, 122.5, 115.4, 115.0, 104.1, 102.1, 95.8, 93.8, 71.7, 70.8, 70.7, 70.5, 70.4, 70.3, 70.2, 70.1 (2 peaks: 70.12, 70.07), 69.4, 68.9, 68.8, 68.4, 67.4, 58.8, 51.5, -0.2, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 28.66-28.39 (m); MS (MALDI) m/z calcd for (M-F)$^+$ $C_{65}H_{83}BFN_2O_{19}Si$ 1253.54, found 1253.58; calcd for M$^+$ $C_{65}H_{83}BF_2N_2O_{19}Si$ 1272.54, found 1272.58; calcd for (M+Na)$^+$ $C_{65}H_{83}BF_2NaN_2O_{19}Si$ 1295.53, found 1295.54.

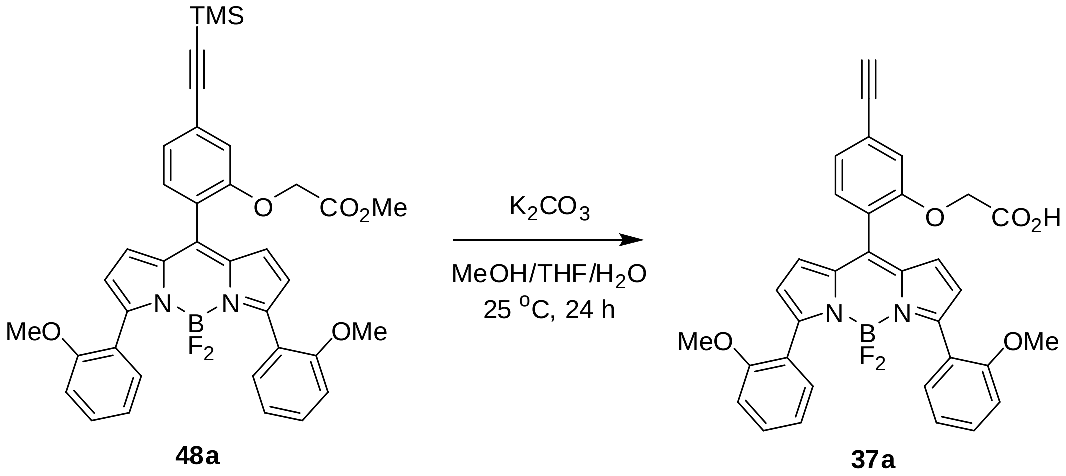

^{1}H NMR (CDCl$_3$)


^{13}C NMR (CDCl$_3$)

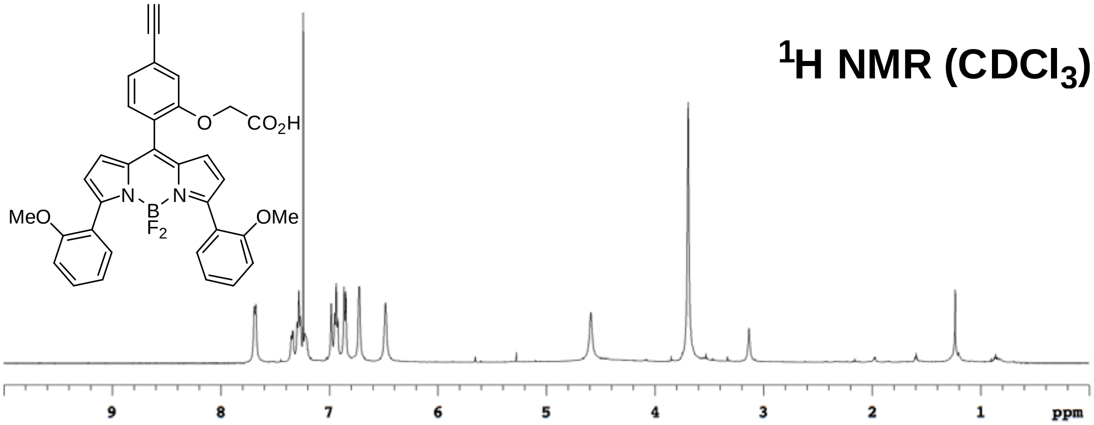

^{19}F NMR (CDCl$_3$)

(5) Syntheses of BODIPYs 37

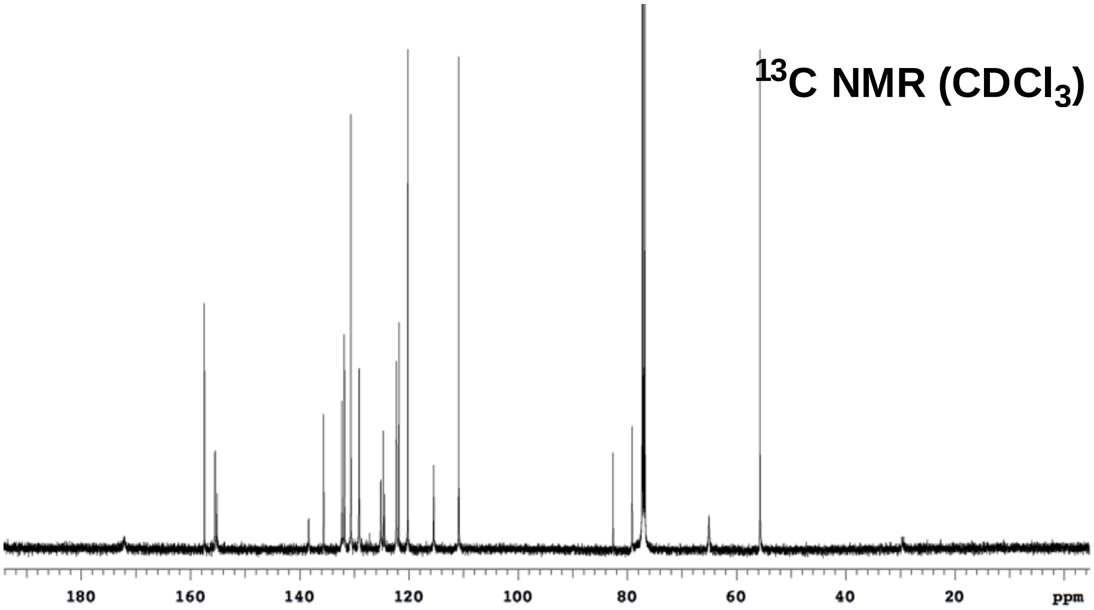
Typical procedures:



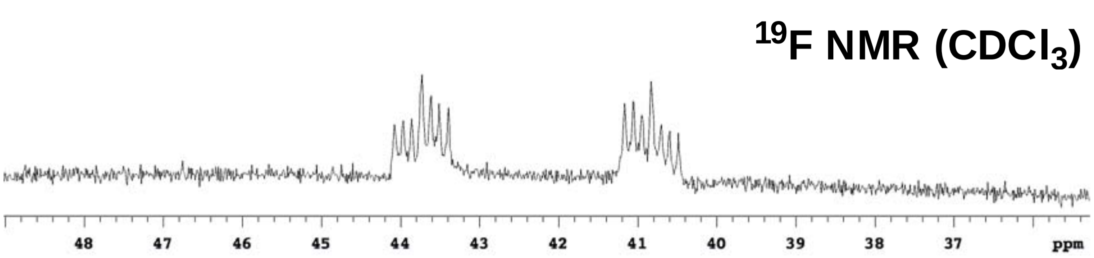
BODIPY **48a** (66 mg, 0.1 mmol) and K_2CO_3 (69 mg, 0.5 mmol) were dissolved in 5 mL MeOH + 5 mL THF + 1 mL H_2O. After stirring at room temperature for 24 h, the reaction mixture was diluted with H_2O, acidified with 2 M HCl and extracted with CH_2Cl_2. The organic layer was washed with H_2O, brine and dried over Na_2SO_4 then evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH_2Cl_2) to afford the product **37a** (49 mg, 85 %) as red solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.69-7.68 (m, 2H), 7.35-7.34 (m, 1H), 7.30-7.26 (m, 2H), 7.23-7.21 (m, 1H), 6.98 (s, 1H), 6.95-6.92 (m, 2H), 6.86-6.85 (m, 2H), 6.72 (br, 2H), 6.48 (br, 2H), 4.59 (s, 2H), 3.69 (s, 6H), 3.13 (s, 1H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.1, 157.4, 155.5, 155.1, 138.3, 135.6, 132.3, 131.8, 130.6, 129.1, 125.1, 124.7, 124.4, 122.3, 121.8, 120.2, 115.4, 110.9, 82.6, 79.1, 65.1, 55.7; ^{19}F NMR (282 MHz, $CDCl_3$) δ 43.73 (dq, 1F, *J* = 97.6, 30.7 Hz), 40.83 (dq, 1F, *J* = 97.6, 30.7 Hz); MS (ESI) m/z calcd for $(M\text{-}H)^-$ $C_{33}H_{24}BF_2N_2O_5$ 577.17; found 577.21.



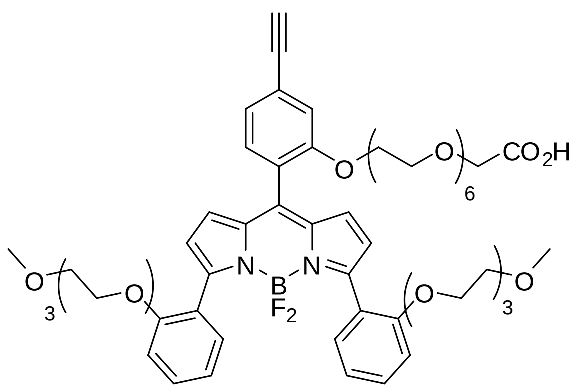
^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

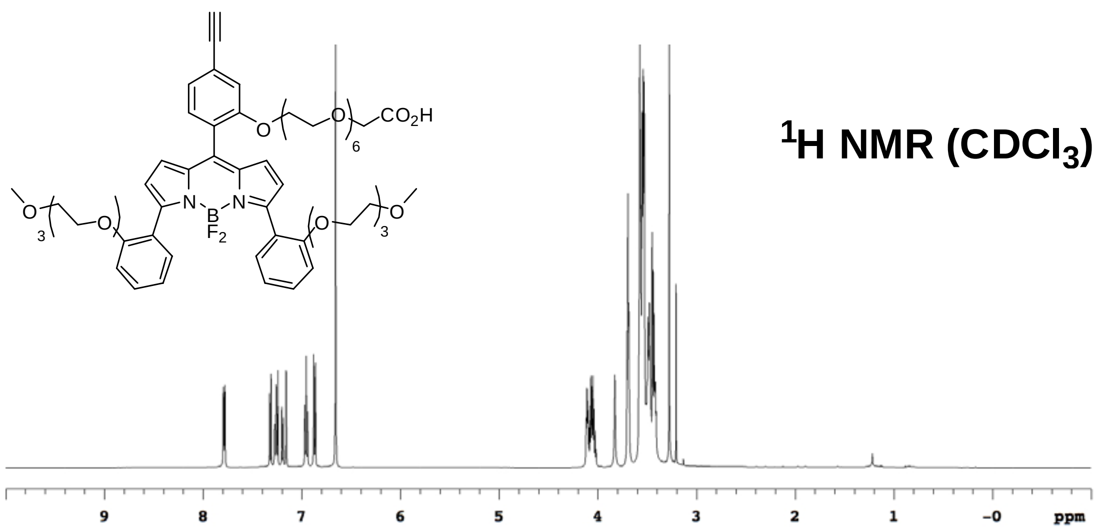


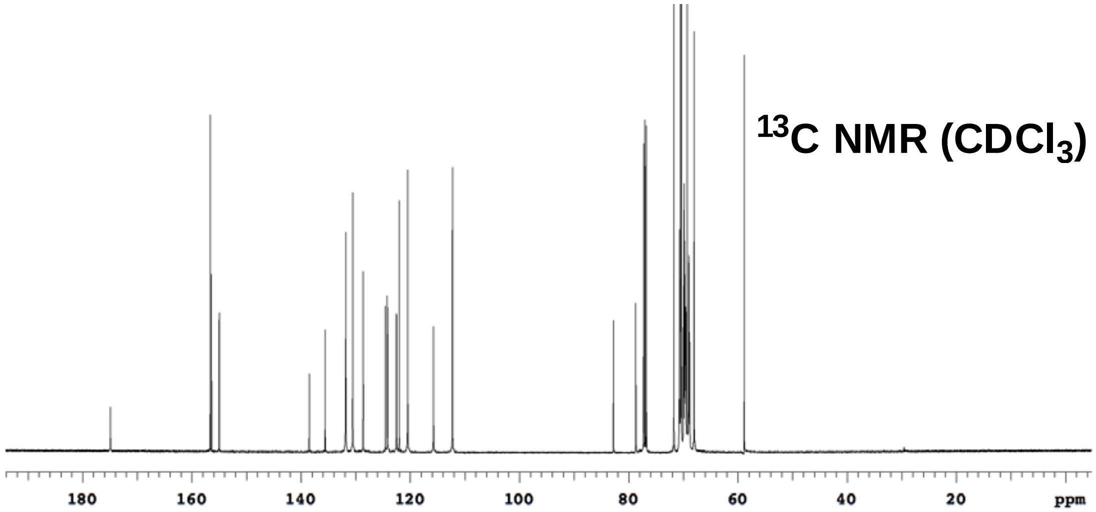
^{19}F NMR ($CDCl_3$)

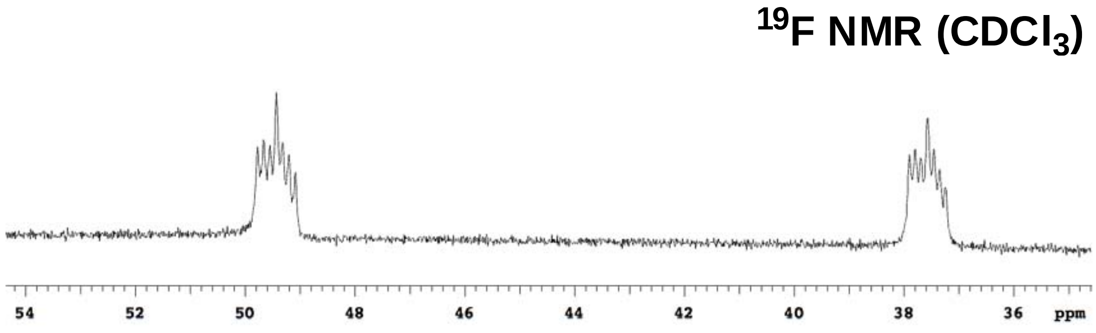


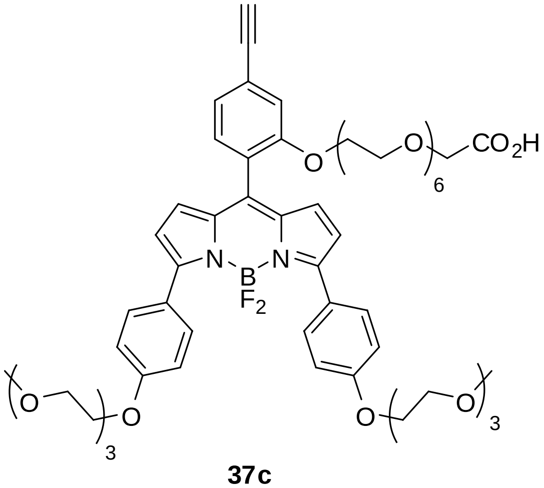
37b

Purple semi-solid (191 mg, 76 %), prepared from **48b** (271 mg, 0.23 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.79 (d, 2H, *J* = 7.6 Hz), 7.32 (d, 1H, *J* = 7.6 Hz), 7.28-7.24 (m, 2H), 7.20-7.19 (m, 1H), 7.16 (br, 1H), 6.97-6.94 (m, 2H), 6.87 (d, 2H, *J* = 8.3 Hz), 6.66 (br, 4H), 4.12-4.02 (m, 6H), 3.83 (s, 2H), 3.70-3.68 (m, 6H), 3.58-3.41 (m, 36H), 3.28 (s, 6H), 3.21 (s, 1H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 174.9, 156.6, 156.4, 155.0, 138.5, 135.6, 131.8, 130.5, 128.6, 124.6, 124.2, 124.1, 122.5, 122.0, 120.5, 115.7, 112.3, 82.8, 78.7, 71.7, 70.7, 70.5 (2 peaks: 70.51, 70.46), 70.3, 70.0, 69.8, 69.7 (3 peaks: 69.74, 69.71, 69.66), 69.6, 69.5, 69.3, 69.0, 68.9, 68.8, 68.0, 58.8, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 49.43 (dq, 1F, *J* = 97.6, 32.4 Hz), 37.58 (dq, 1F, *J* = 97.6, 32.4 Hz); MS (ESI) m/z calcd for (M-H)$^-$ $C_{57}H_{72}BF_2N_2O_{17}$ 1105.49; found 1105.45.


^{1}H NMR (CDCl$_3$)
O
6
CO$_2$H
N
B
N
F$_2$
O
3
O
O
3
O
9
8
7
6
5
4
3
2
1
-0
ppm

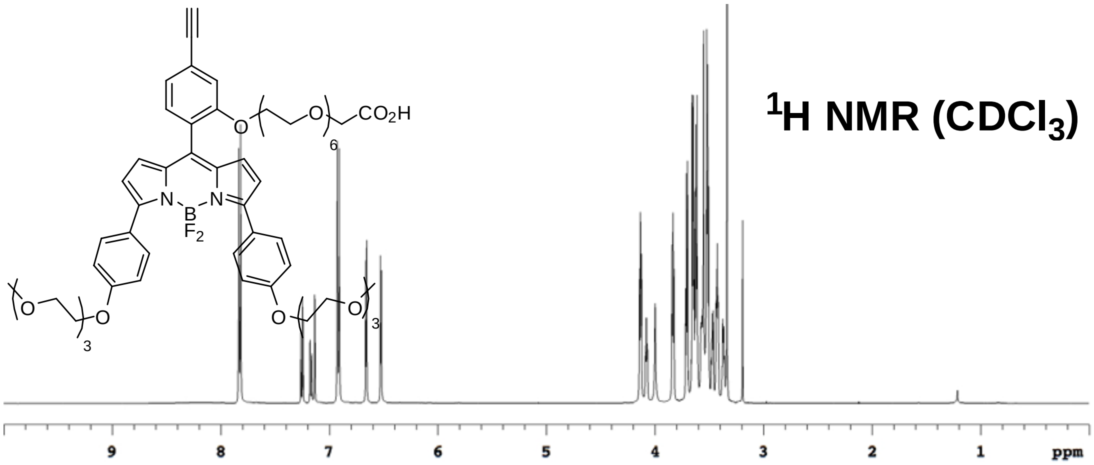

^{13}C NMR (CDCl$_3$)
180
160
140
120
100
80
60
40
20
ppm

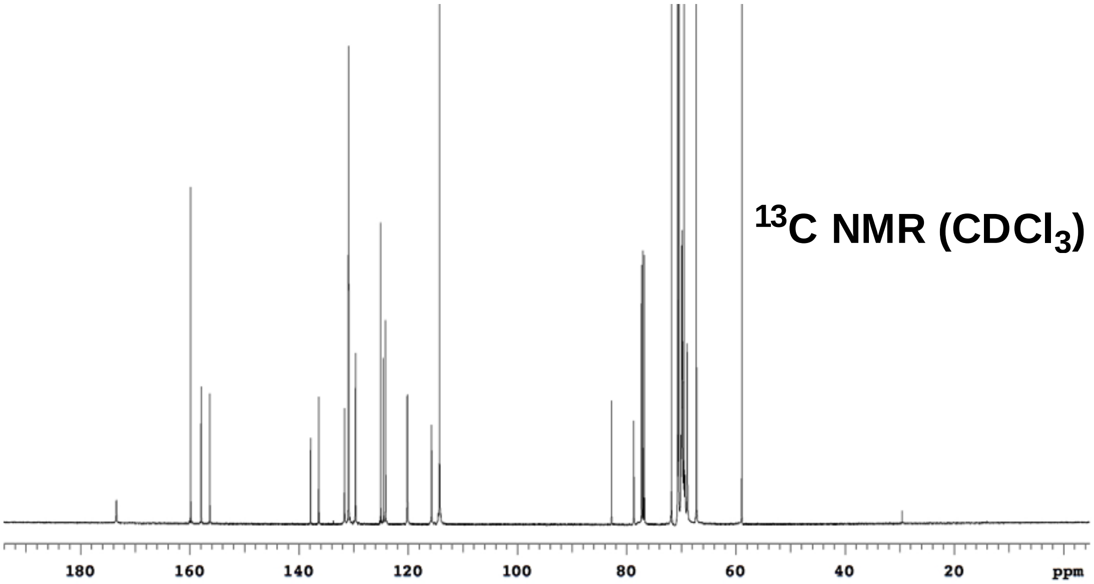

^{19}F NMR (CDCl$_3$)
54
52
50
48
46
44
42
40
38
36
ppm

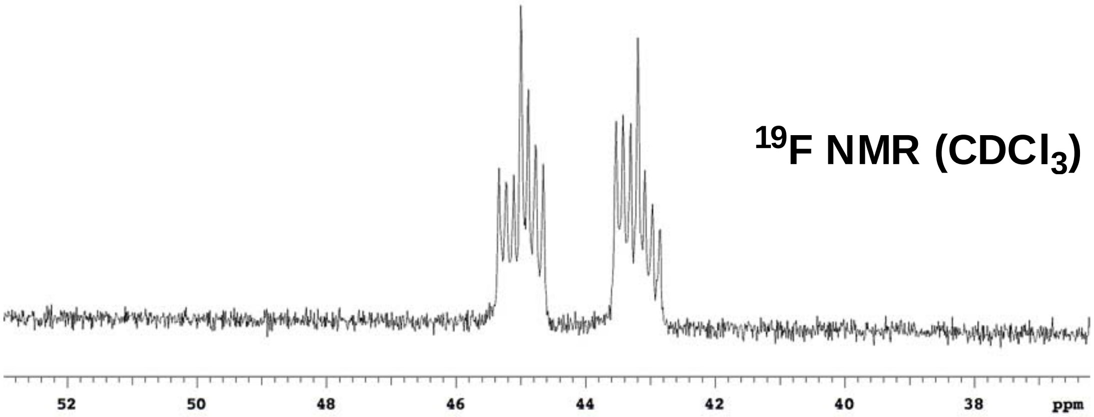


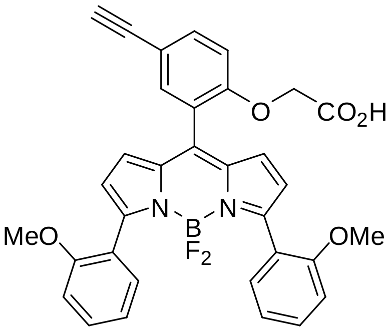
37c

Blue semi-solid (512 mg, 67 %), prepared from **48c** (815 mg, 0.69 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.82 (d, 4H, *J* = 9.0 Hz), 7.25 (d, 1H, *J* = 7.8 Hz), 7.18-7.16 (m, 1H), 7.14-7.13 (m, 1H), 6.92 (d, 4H, *J* = 9.0 Hz), 6.66 (d, 2H, *J* = 4.3 Hz), 6.52 (d, 2H, *J* = 4.3 Hz), 4.14-4.12 (m, 4H), 4.09-4.07 (m, 2H), 4.00 (s, 2H), 3.84-3.82 (m, 4H), 3.72-3.69 (m, 4H), 3.66-3.51 (m, 26H), 3.48-3.46 (m, 2H), 3.44-3.42 (m, 4H), 3.38-3.36 (m, 2H), 3.34 (s, 6H), 3.19 (s, 1H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 173.4, 159.8, 157.9, 156.3, 137.9, 136.4, 131.7, 130.9 (t, *J* = 4.2 Hz), 129.7, 125.0, 124.6, 124.2, 124.1, 120.2, 115.7, 114.3, 82.8, 78.7, 71.8, 70.7, 70.6, 70.5, 70.4, 70.1, 69.9 (2 peaks: 69.93, 69.88), 69.8 (2 peaks: 69.83, 69.80), 69.7, 69.6, 69.5, 69.3, 68.9 (2 peaks: 68.93, 68.85), 67.2, 58.9, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 45.00 (dq, 1F, *J* = 94.9, 32.0 Hz), 43.19 (dq, 1F, *J* = 94.9, 32.0 Hz); MS (ESI) m/z calcd for $(M\text{-}H)^-$ $C_{57}H_{72}BF_2N_2O_{17}$ 1105.49; found 1105.50.


O
CO2H
6
N
B
N
F2
O
3
O
O
3
1H NMR (CDCl3)
9
8
7
6
5
4
3
2
1
ppm

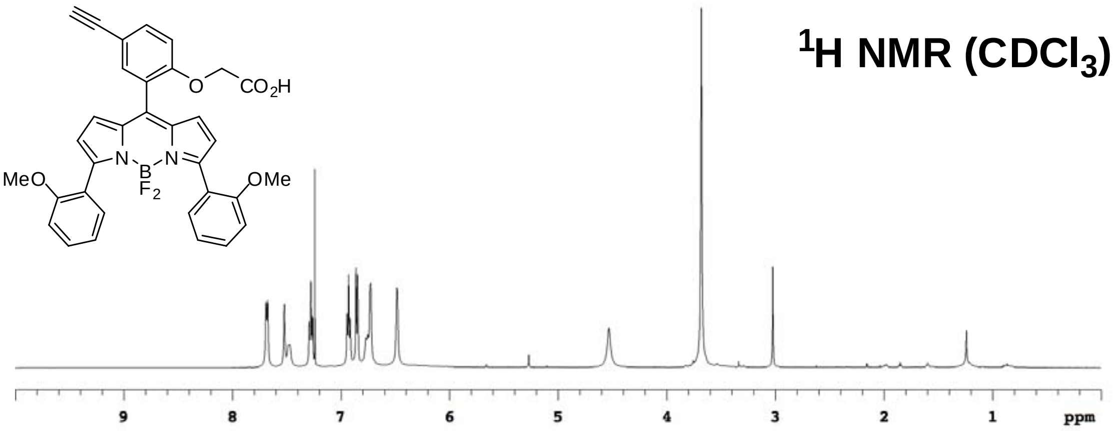

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm

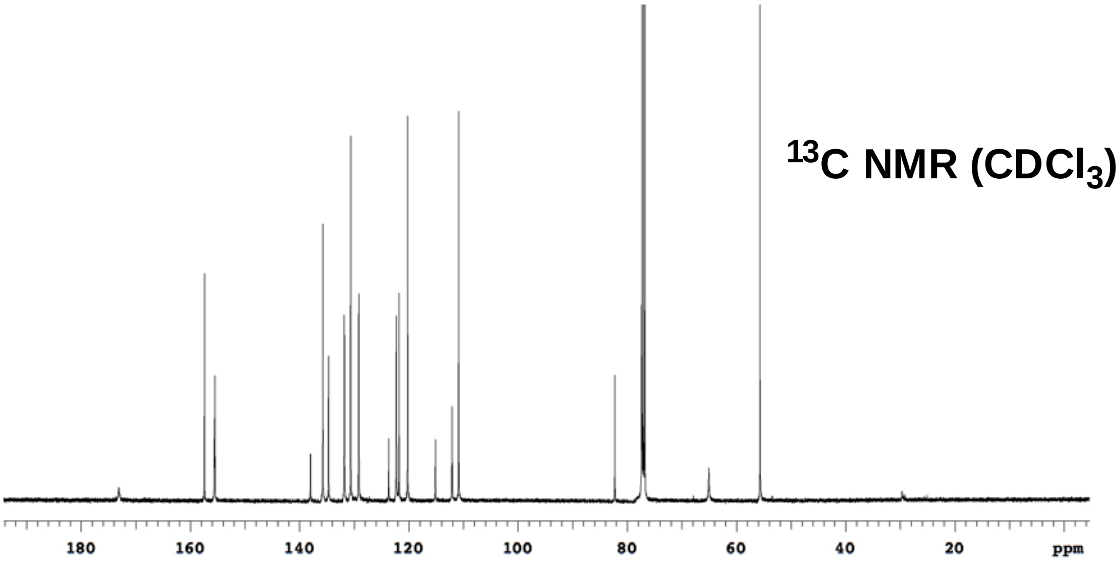

19F NMR (CDCl3)
52
50
48
46
44
42
40
38
ppm

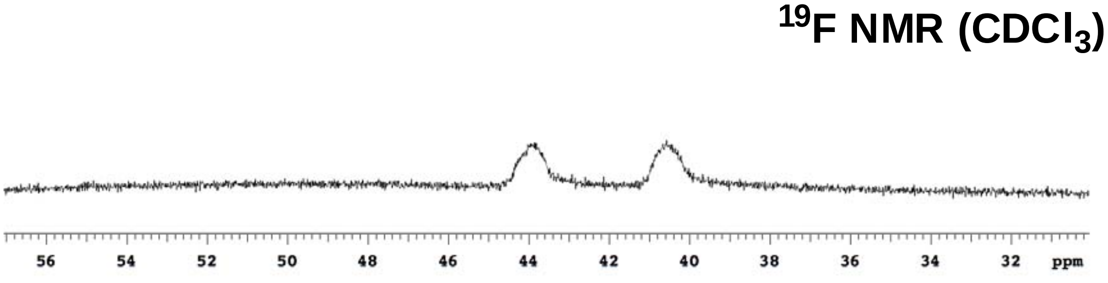


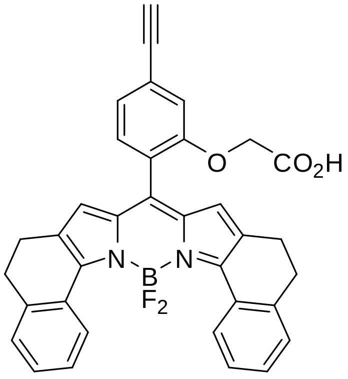
37d

Red solid (83 mg, 95 %), prepared from **48d** (100 mg, 0.15 mmol). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.68 (d, 2H, *J* = 7.3 Hz), 7.52-7.47 (m, 2H), 7.29-7.26 (m, 2H), 6.94-6.91 (m, 2H), 6.85 (d, 2H, *J* = 8.4 Hz), 6.77-6.73 (m, 3H), 6.48 (br, 2H), 4.53 (s, 2H), 3.68 (s, 6H), 3.02 (s, 1H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.1, 157.4, 155.6, 155.5, 138.0, 135.7, 134.7, 131.8, 130.6, 129.2, 123.7, 122.3, 121.8, 120.2, 115.1, 112.1, 110.9, 82.3, 77.4, 65.1, 55.7, (one carbon is ambiguous, some peak accounts for two carbons); ^{19}F NMR (282 MHz, $CDCl_3$) δ 43.97 (br, 1F), 40.56 (br, 1F); HRMS (ESI) m/z calcd for $(M-H)^-$ $C_{33}H_{24}BF_2N_2O_5$ 577.1746; found 577.1743.


MeO
OMe
N
B
N
F2
O
CO2H
¹H NMR (CDCl₃)
9
8
7
6
5
4
3
2
1
ppm

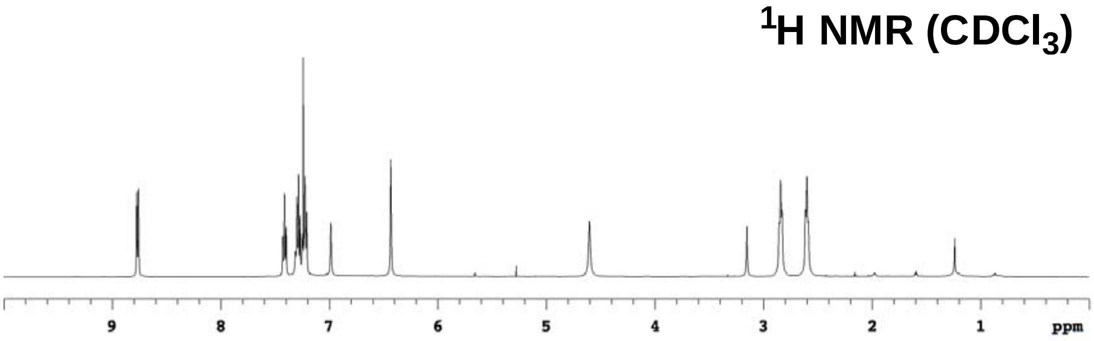

¹³C NMR (CDCl₃)
180
160
140
120
100
80
60
40
20
ppm


¹⁹F NMR (CDCl₃)
56
54
52
50
48
46
44
42
40
38
36
34
32
ppm

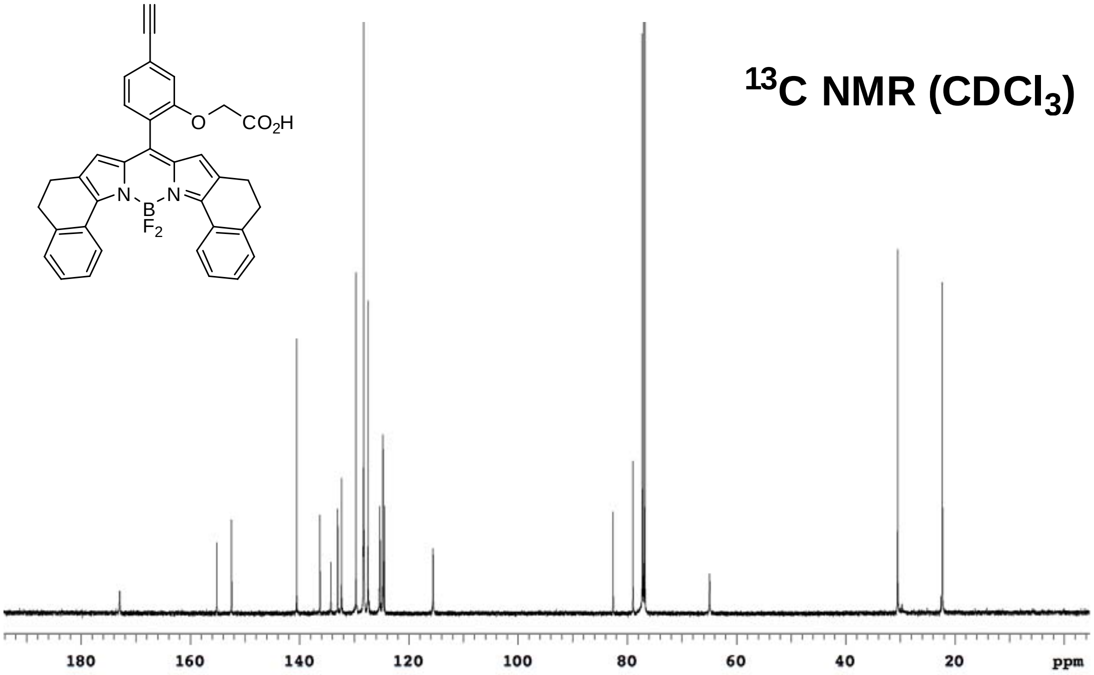


37e

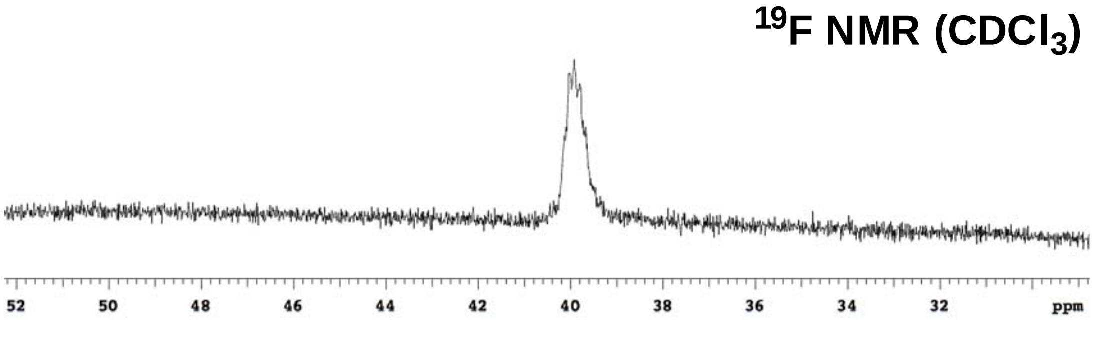
Blue solid (36 mg, 80 %), prepared from **48e** (52 mg, 0.08 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.77 (d, 2H, *J* = 8.1 Hz), 7.43-7.40 (m, 2H), 7.31-7.27 (m, 3H), 7.24-7.21 (m, 3H), 6.99 (s, 1H), 6.43 (s, 2H), 4.60 (s, 2H), 3.15 (s, 1H), 2.84 (t, 4H, *J* = 7.0 Hz), 2.60 (t, 4H, *J* = 7.0 Hz); ^{13}C NMR (125 MHz, CDCl$_3$) δ 172.9, 155.2, 152.5, 140.5, 136.3, 134.3, 133.0, 132.3, 129.7, 128.4, 128.3, 128.2, 127.5, 125.4, 125.2, 124.7, 124.4, 115.5, 82.6, 78.9, 64.9, 30.5, 22.3; ^{19}F NMR (282 MHz, CDCl$_3$) δ 40.14-39.64 (m); MS (ESI) m/z calcd for (M-H)$^-$ $C_{35}H_{24}BF_2N_2O_3$ 569.18; found 569.17.



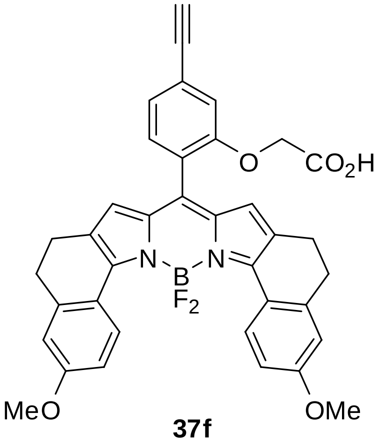
^{1}H NMR ($CDCl_3$)



^{13}C NMR ($CDCl_3$)

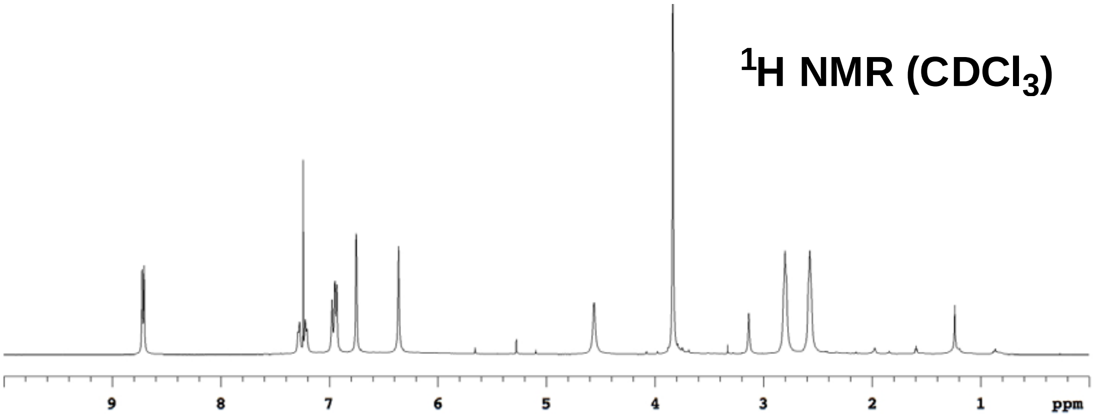


^{19}F NMR ($CDCl_3$)

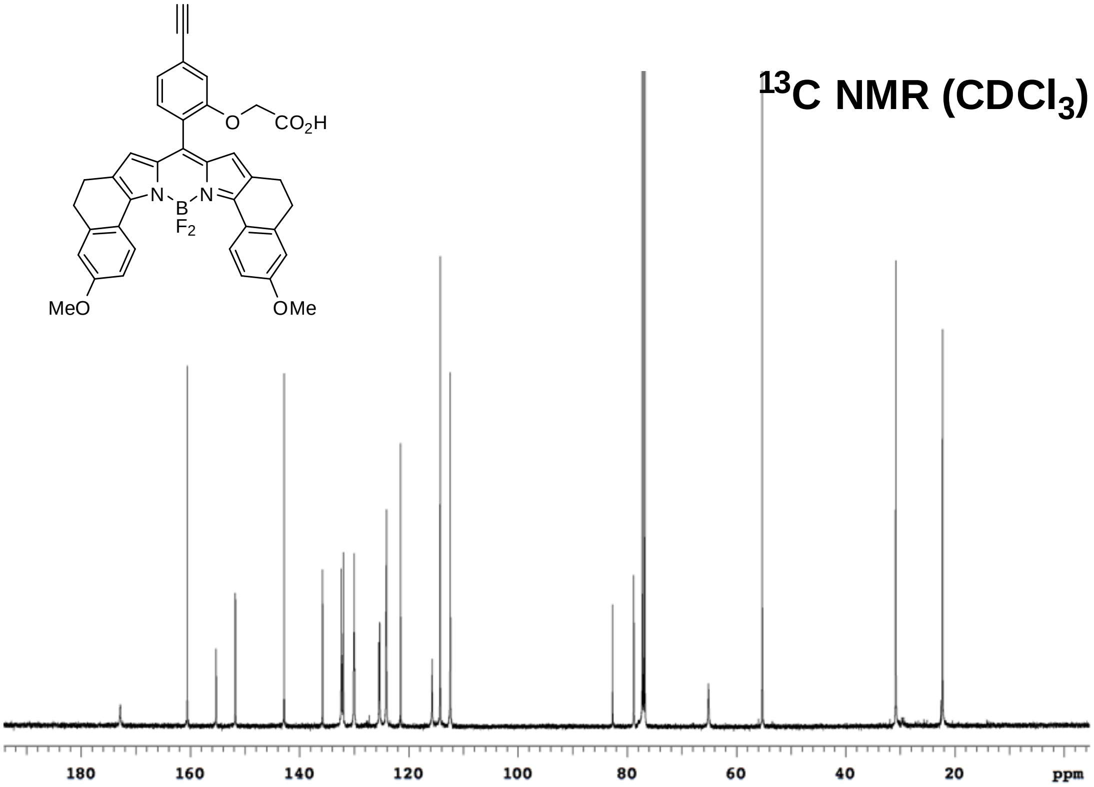


37f

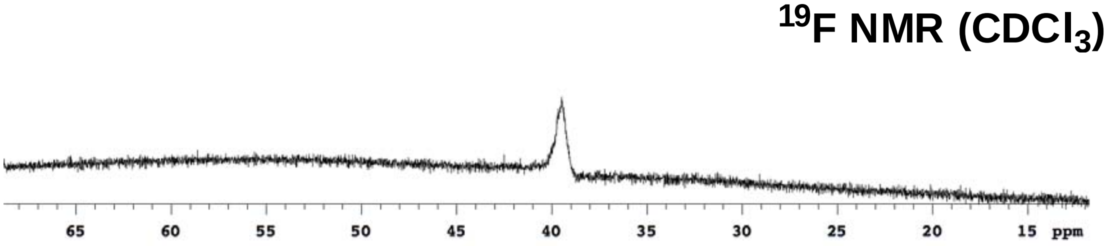
Blue solid (63 mg, *quant.*), prepared from **48f** (72 mg, 0.1 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.72 (d, 2H, *J* = 8.9 Hz), 7.29-7.27 (m, 1H), 7.22-7.21 (m, 1H), 6.98-6.93 (m, 3H), 6.75 (br, 2H), 6.36 (br, 2H), 4.56 (s, 2H), 3.83 (s, 6H), 3.14 (s, 1H), 2.82-2.79 (m, 4H), 2.59-2.56 (m, 4H); ^{13}C NMR (125 MHz, CDCl$_3$) δ172.8, 160.6, 155.3, 151.8, 142.8, 135.8, 132.4, 132.2, 132.0, 130.0 (t, *J* = 10.9 Hz), 125.5, 125.3, 124.2, 124.1, 121.5, 115.7, 114.2, 112.4, 82.7, 78.8, 65.1, 55.3, 30.8, 22.3; ^{19}F NMR (282 MHz, CDCl$_3$) δ 39.50 (br); MS (ESI) m/z calcd for (M-H)$^-$ $C_{37}H_{28}BF_2N_2O_5$ 629.21; found 629.18.



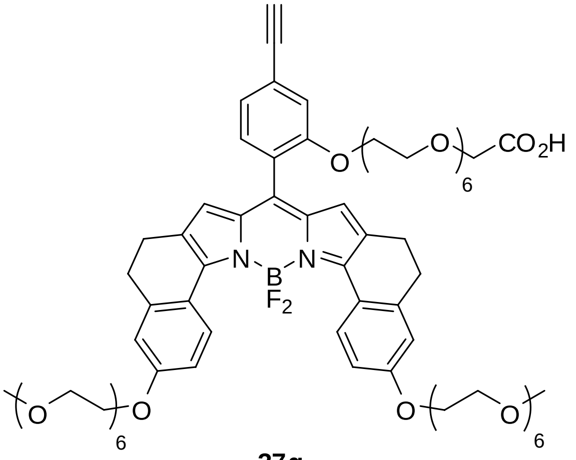
^{1}H NMR (CDCl$_3$)
9
8
7
6
5
4
3
2
1
ppm



CO$_2$H
O
N
B
N
F$_2$
MeO
OMe
^{13}C NMR (CDCl$_3$)
180
160
140
120
100
80
60
40
20
ppm

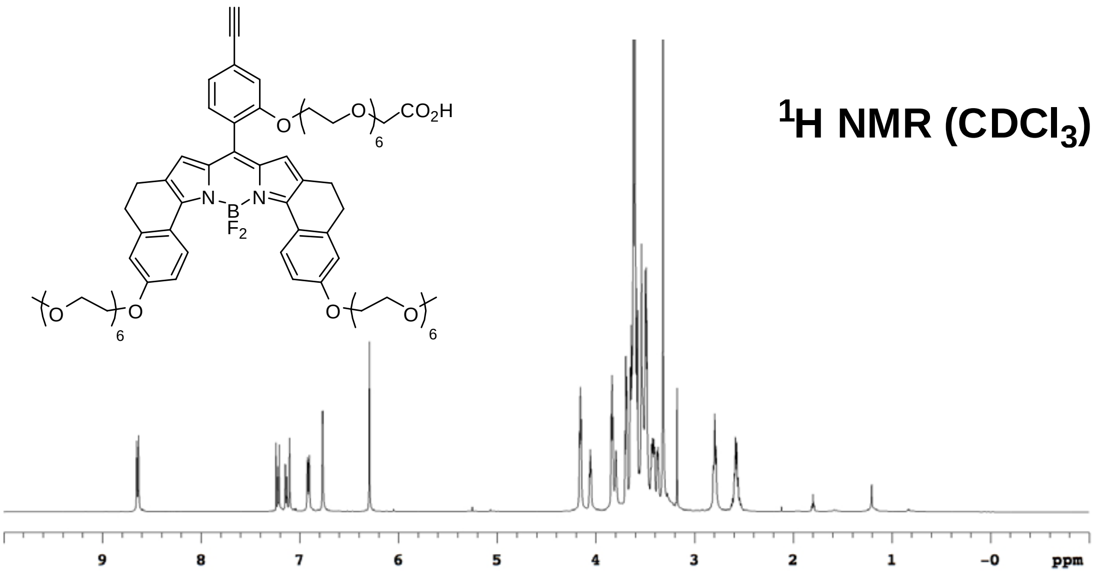


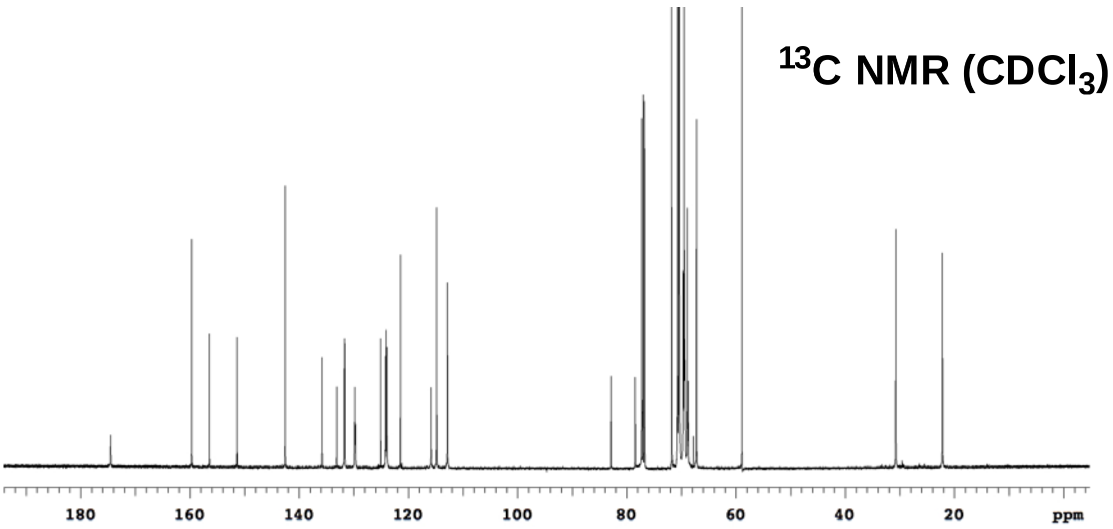
^{19}F NMR (CDCl$_3$)
65
60
55
50
45
40
35
30
25
20
15
ppm

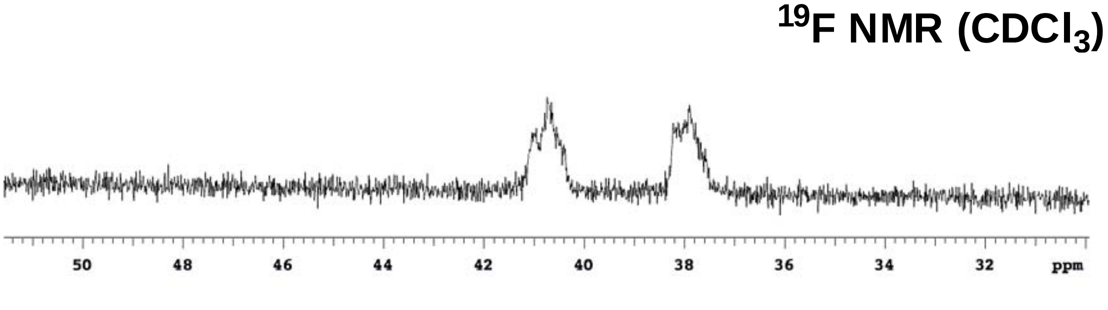


37g

Green sticky solid (315 mg, 80 %), prepared from **48g** (416 mg, 0.28 mmol). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.64 (d, 2H, *J* = 9.1 Hz), 7.22 (d, 1H, *J* = 7.7 Hz), 7.14 (d, 1H, *J* = 7.7 Hz), 7.10 (s, 1H), 6.91 (dd, 2H, *J* = 9.1, 2.5 Hz), 6.77 (d, 2H, *J* = 2.5 Hz), 6.29 (s, 2H), 4.17-4.15 (m, 4H), 4.06-4.05 (m, 2H), 3.85-3.83 (m, 4H), 3.80 (s, 2H), 3.70-3.37 (m, 62H), 3.32 (s, 6H), 3.18 (s, 1H), 2.81-2.78 (m, 4H), 2.60-2.56 (m, 4H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 174.5, 159.7, 156.4, 151.4, 142.5, 135.8, 133.1, 131.7, 131.6, 129.8 (t, *J* = 11.0 Hz), 125.0, 124.2, 124.1, 123.9, 121.4, 115.8, 114.8, 112.8, 82.9, 78.4,71.8, 70.7, 70.5, 70.4 (4 peaks: 70.43, 70.43, 70.40, 70.35), 69.7, 69.5 (2 peaks: 69.53, 69.48), 69.4 (2 peaks: 69.44, 69.35), 68.9, 68.7, 67.2, 58.9, 30.7, 22.2, (several peaks account for more than one carbon); ^{19}F NMR (282 MHz, CDCl$_3$) δ 41.09-40.38 (m, 1F), 38.23-37.58 (m, 1F); MS (MALDI) m/z calcd for $(M+Na)^+$ $C_{73}H_{101}BF_2NaN_2O_{23}$ 1445.68; found 1445.79.


O
CO2H
6
N
B
N
F2
O
O
6
6
1H NMR (CDCl3)
9
8
7
6
5
4
3
2
1
−0
ppm

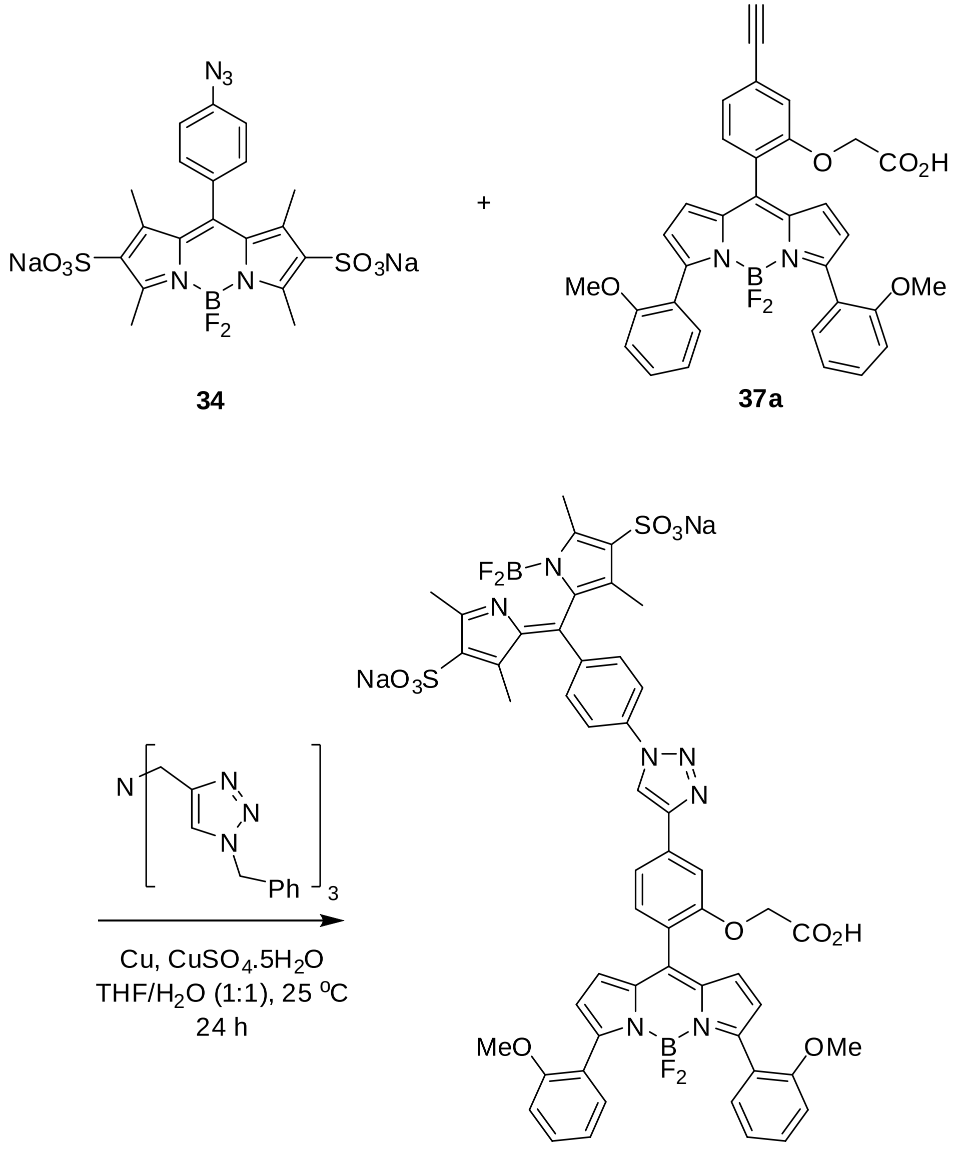

13C NMR (CDCl3)
180
160
140
120
100
80
60
40
20
ppm


19F NMR (CDCl3)
50
48
46
44
42
40
38
36
34
32
ppm

Syntheses of through-bond energy transfer cassettes

(1) Syntheses of cassettes 31a-g

Typical procedures:

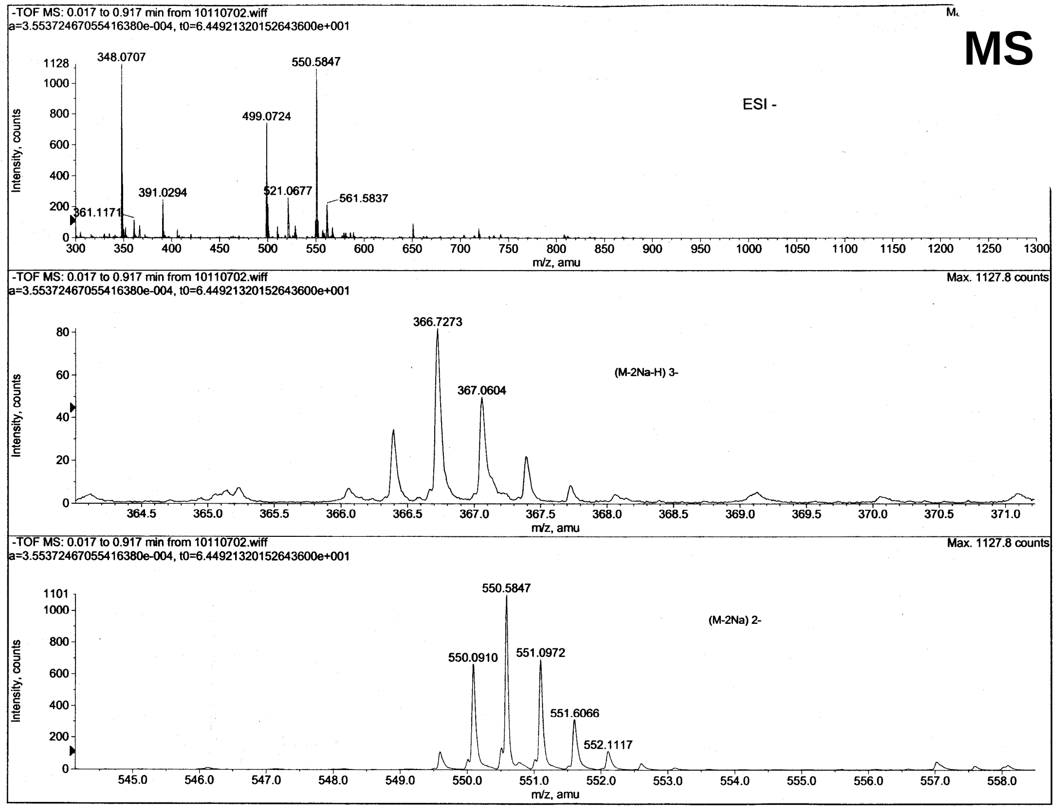

N3
NaO3S
SO3Na
N
B
F2
+
CO2H
O
MeO
OMe
Ph
3
Cu, CuSO4.5H2O
THF/H2O (1:1), 25 °C
24 h

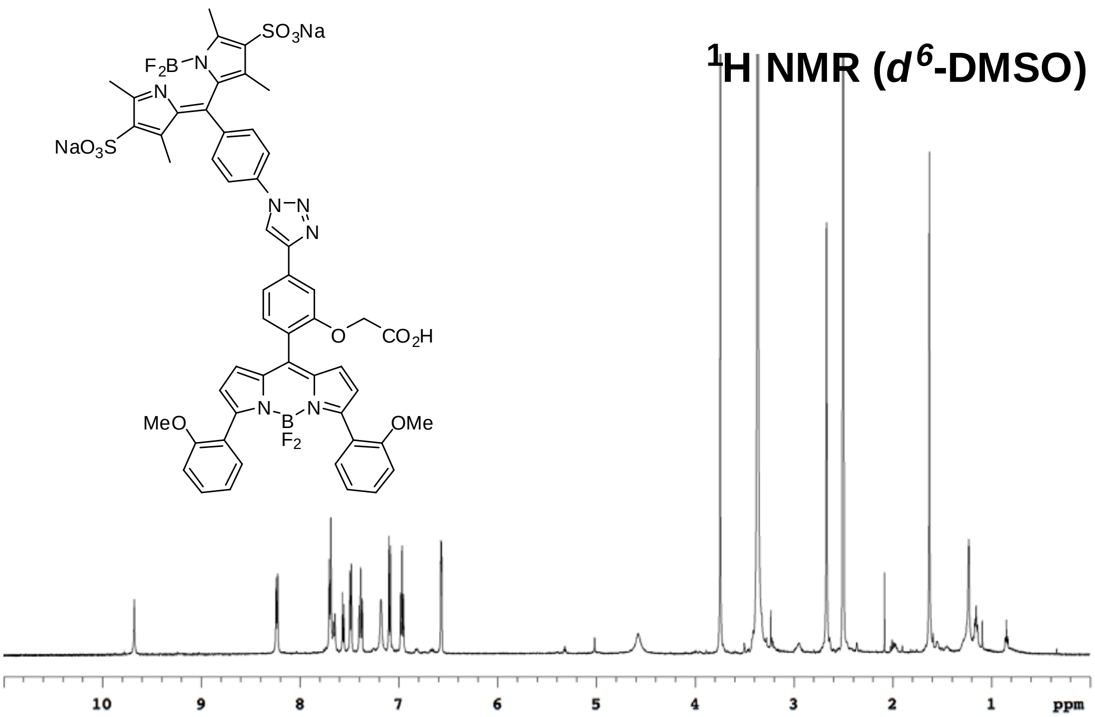
34

37a

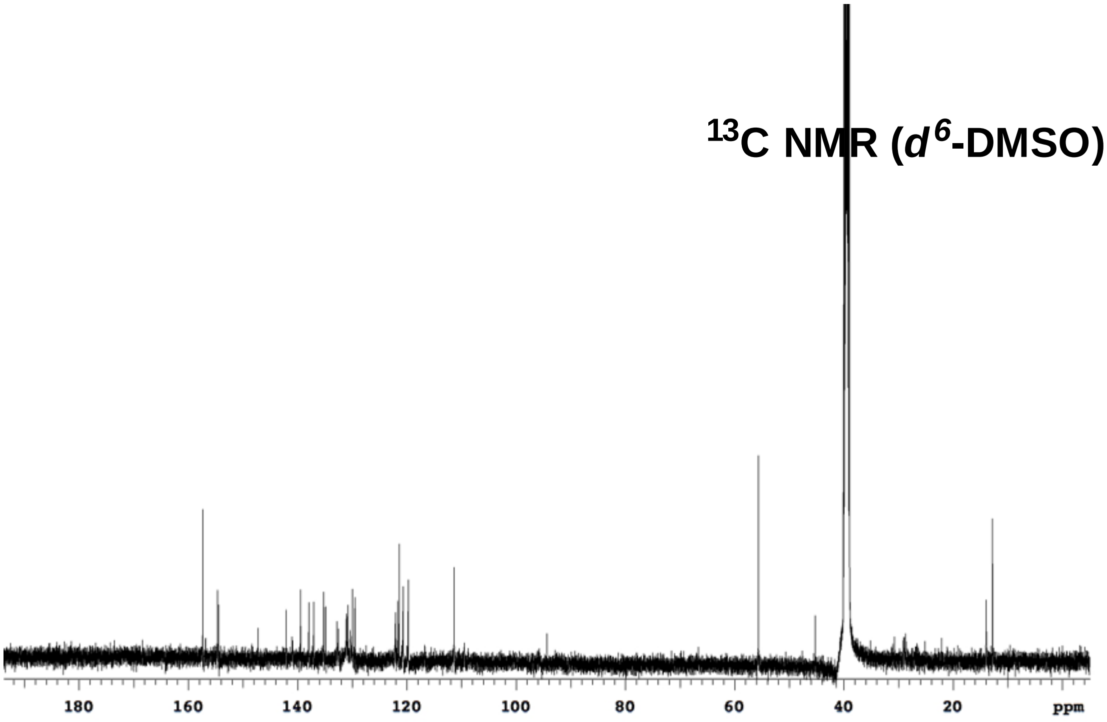
31a 75 %

BODIPY donor **34** (6 mg, 0.01 mmol), BODIPY acceptor **37a** (11 mg, 0.02 mmol), TBTA ligand (0.5 mg, 0.001 mmol), Cu (0.6 mg, 0.01 mmol) and $CuSO_4{\cdot}5H_2O$ (0.3 mg, 0.001 mmol) were dissolved in 5 mL THF + 1 mL H_2O and stirred at room temperature for 24 h. The solvents were removed under reduced pressure and the residue was purified by MPLC on reverse phase silica gel (H_2O to 50 % CH_3CN/H_2O) to afford the product (9 mg, 75 %) as red solid. ^{1}H NMR (500 MHz, d^6-DMSO) δ 9.68 (s, 1H), 8.23 (d, 2H, *J* = 8.4 Hz), 7.70-7.65 (m, 4H), 7.56 (d, 1H, *J* = 7.7 Hz), 7.49 (d, 2H, *J* = 7.7 Hz), 7.40-7.37 (m, 2H), 7.18 (br, 2H), 7.09 (d, 2H, *J* = 8.4 Hz), 6.98-6.95 (m, 2H), 6.57 (d, 2H, *J* = 4.2 Hz), 4.58 (s, 2H), 3.74 (s, 6H), 2.67 (s, 6H), 1.63 (s, 6H); ^{13}C NMR (125 MHz, d^6-DMSO) δ 157.3, 154.6, 154.4, 147.2, 142.1, 139.4, 137.9, 137.1, 135.2, 134.8, 132.8, 132.5, 131.1, 130.8, 130.4, 130.0, 129.5, 122.1, 121.7, 121.4, 120.7, 120.6, 119.7, 111.3, 94.4, 55.6, 45.3, 13.9, 12.8, (due to poor solubility, several carbon did not show up); MS (ESI) calcd for $(M\text{-}2Na)^{2-}$ $C_{52}H_{41}B_2F_4N_7O_{11}S_2$ m/z 550.62, found 550.58; calcd for $(M\text{-}2Na\text{-}H)^{3-}$ $C_{52}H_{40}B_2F_4N_7O_{11}S_2$ m/z 366.74, found 366.73.

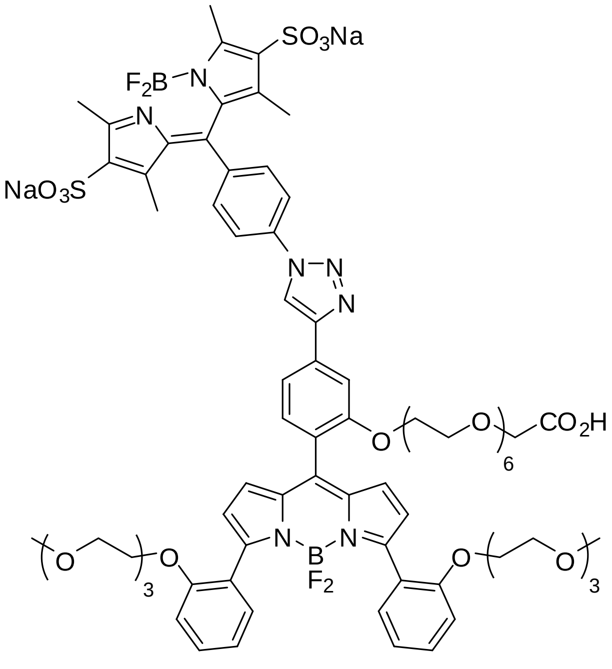




^{1}H NMR (d^{6}-DMSO)

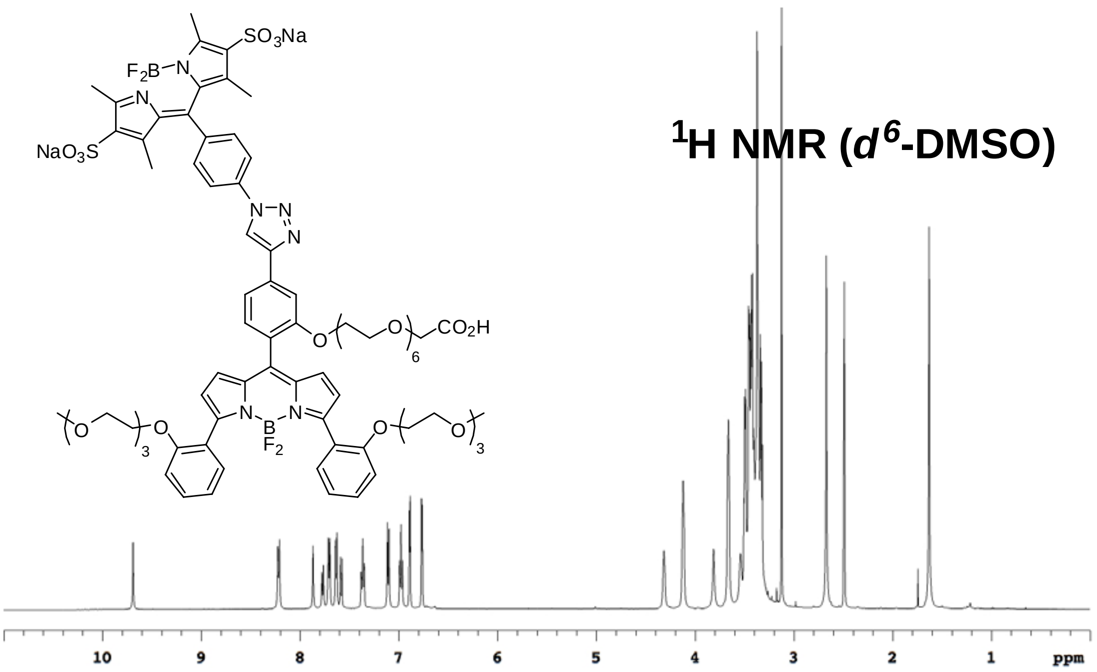


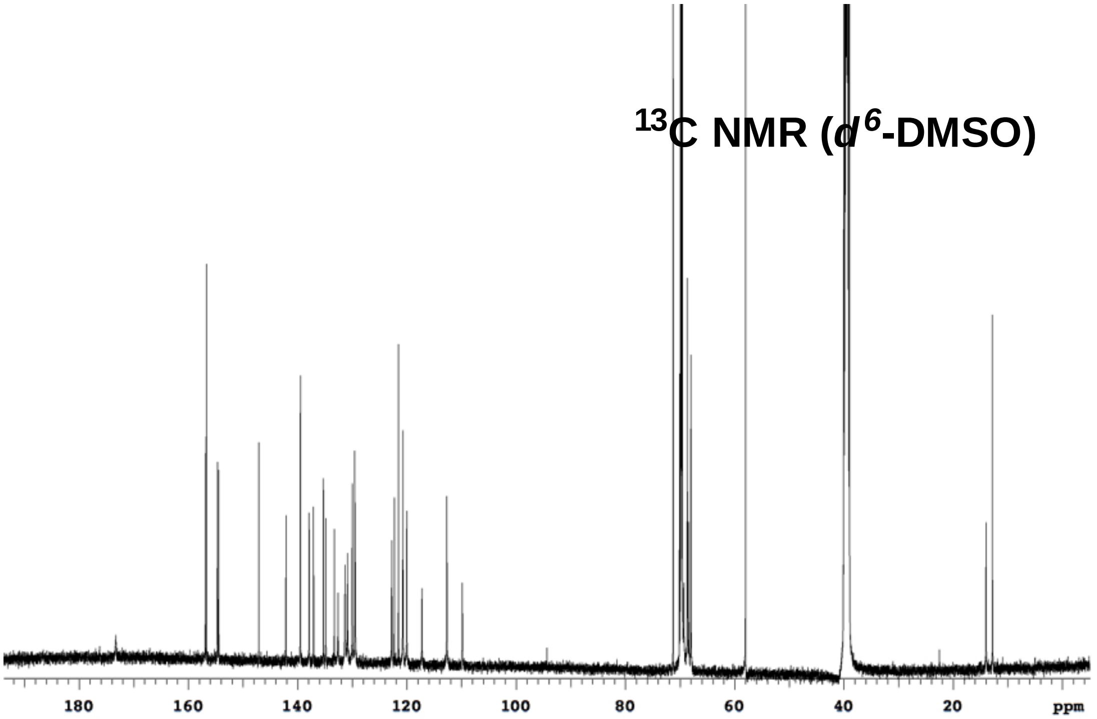
^{13}C NMR (d^{6}-DMSO)

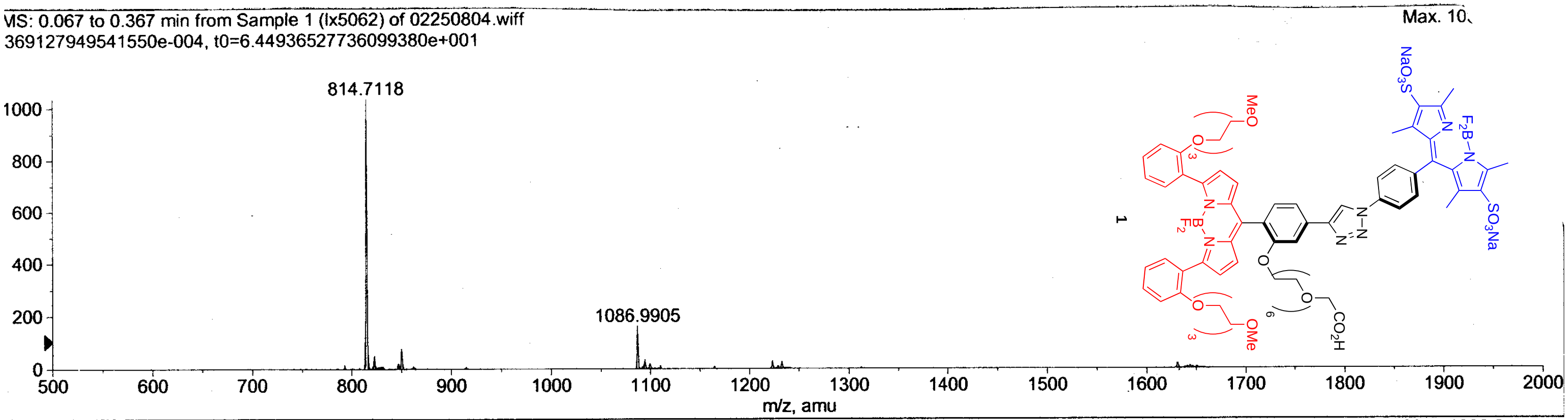


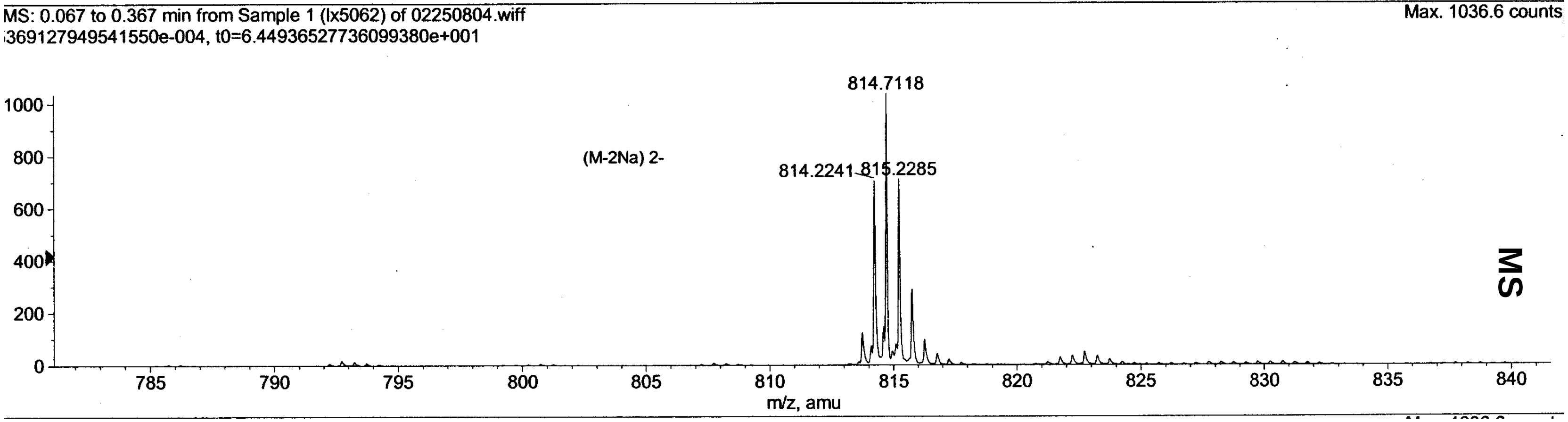
31b

Red solid (46 mg, 78 %), prepared from donor **34** (20 mg, 0.035 mmol) and BODIPY acceptor **37b** (43 mg, 0.039 mmol). ^{1}H NMR (500 MHz, *d*^{6}-DMSO) δ 9.69 (s, 1H), 8.22 (d, 2H, *J* = 8.4 Hz), 7.87 (s, 1H), 7.77 (d, 1H, *J* = 7.9 Hz), 7.71 (d, 2H, *J* = 8.2 Hz), 7.63 (d, 2H, *J* = 7.5 Hz), 7.58 (d, 1H, *J* = 7.9 Hz), 7.38-7.35 (m, 2H), 7.11 (d, 2H, *J* = 8.4 Hz), 6.99-6.96 (m, 2H), 6.89 (d, 2H, *J* = 4.3 Hz), 6.77 (d, 2H, *J* = 4.3 Hz), 4.32 (s, 2H), 4.12 (br, 4H), 3.81 (s, 2H), 3.66 (br, 6H), 3.54-3.32 (m, 36H), 3.12 (s, 6H), 2.67 (s, 6H), 1.63 (s, 6H); ^{13}C NMR (125 MHz, *d*^{6}-DMSO) δ 173.3, 156.9, 156.7, 154.7, 154.4, 147.1, 142.1, 139.5, 137.9, 137.1, 135.3, 134.8, 133.3, 132.6, 131.3, 130.8, 130.0, 129.5, 129.4, 122.7, 122.3, 121.5, 120.7, 120.0, 117.2, 112.7, 109.8, 71.3, 70.1, 69.9, 69.8 (2 peaks: 69.81, 69.78), 69.7 (3 peaks: 69.71, 69.69, 69.66), 69.6, 69.3, 68.7, 68.6, 68.4, 68.0, 58.0, 14.0, 12.8. (7C are ambiguous, several peaks account for more than one carbon). MS (ESI) m/z calcd for $(M\text{-}2Na)^{2-}$ $C_{76}H_{89}B_2F_4N_7O_{23}S_2$ M/Z: 814.78; found 814.71.

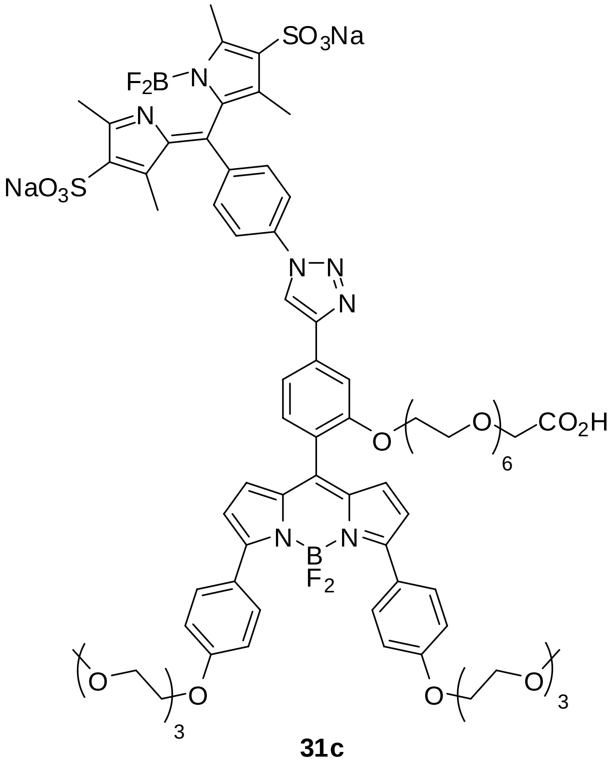

SO3Na
F2B
N
N
NaO3S
N
N
N
O
CO2H
6
O
N
B
F2
N
O
3
O
O
3
1H NMR (d6-DMSO)
10
9
8
7
6
5
4
3
2
1
ppm


13C NMR (d6-DMSO)
180
160
140
120
100
80
60
40
20
ppm


MS: 0.067 to 0.367 min from Sample 1 (lx5062) of 02250804.wiff
369127949541550e-004, t0=6.44936527736099380e+001
Max. 10
814.7118
1086.9905
1000
800
600
400
200
0
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
m/z, amu
NaO$_3$S
SO$_3$Na
F$_2$B
N
MeO
OMe
O
3
6
CO$_2$H
F$_2$
1

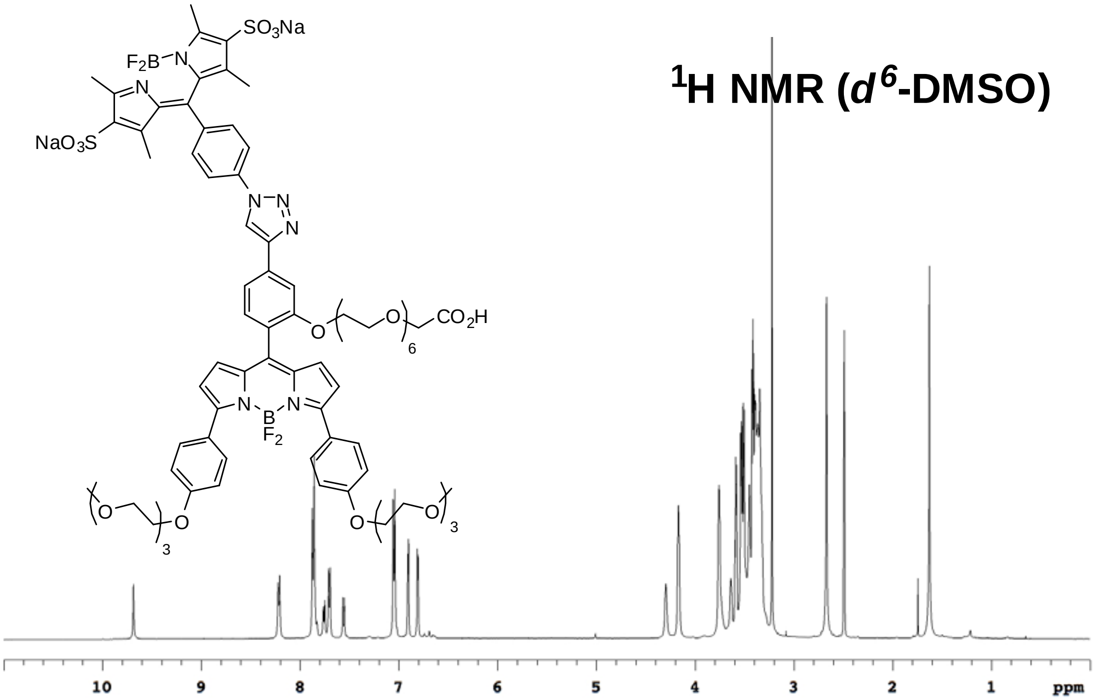

MS: 0.067 to 0.367 min from Sample 1 (lx5062) of 02250804.wiff
369127949541550e-004, t0=6.44936527736099380e+001
Max. 1036.6 counts
814.7118
(M-2Na) 2-
814.2241
815.2285
1000
800
600
400
200
0
785
790
795
800
805
810
815
820
825
830
835
840
m/z, amu

MS

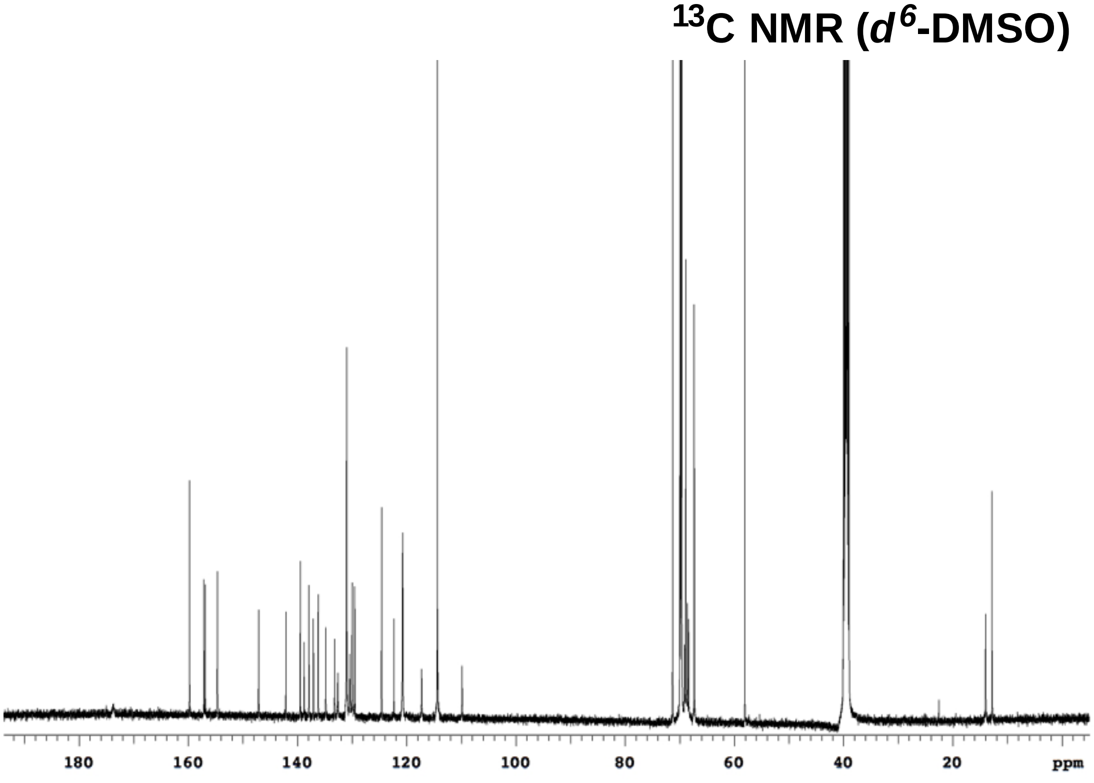


31c

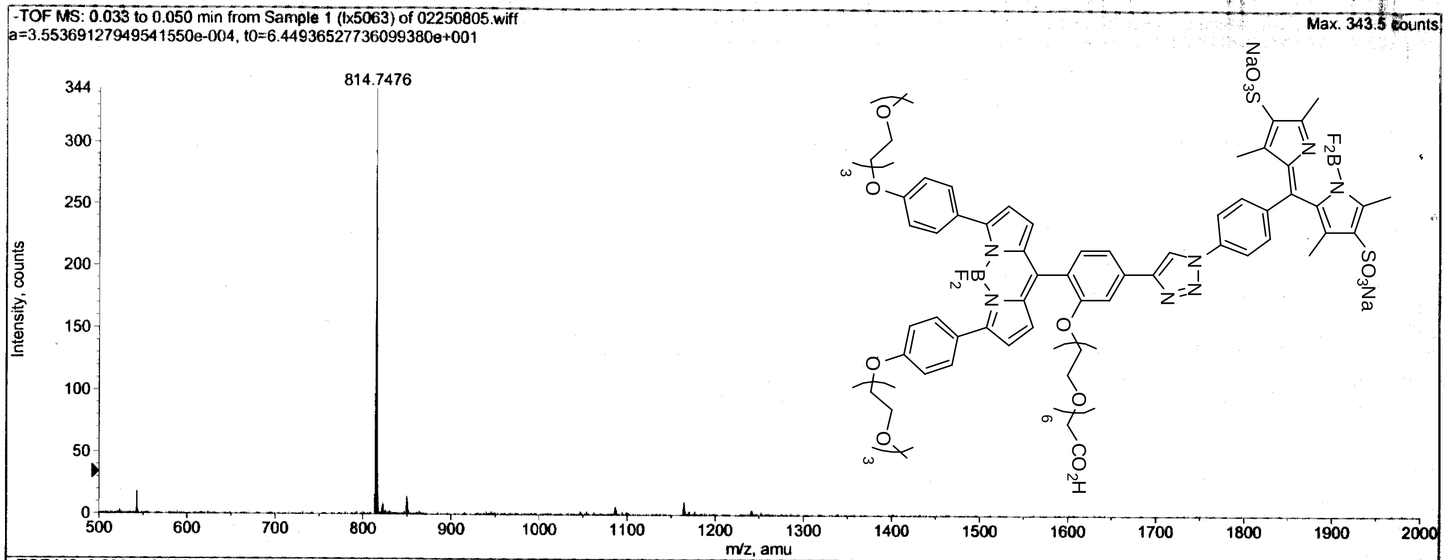
Blue solid (54 mg, 80 %), prepared from donor **34** (23 mg, 0.040 mmol) and BODIPY acceptor **37c** (49 mg, 0.044 mmol). ^{1}H NMR (500 MHz, d^6-DMSO) δ 9.69 (s, 1H), 8.21 (d, 2H, *J* = 8.2 Hz), 7.87-7.83 (m, 5H), 7.76 (d, 1H, *J* = 7.8 Hz), 7.70 (d, 2H, *J* = 8.2 Hz), 7.56 (d, 1H, *J* = 7.8 Hz), 7.05 (d, 4H, *J* = 8.8 Hz), 6.90 (d, 2H, *J* = 4.3 Hz), 6.81 (d, 2H, *J* = 4.3 Hz), 4.30 (s, 2H), 4.18-4.16 (m, 4H), 3.77-3.75 (m, 6H), 3.64 (br, 2H), 3.60-3.35 (m, 36H), 3.22 (s, 6H), 2.67 (s, 6H), 1.63 (s, 6H); ^{13}C NMR (125 MHz, d^6-DMSO) δ 173.7, 159.8, 157.1, 156.9, 154.7, 147.1, 142.1, 139.5, 138.8, 137.9, 137.1, 136.2, 134.8, 133.2, 132.6, 131.0, 130.4, 130.0, 129.5, 124.6, 122.3, 120.8, 120.7, 117.3, 114.4, 109.9, 71.3, 70.0 (2 peaks: 70.01, 69.96), 69.8 (2 peaks: 69.82, 69.76), 69.7, 69.6 (2 peaks: 69.64, 69.62), 69.1, 68.9, 68.6, 68.4, 67.3, 58.1, 14.0, 12.8, (several peaks account for more than one carbon); MS (ESI) calcd for $(M\text{-}2Na)^{2-}$ $C_{76}H_{89}B_2F_4N_7O_{23}S_2$ m/z 814.78; found 814.75.

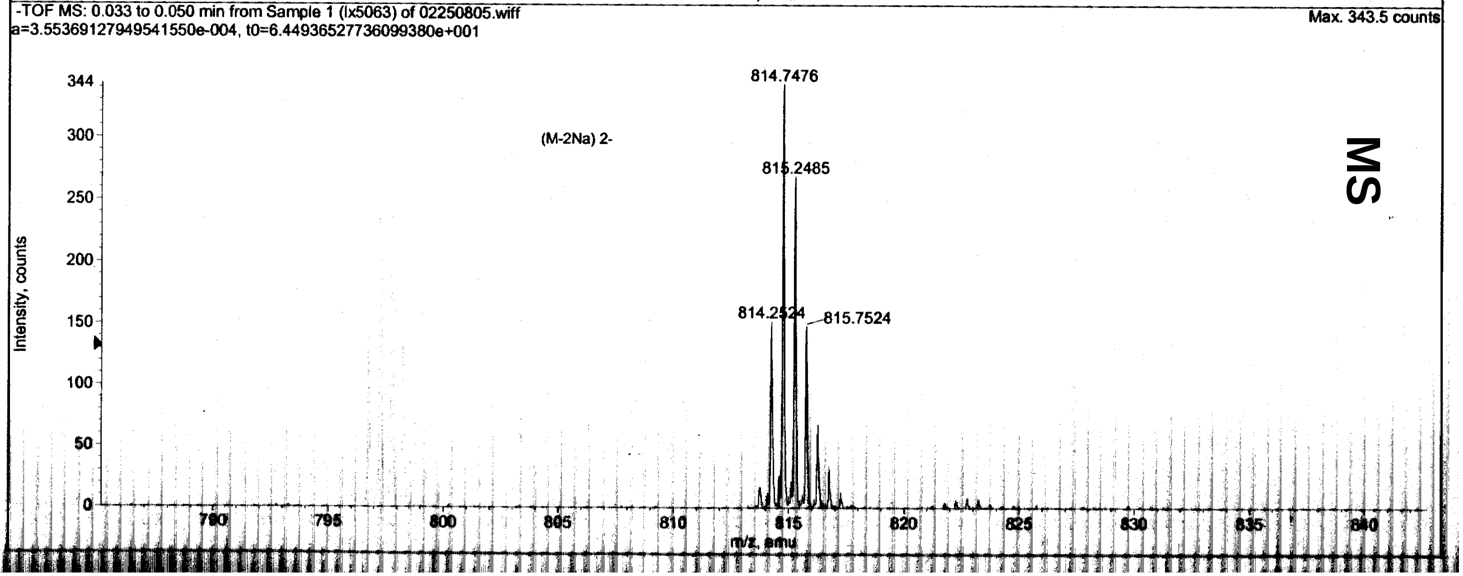


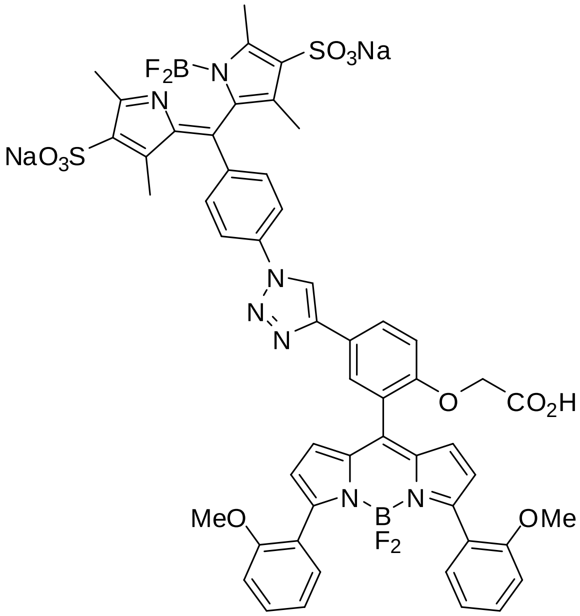
^{1}H NMR (d^6-DMSO)



^{13}C NMR (d^6-DMSO)

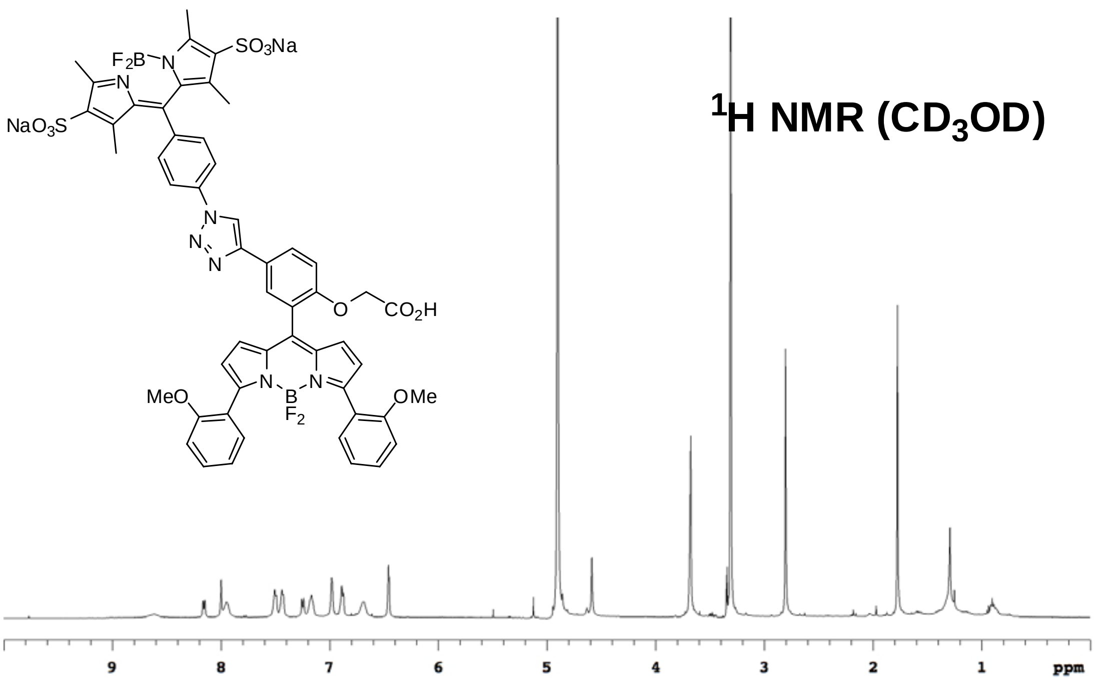

-TOF MS: 0.033 to 0.050 min from Sample 1 (lx5063) of 02250805.wiff
a=3.55369127949541550e-004, t0=6.44936527736099380e+001
Max. 343.5 counts
814.7476
Intensity, counts
344
300
250
200
150
100
50
0
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
m/z, amu
NaO3S
SO3Na
F2B
N
N-N
N
O
CO2H
6
B
F2
3
3

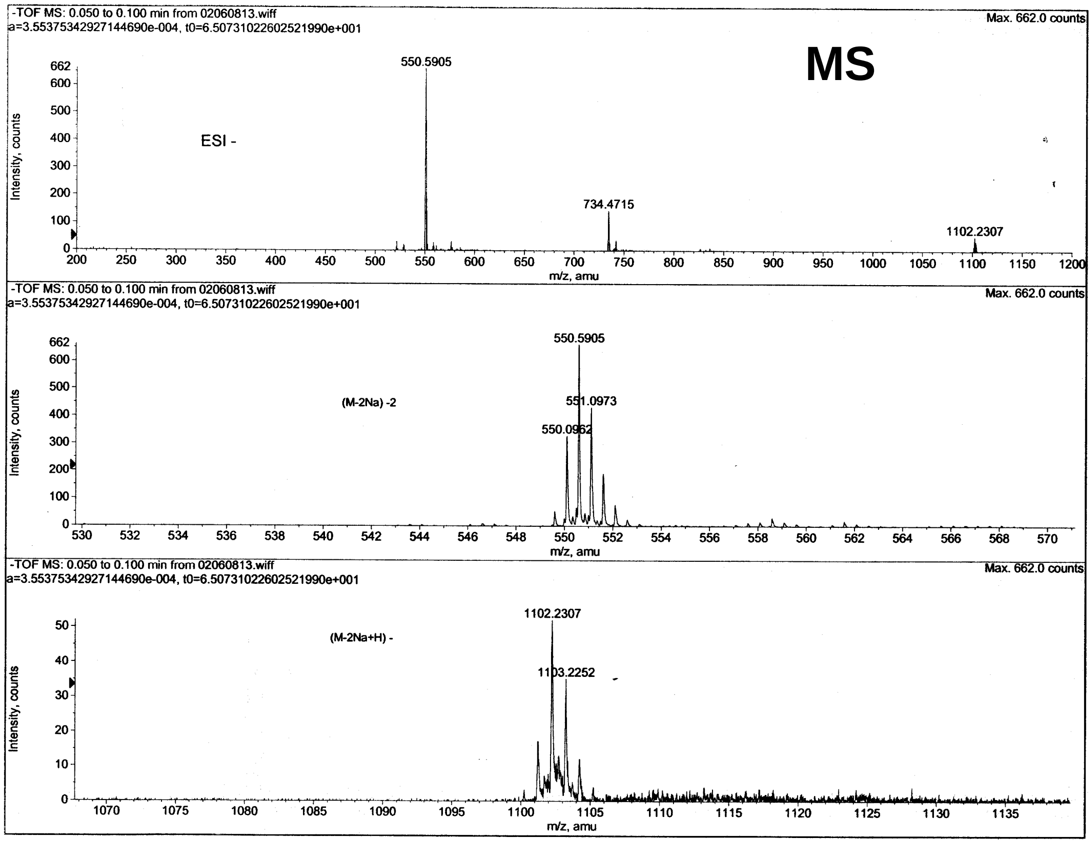

-TOF MS: 0.033 to 0.050 min from Sample 1 (lx5063) of 02250805.wiff
a=3.55369127949541550e-004, t0=6.44936527736099380e+001
Max. 343.5 counts
MS
(M-2Na) 2-
814.7476
815.2485
814.2524
815.7524
Intensity, counts
344
300
250
200
150
100
50
0
790
795
800
805
810
815
820
825
830
835
840
m/z, amu

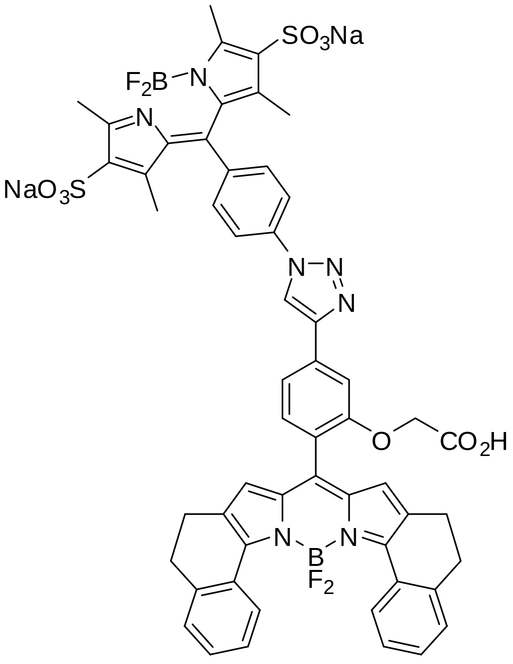


31d

Red solid (11 mg, 48 %), prepared from donor **34** (11 mg, 0.02 mmol) and BODIPY acceptor **37d** (17 mg, 0.03 mmol). ^{1}H NMR (500 MHz, CD_3OD) δ 8.61 (br, 1H), 8.17-8.15 (m, 1H), 8.00-7.95 (m, 3H), 7.51-7.43 (m, 4H), 7.25 (d, 1H, *J* = 8.7 Hz), 7.18-7.15 (m, 2H), 6.98 (br, 2H), 6.89-6.87 (m, 2H), 6.69 (br, 2H), 6.46-6.45 (m, 2H), 4.59 (s, 2H), 3.68 (s, 6H), 2.80 (s, 6H), 1.77 (s, 6H); Due to poor solubility, ^{13}C NMR with good resolution could not be obtained; MS (ESI) calcd for $(M\text{-}2Na)^{2-}$ $C_{52}H_{41}B_2F_4N_7O_{11}S_2$ m/z 550.62, found 550.59; calcd for $(M\text{-}2Na\text{+}H)^-$ $C_{52}H_{42}B_2F_4N_7O_{11}S_2$ m/z 1102.25, found 1102.23.

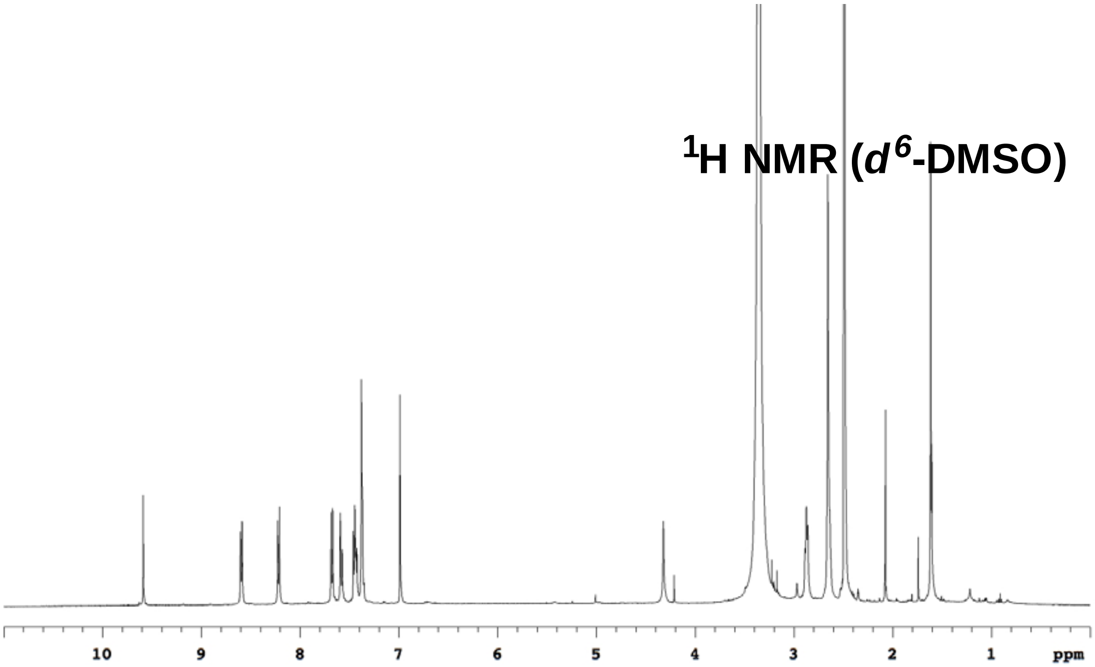

F$_2$B
SO$_3$Na
NaO$_3$S
N
CO$_2$H
MeO
OMe
B
F$_2$
^{1}H NMR (CD$_3$OD)
9
8
7
6
5
4
3
2
1
ppm

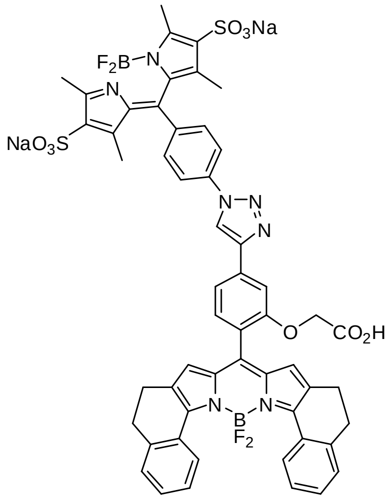

MS
-TOF MS: 0.050 to 0.100 min from 02060813.wiff
a=3.55375342927144690e-004, t0=6.50731022602521990e+001
Max. 662.0 counts
ESI -
550.5905
734.4715
1102.2307
Intensity, counts
m/z, amu
(M-2Na) -2
550.5905
551.0973
550.0962
(M-2Na+H) -
1102.2307
1103.2252

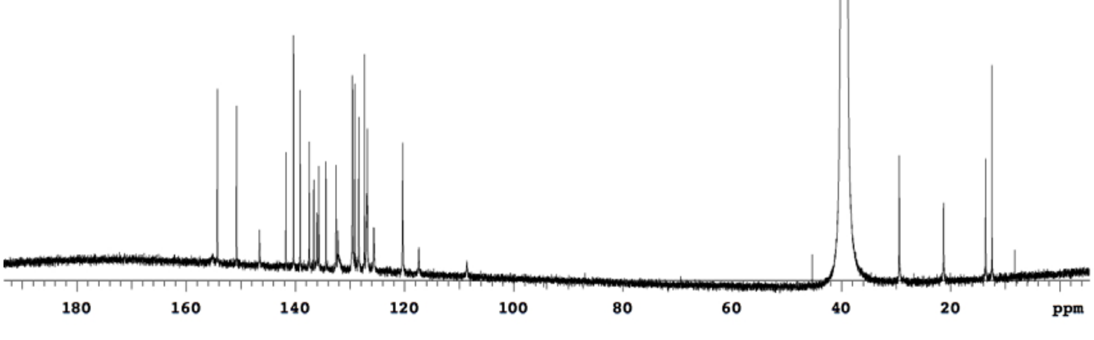


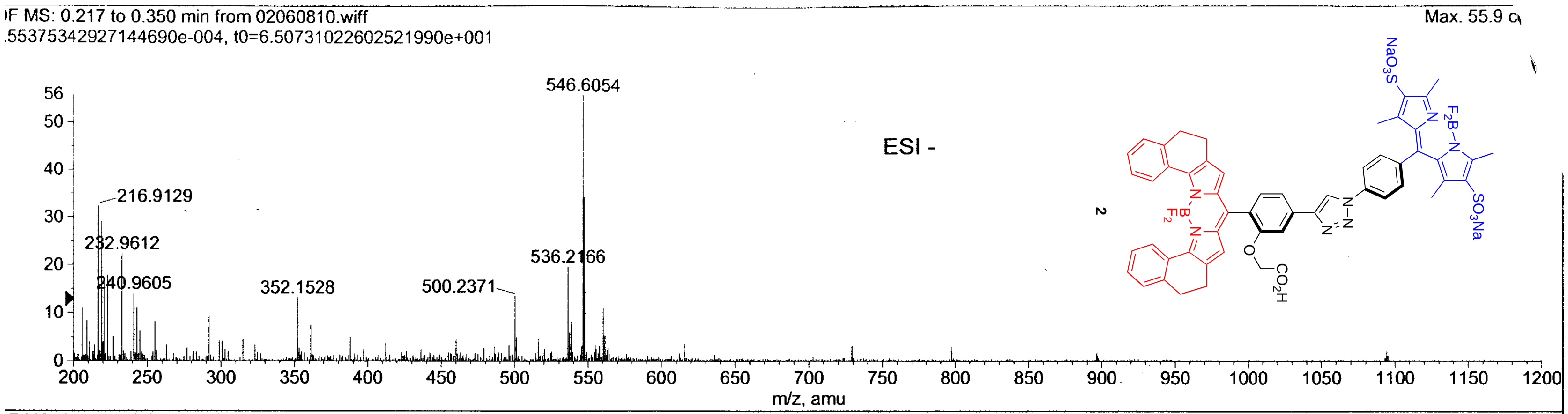
31e

Green solid (36 mg, 60 %), prepared from donor **34** (30 mg, 0.053 mmol) and BODIPY acceptor **37e** (36 mg, 0.063 mg). ^{1}H NMR (500 MHz, *d*^{6}-DMSO) δ 9.59 (s, 1H), 8.59 (d, 2H, *J* = 8.2 Hz), 8.22 (d, 2H, *J* = 8.5 Hz), 7.68 (d, 2H, *J* = 8.5 Hz), 7.59-7.57 (m, 2H), 7.46-7.42 (m, 3H), 7.39-7.35 (m, 4H), 6.99 (s, 2H), 4.32 (s, 2H), 2.89-2.86 (m, 4H), 2.66 (br, 10H), 1.61 (s, 6H); ^{13}C NMR (125 MHz, *d*^{6}-DMSO) δ 154.3, 150.8, 146.6, 141.7, 140.3, 139.1, 137.4, 136.6, 136.1, 135.7, 134.4, 132.5, 132.2, 129.5, 129.1, 128.4, 127.3, 126.9, 126.8, 125.6, 120.4, 117.4, 108.6, 45.4, 29.4, 21.3, 13.6, 12.4. (6C are ambiguous, several peaks account for more than one carbon). MS (ESI) m/z calcd for $(M\text{-}2Na)^{2-}$ $C_{54}H_{41}B_2F_4N_7O_9S_2$ M/Z: 546.63; found 546.61.

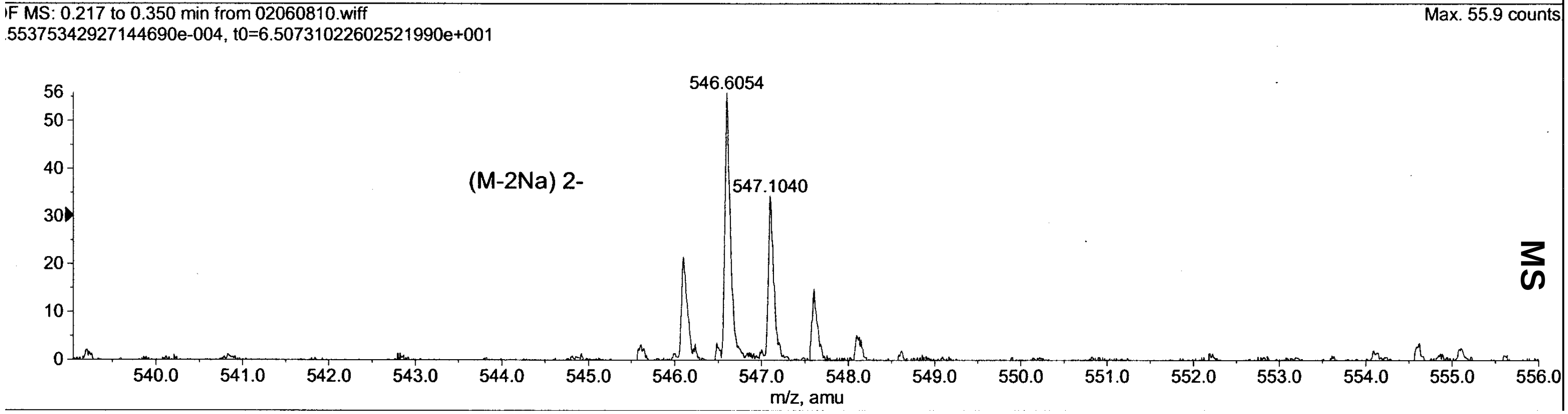

^{1}H NMR (d^6-DMSO)
10
9
8
7
6
5
4
3
2
1
ppm

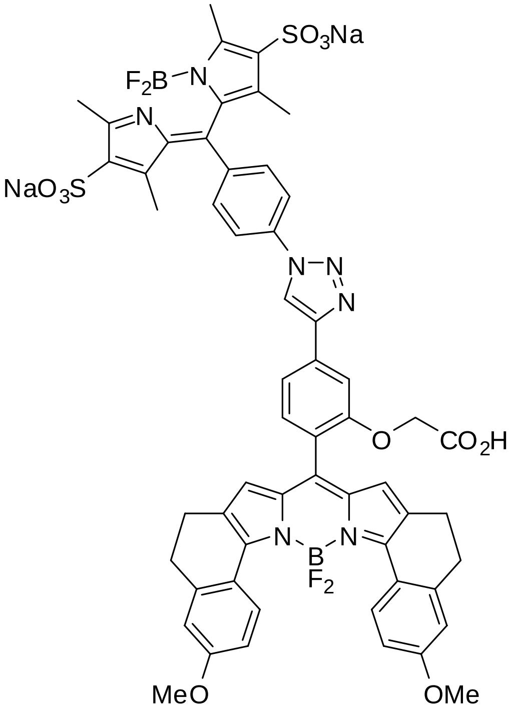

SO3Na
F2B
N
N
NaO3S
N
N
N
O
CO2H
N
B
N
F2


^{13}C NMR (d^6-DMSO)
180
160
140
120
100
80
60
40
20
ppm


ʾF MS: 0.217 to 0.350 min from 02060810.wiff
.55375342927144690e-004, t0=6.50731022602521990e+001
Max. 55.9 c
ESI -
546.6054
536.2166
500.2371
352.1528
216.9129
232.9612
240.9605
0
10
20
30
40
50
56
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
1050
1100
1150
1200
m/z, amu
NaO$_3$S
SO$_3$Na
F$_2$B
N
N-N
N
CO$_2$H
O
B
F$_2$

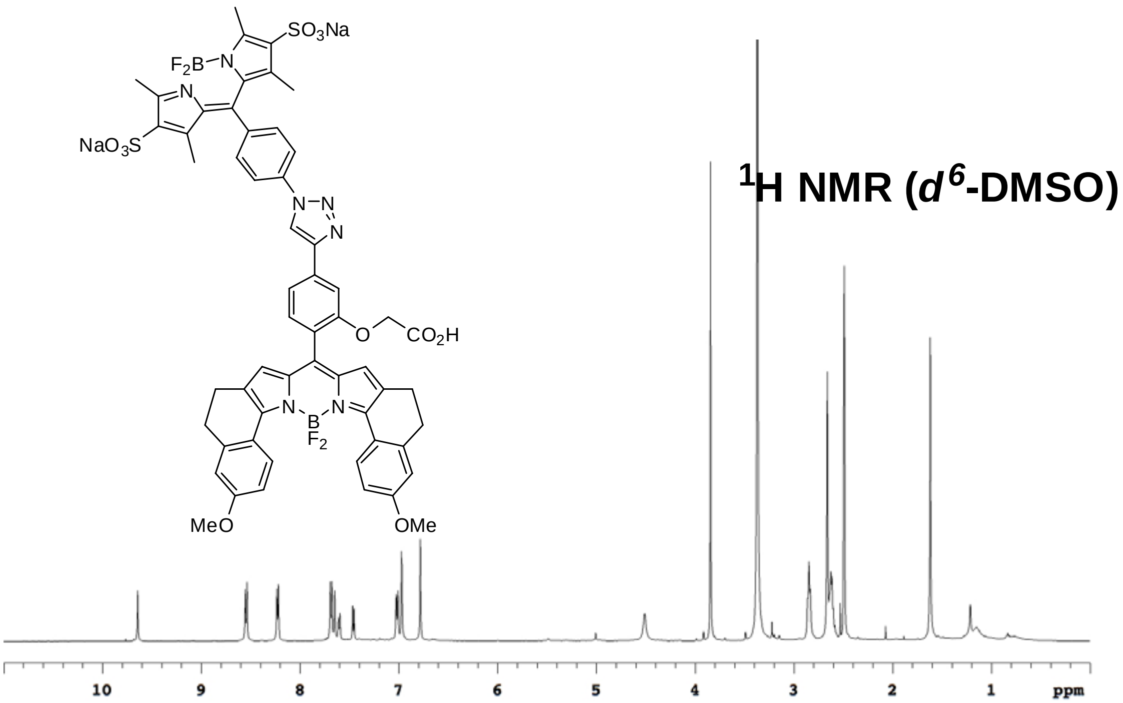
2

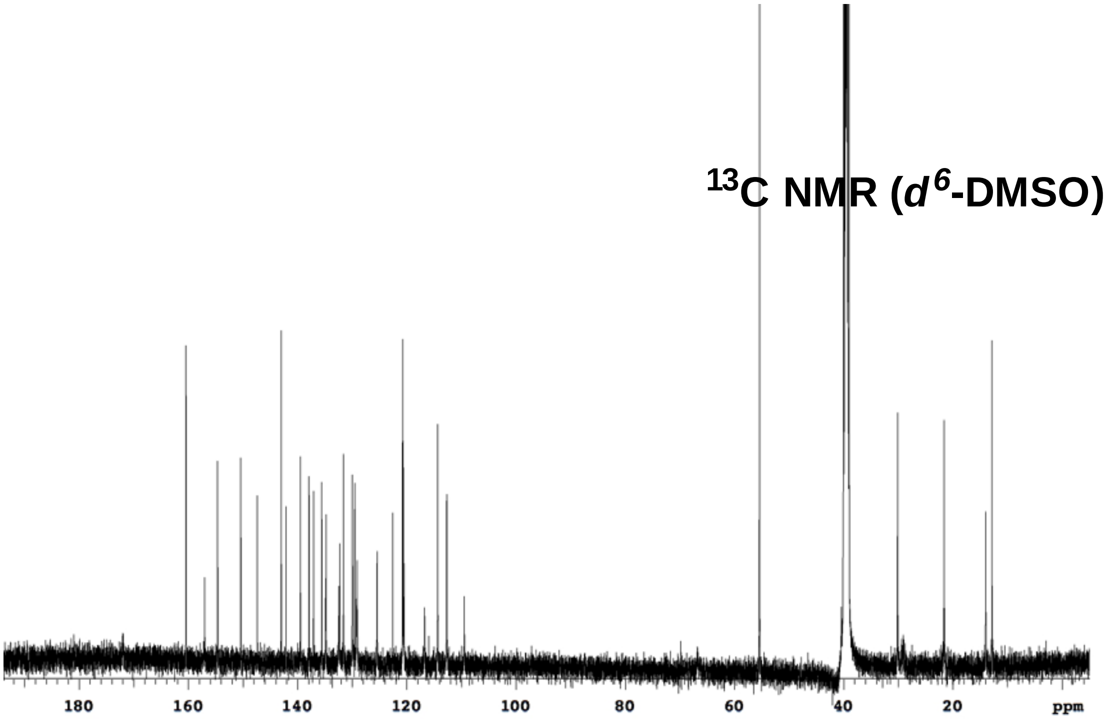

ʾF MS: 0.217 to 0.350 min from 02060810.wiff
.55375342927144690e-004, t0=6.50731022602521990e+001
Max. 55.9 counts
(M-2Na) 2-
546.6054
547.1040
0
10
20
30
40
50
56
540.0
541.0
542.0
543.0
544.0
545.0
546.0
547.0
548.0
549.0
550.0
551.0
552.0
553.0
554.0
555.0
556.0
m/z, amu
MS

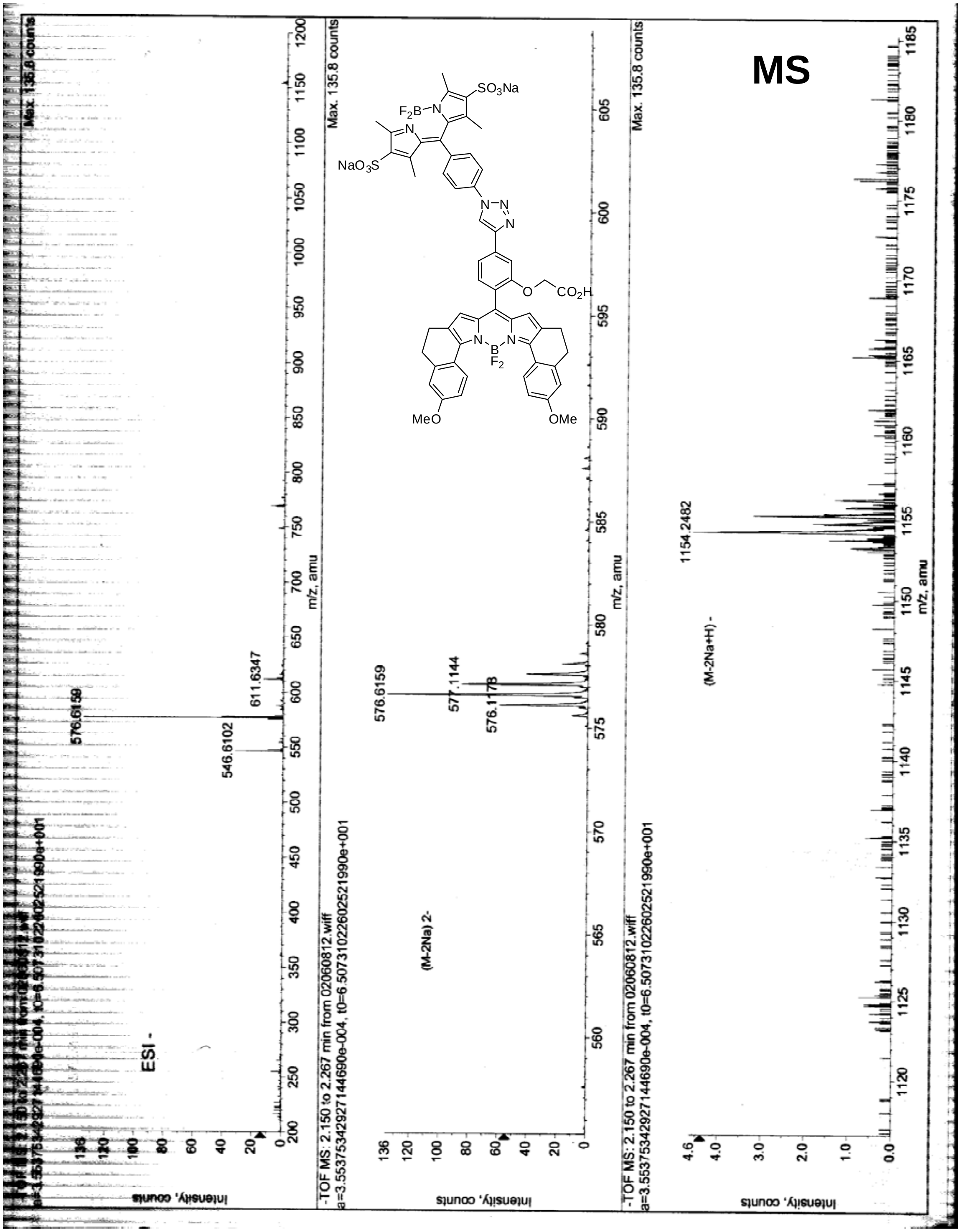


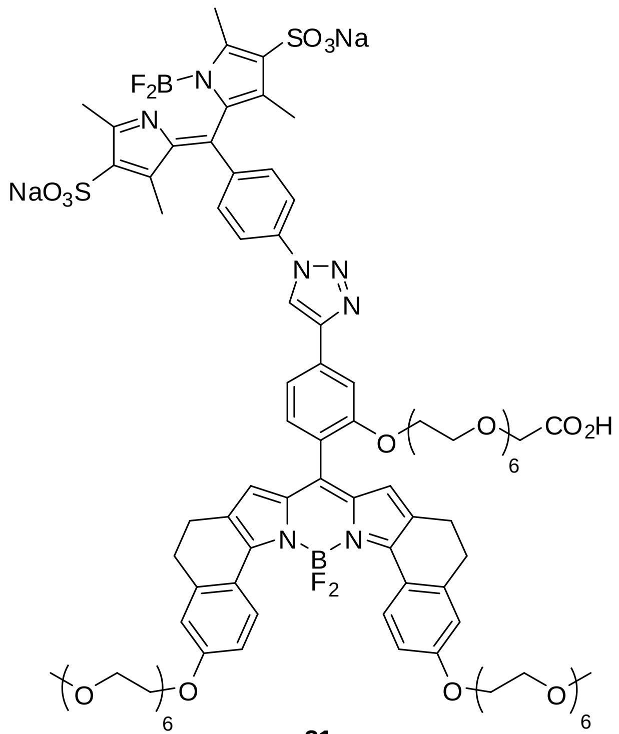
31f

Green solid (14 mg, 35 %), prepared from donor **34** (19 mg, 0.033 mmol) and BODIPY acceptor **37f** (32 mg, 0.051 mg). ^{1}H NMR (500 MHz, *d*^{6}-DMSO) δ 9.64 (s, 1H), 8.54 (d, 2H, *J* = 9.0 Hz), 8.23 (d, 2H, *J* = 8.4 Hz), 7.68 (d, 2H, *J* = 8.4 Hz), 7.65 (s, 1H), 7.60 (d, 1H, *J* = 7.7 Hz), 7.46 (d, 1H, *J* = 7.7 Hz), 7.02 (dd, 2H, *J* = 9.0, 2.0 Hz), 6.97 (d, 2H, *J* = 2.0 Hz), 6.78 (s, 2H), 4.51 (s, 2H), 3.84 (s, 6H), 2.86-2.83 (m, 4H), 2.66 (s, 6H), 2.64-2.60 (m, 4H), 1.62 (s, 6H); ^{13}C NMR (125 MHz, *d*^{6}-DMSO) δ 172.0, 160.4, 157.0, 154.6, 150.4, 147.4, 143.0, 142.1, 139.5, 137.9, 137.1, 135.6, 134.9, 134.8, 132.4, 132.3, 131.6, 130.0, 129.5, 129.1, 125.4, 122.6, 120.8, 120.6, 120.5, 116.7, 114.3, 112.7, 109.5, 55.4, 30.1, 21.6, 14.0, 12.8, (one carbon is ambiguous, some peak might account for two carbons); MS (ESI) calcd for $(M\text{-}2Na)^{2-}$ $C_{56}H_{45}B_2F_4N_7O_{11}S_2$ m/z 576.64; found 576.62.


SO$_3$Na
F$_2$B
N
NaO$_3$S
N
N
N
O
CO$_2$H
N
B
F$_2$
N
MeO
OMe
^{1}H NMR (d^6-DMSO)
10
9
8
7
6
5
4
3
2
1
ppm

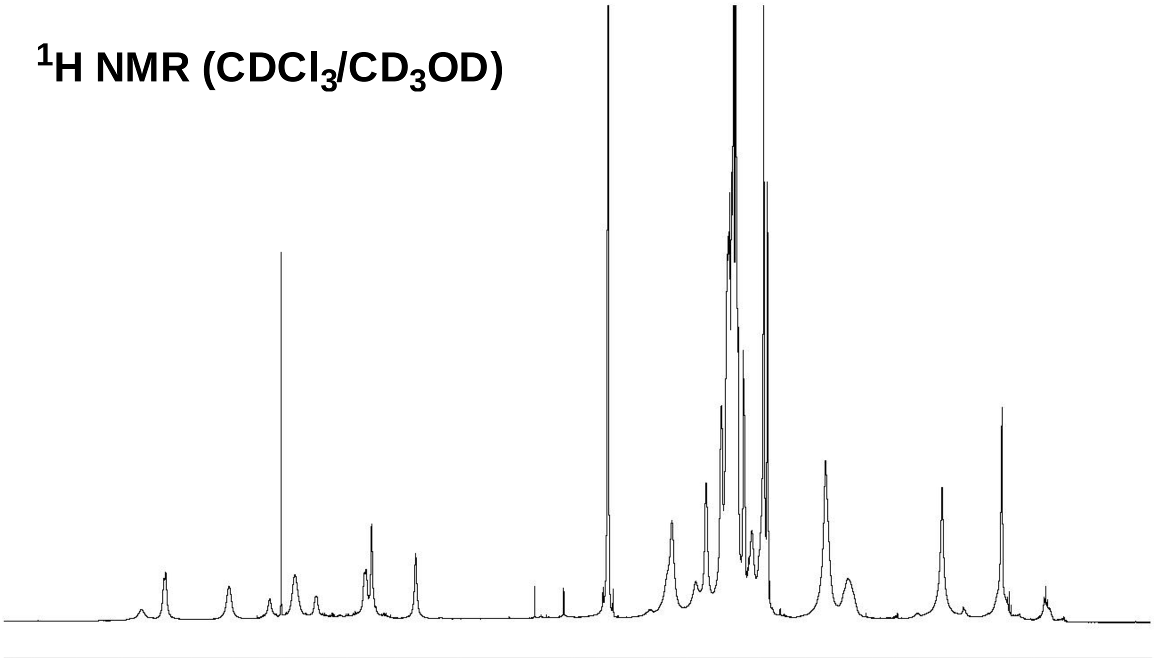

^{13}C NMR (d^6-DMSO)
180
160
140
120
100
80
60
40
20
ppm


MS
ESI -
-TOF MS: 2.150 to 2.267 min from 02060812.wiff
a=3.55375342927144690e-004, t0=6.50731022602521990e+001
Max. 135.8 counts
Intensity, counts
m/z, amu
576.6159
546.6102
611.6347
577.1144
576.1178
(M-2Na) 2-
1154.2482
(M-2Na+H) -
SO$_3$Na
F$_2$B
NaO$_3$S
CO$_2$H
MeO
OMe
F$_2$

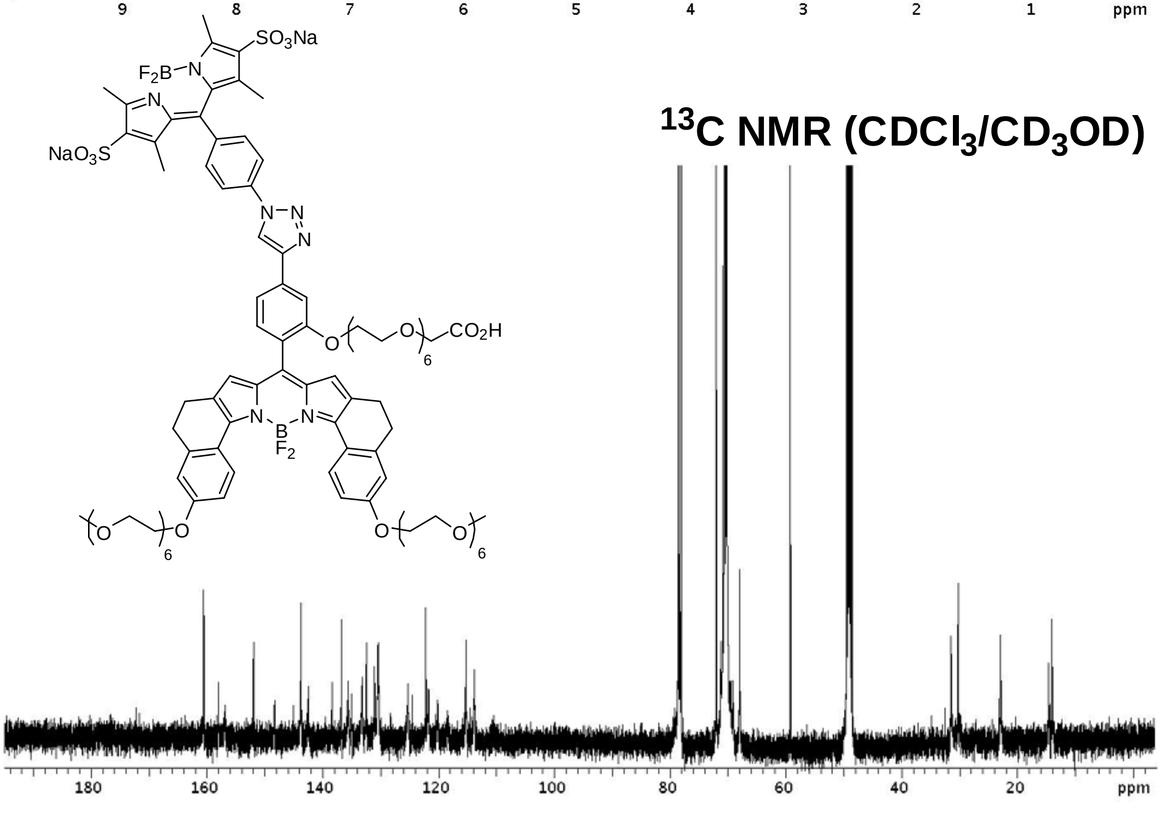


31g

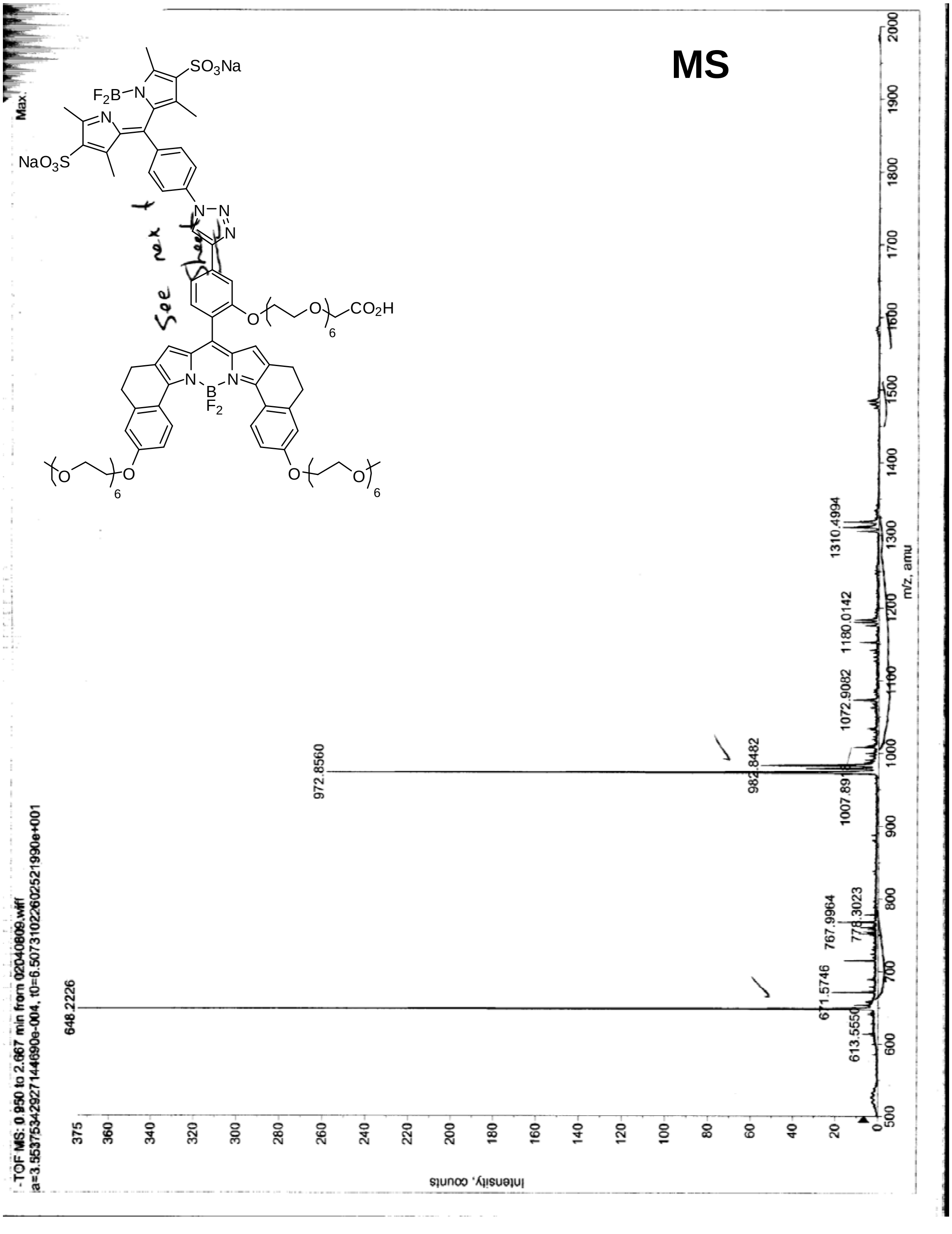
Green solid (30 mg, 25 %), prepared from donor **34** (34 mg, 0.06 mmol) and BODIPY acceptor **37g** (111 mg, 0.078 mg). ^{1}H NMR (500 MHz, 1:1 $CDCl_3/CD_3OD$) δ 8.83 (br, 1H), 8.63-8.62 (m, 2H), 8.06 (br, 2H), 7.70 (br, 1H), 7.48 (br, 3H), 7.29 (br, 1H), 6.86-6.80 (m, 4H), 6.41 (s, 2H), 4.15 (br, 6H), 3.94-3.85 (m, 8H), 3.72-3.34 (m, 66H), 2.80 (br, 10H), 2.60 (br, 4H), 1.77 (s, 6H); ^{13}C NMR (125 MHz, 1:1 $CDCl_3/CD_3OD$) δ 160.5, 157.9, 156.8, 151.9, 148.3, 143.7, 142.4, 138.4, 136.8, 135.6, 135.0, 133.2, 132.4, 131.0, 130.4, 125.2, 124.5, 122.1, 121.7, 120.1, 118.3, 115.2, 113.8, 72.0, 71.2, 70.9, 70.6, 70.5, 70.3, 70.2, 70.1, 67.9, 59.2, 31.4, 22.9, 14.6, 14.0, (due to poor solubility, 22C are ambiguous, several peaks might account for more than one carbon); MS (ESI) m/z calcd for $(M\text{-}2Na)^{2-}$ $C_{92}H_{117}B_2F_4N_7O_{29}S_2$ M/Z: 972.87, found 972.86; calcd for $(M\text{-}2Na\text{-}H)^{3-}$ $C_{92}H_{116}B_2F_4N_7O_{29}S_2$ M/Z: 648.26, found 648.22.



^{1}H NMR ($CDCl_3/CD_3OD$)

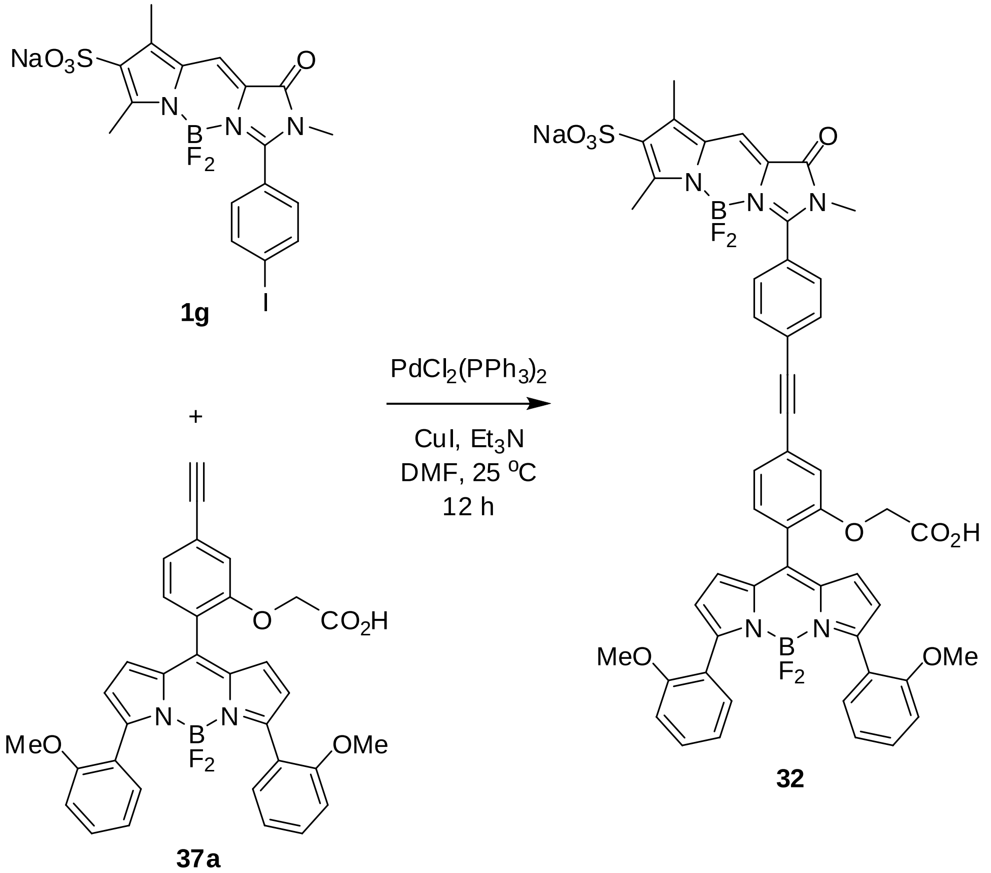


^{13}C NMR ($CDCl_3/CD_3OD$)


-TOF MS: 0.950 to 2.667 min from 02040809.wiff
a=3.55375342927144690e-004, t0=6.50731022602521990e+001
Max.
MS
See next sheet
F2B
N
SO3Na
NaO3S
N–N
N
O
CO2H
6
B
F2
6
6
Intensity, counts
m/z, amu
648.2226
613.5550
671.5746
767.9964
778.3023
972.8560
982.8482
1007.89
1072.9082
1180.0142
1310.4994
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
375
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

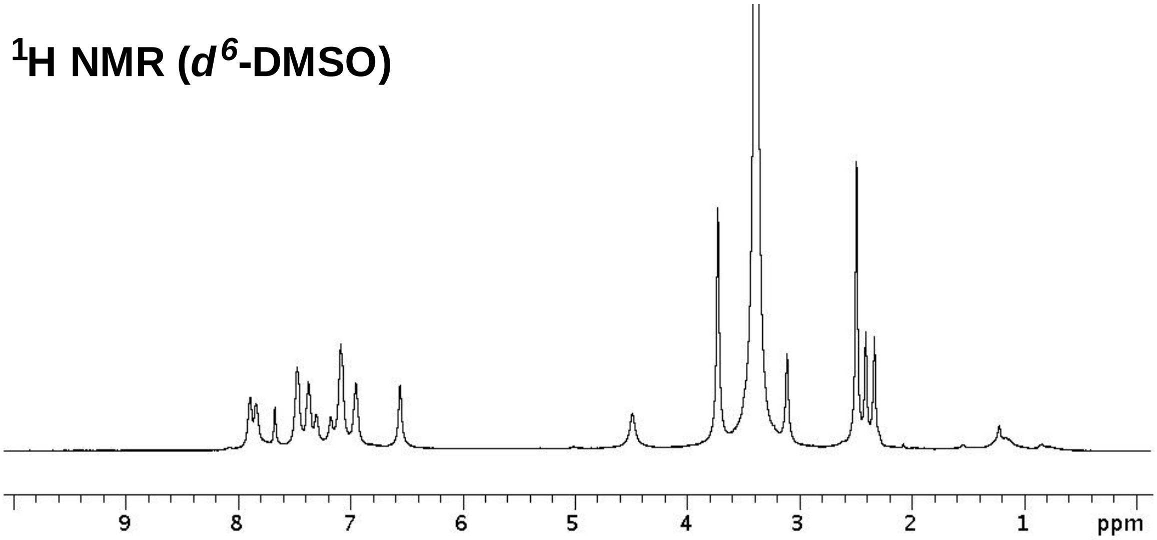
(2) Syntheses of cassettes 32-33

Typical procedures:

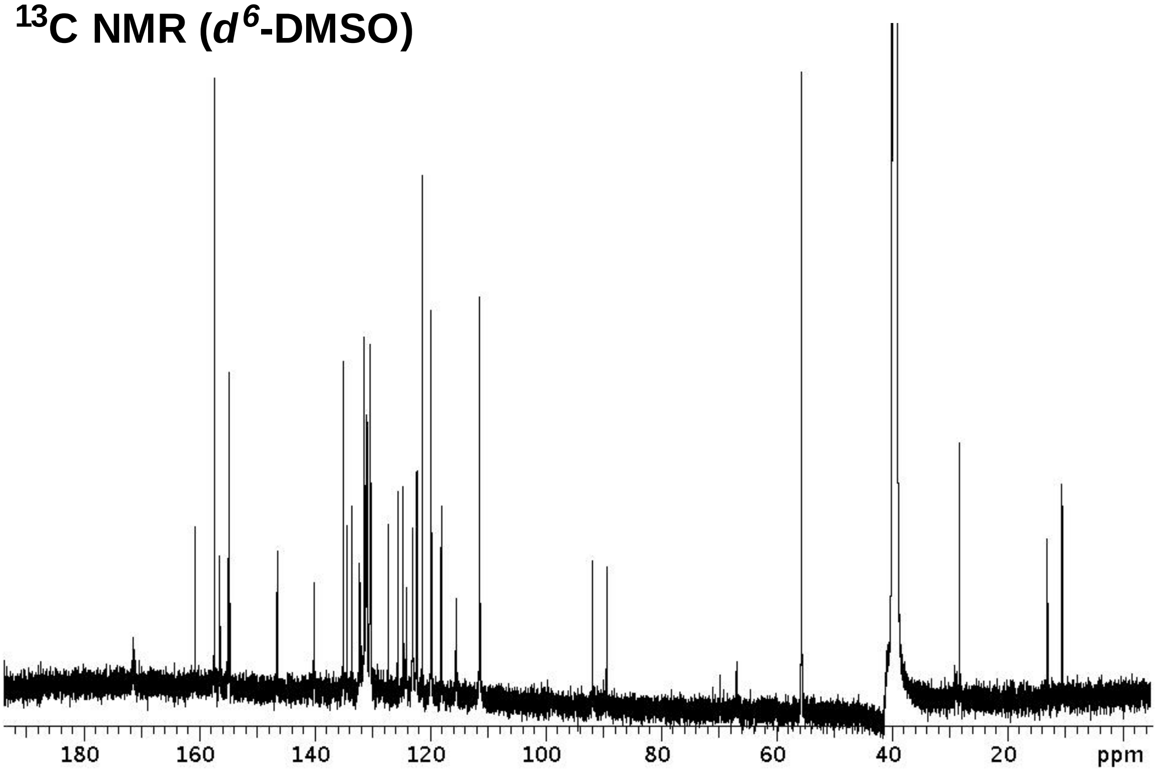


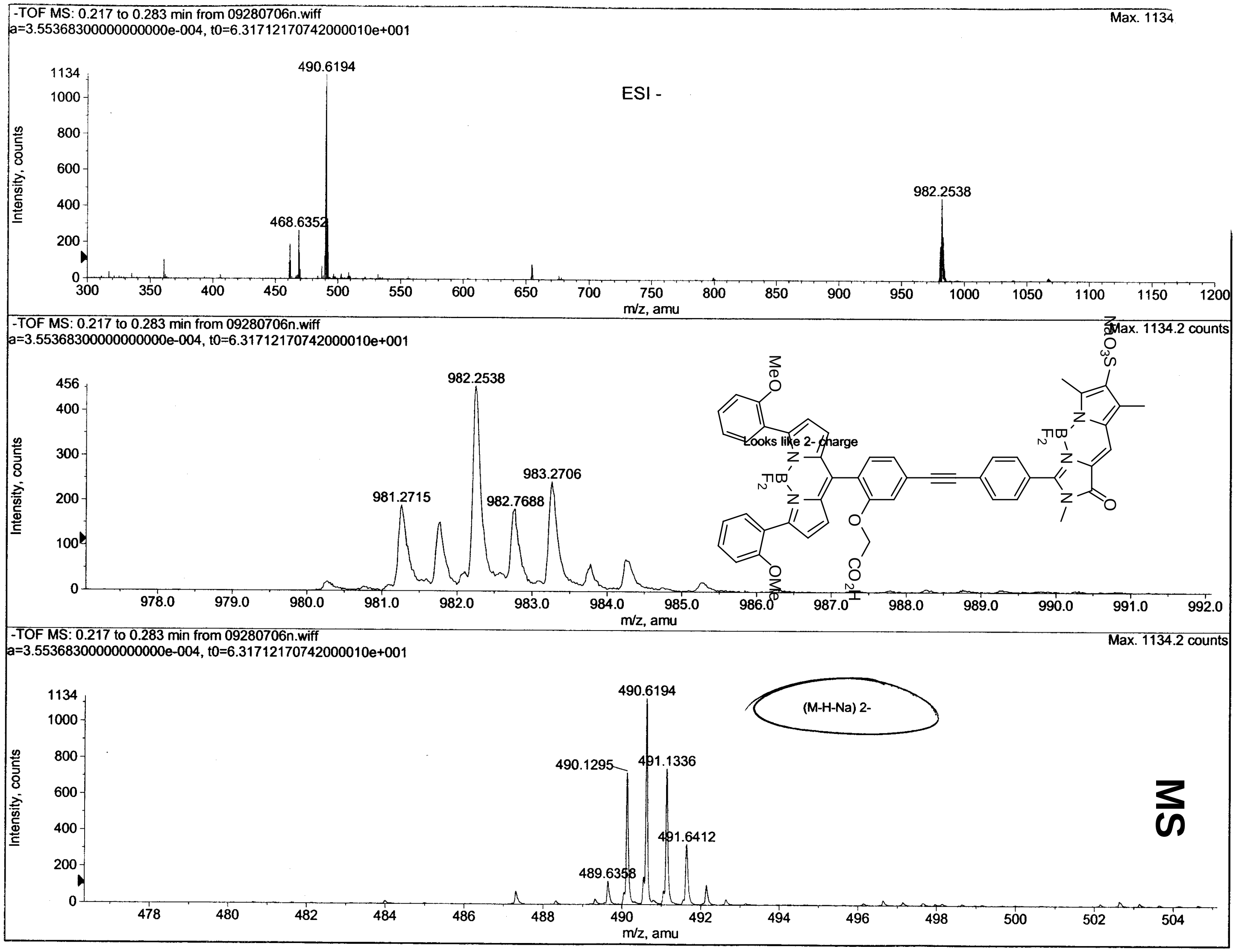
A solution of donor **1g** (14 mg, 0.025 mmol), BODIPY acceptor **37a** (20 mg, 0.035 mmol), $PdCl_2(PPh_3)_2$ (2 mg, 0.003 mmol) and CuI (0.6 mg, 0.003 mmol) in 3 mL DMF was degassed and filled with N_2 then Et_3N (35 μL, 0.25 mmol) was added. The reaction mixture was stirred at 25 ^{o}C for 24 h then concentrated under reduced pressure. The residue was purified by flash chromatography (20 % $MeOH/CH_2Cl_2$) to afford the product **32** (20 mg, 79 %) as a red solid. ^{1}H NMR (500 MHz, d^6-DMSO) δ 7.89-7.84 (m, 3H), 7.67 (s, 1H), 7.47 (br, 3H), 7.37-7.30 (m, 4H), 7.17 (s, 1H), 7.08-7.07 (m, 4H), 6.95 (br, 2H), 6.56 (br, 2H), 4.48 (s, 2H), 3.73 (s, 6H), 3.11 (s, 3H), 2.41 (s, 3H), 2.33 (s, 3H); ^{13}C NMR (500 MHz, d^6-DMSO) δ 171.3, 160.7, 157.3, 156.4, 155.0, 154.7, 146.4, 140.1, 140.0, 135.0, 134.4, 133.5, 132.2, 131.4, 131.1, 130.9, 130.4, 130.2, 127.2, 125.6, 124.6, 124.1, 123.0, 122.3, 121.3, 119.8, 118.1, 117.9, 115.5, 111.4, 91.8, 89.4, 66.9, 55.6, 28.2, 13.0, 10.5; MS (ESI) m/z calcd for $(M\text{-}Na)^{-}$ $C_{50}H_{38}B_2F_4N_5O_9S$ 982.25; found 982.25. m/z calcd for $(M\text{-}Na\text{-}H)^{2-}$ $C_{50}H_{37}B_2F_4N_5O_9S$ 490.62; found 490.62.

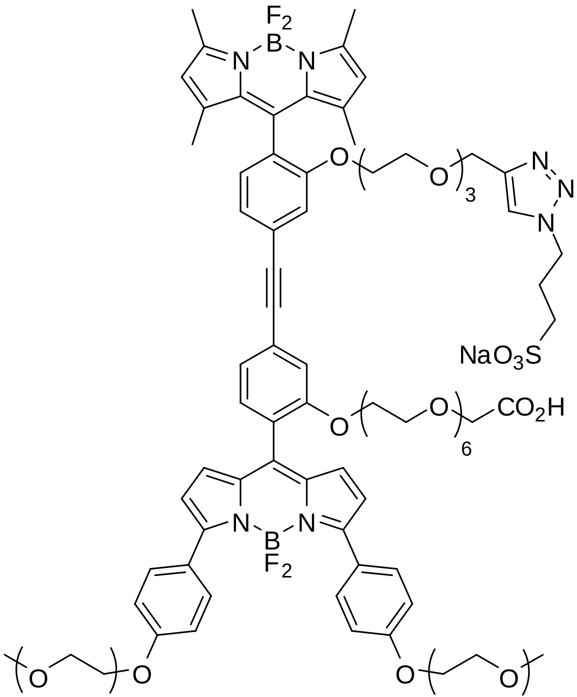
^{1}H NMR (d^{6}-DMSO)



^{13}C NMR (d^{6}-DMSO)

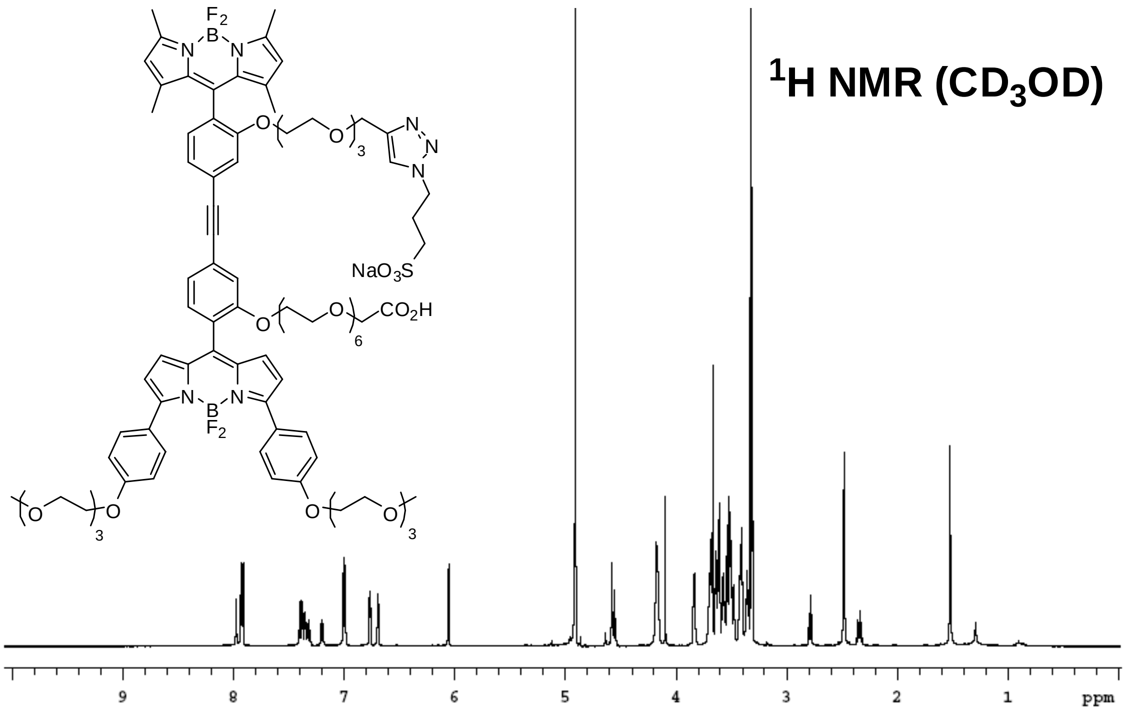



-TOF MS: 0.217 to 0.283 min from 09280706n.wiff
a=3.55368300000000000e-004, t0=6.31712170742000010e+001
Max. 1134
ESI -
Intensity, counts
490.6194
468.6352
982.2538
m/z, amu
-TOF MS: 0.217 to 0.283 min from 09280706n.wiff
a=3.55368300000000000e-004, t0=6.31712170742000010e+001
Max. 1134.2 counts
982.2538
981.2715
982.7688
983.2706
Looks like 2- charge
NaO3S
MeO
OMe
CO2H
F2
-TOF MS: 0.217 to 0.283 min from 09280706n.wiff
a=3.55368300000000000e-004, t0=6.31712170742000010e+001
Max. 1134.2 counts
490.6194
490.1295
491.1336
491.6412
489.6358
(M-H-Na) 2-
MS

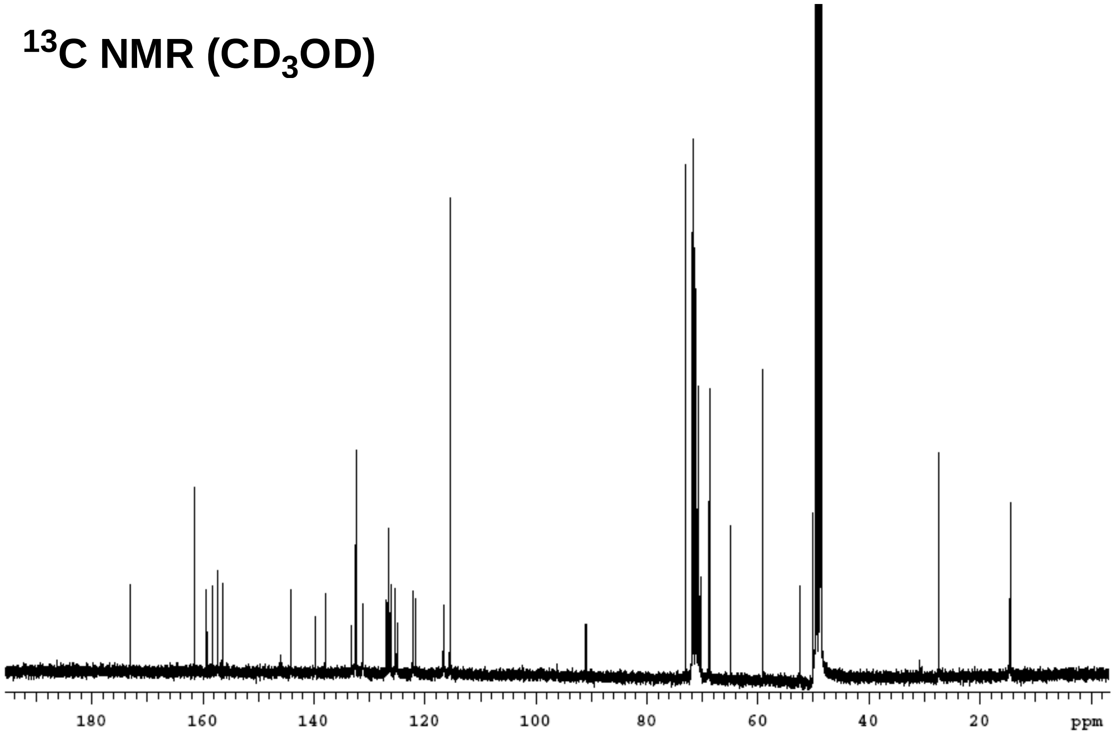


33

Blue sticky solid (45 mg, 45 %), prepared from BODIPY donor **35** (46 mg, 0.056 mmol) and BODIPY acceptor **37c** (75 mg, 0.068 mmol). ^{1}H NMR (500 MHz, CD_3OD) δ 7.97 (s, 1H), 7.92 (d, 4H, *J* = 8.8 Hz), 7.40-7.30 (m, 5H), 7.20 (d, 1H, *J* = 7.5 Hz), 7.00 (d, 4H, *J* = 8.8 Hz), 6.77 (d, 2H, *J* = 4.3 Hz), 6.69 (d, 2H, *J* = 4.3 Hz), 6.05 (s, 2H), 4.58-4.54 (m, 4H), 4.18-4.16 (m, 8H), 4.09 (s, 2H), 3.84-3.82 (m, 4H), 3.70-3.30 (m, 54H), 2.80-2.77 (m, 2H), 2.48 (s, 6H), 2.37-2.31 (m, 2H), 1.52 (s, 6H); ^{13}C NMR (125 MHz, CD_3OD) δ 173.0, 161.6, 159.3, 158.2, 157.3, 156.5, 144.1, 139.8, 139.7, 137.9, 133.2, 132.6, 132.4, 131.2, 131.1, 127.0, 126.7, 126.5, 126.3, 126.0, 125.3, 125.2, 125.0, 122.1, 121.6, 116.6 (2 peaks: 116.62, 116.58), 115.5, 91.2, 90.9, 73.0, 72.0, 71.8 (2 peaks: 71.83, 71.76), 71.6 (2 peaks: 71.64, 71.57), 71.4, 71.3 (2 peaks: 71.30, 71.26), 71.1 (3 peaks: 71.13, 71.09, 71.08), 71.0, 70.9, 70.8, 70.7, 70.5, 70.3, 70.2, 68.9, 68.7, 64.9, 59.1, 52.4, 50.0, 49.1, 27.3, 14.7, 14.4, (4C are ambiguous, several peaks might account for more than one carbon).



^{1}H NMR (CD$_3$OD)



^{13}C NMR (CD$_3$OD)

Spectroscopic properties of through-bond energy transfer cassettes 31-33

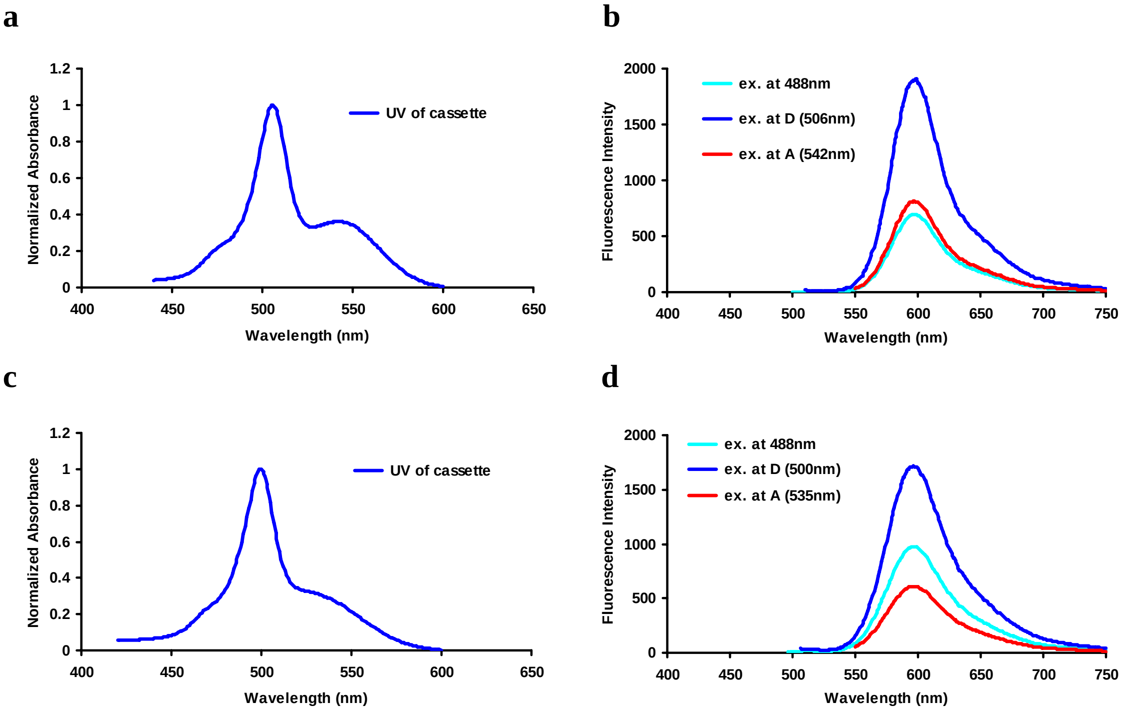


Figure F1. Spectra for cassette **31a**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

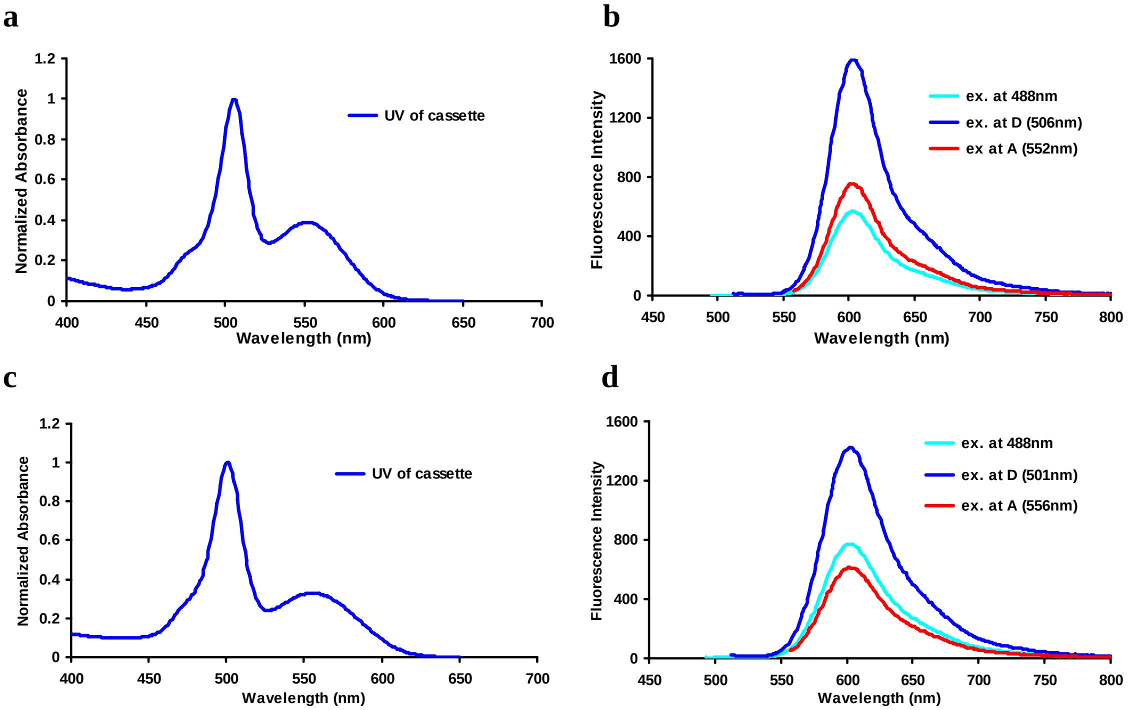


Figure F2. Spectra for cassette **31b**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

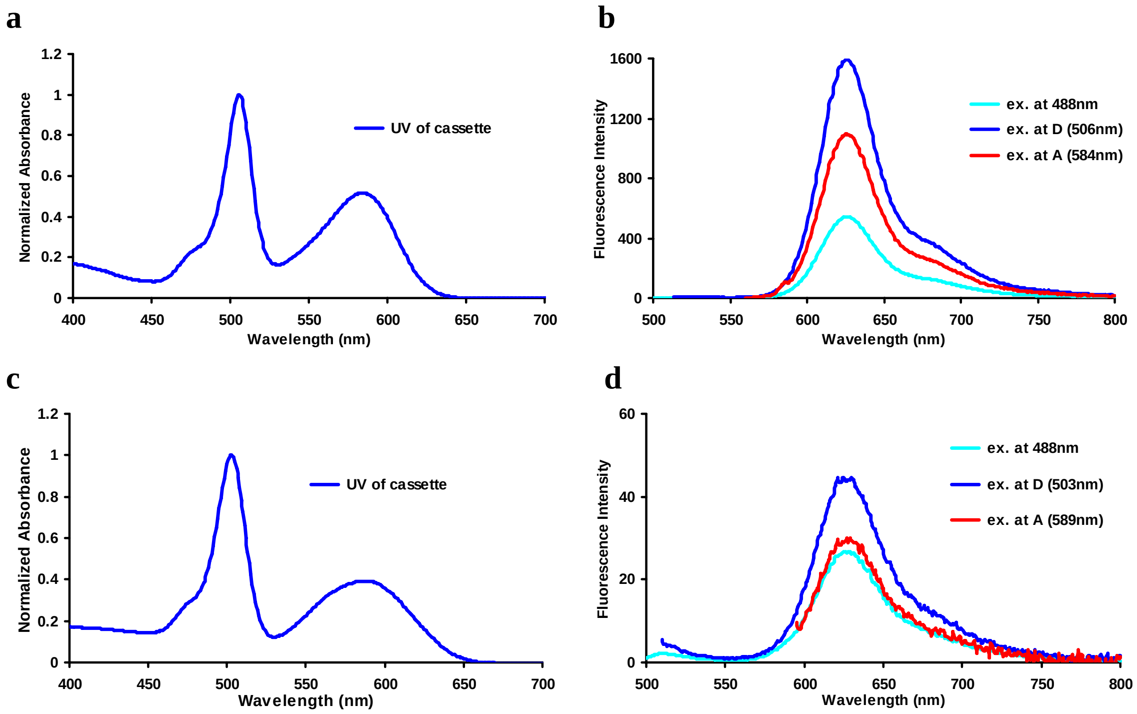


Figure F3. Spectra for cassette **31c**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

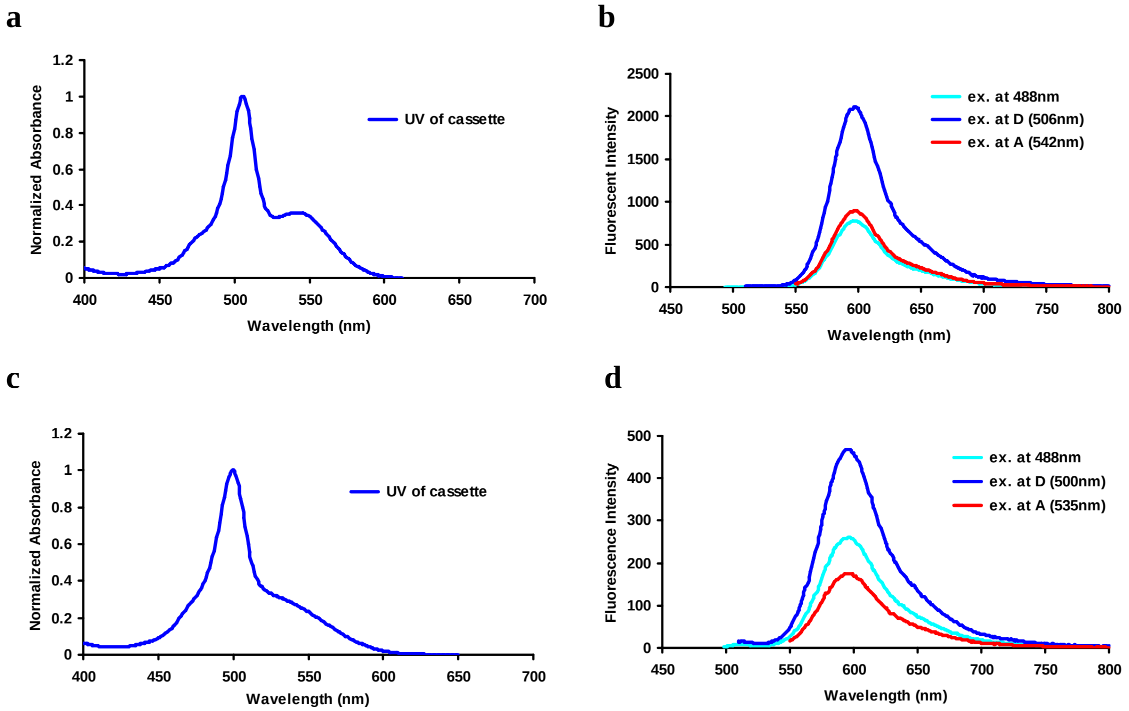


Figure F4. Spectra for cassette **31d**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

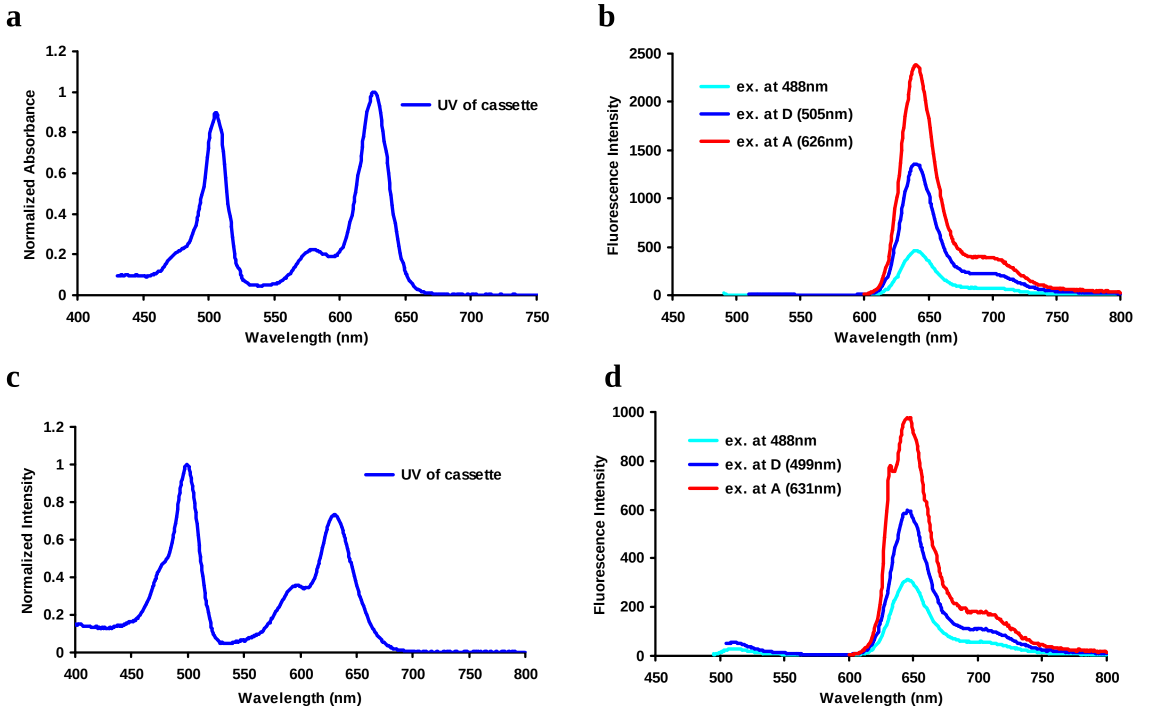


Figure F5. Spectra for cassette **31e**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

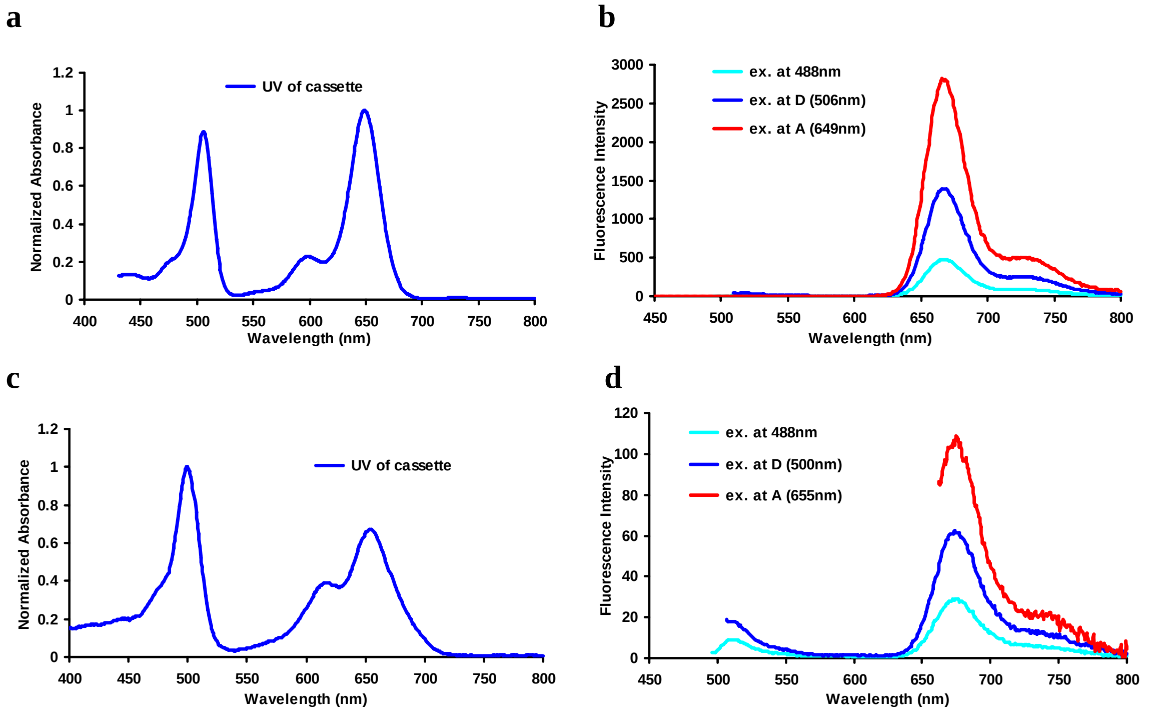


Figure F6. Spectra for cassette **31f**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

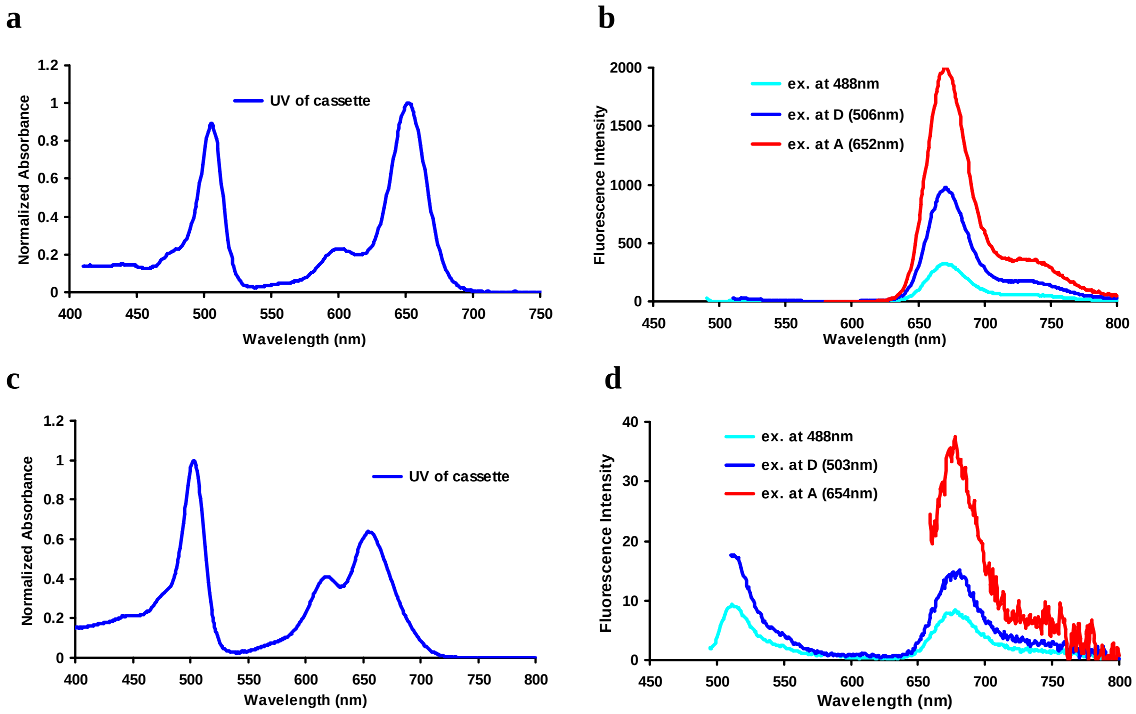


Figure F7. Spectra for cassette **31g**, in MeOH: **a** absorbance and, **b** fluorescence; in 0.1 M phosphate buffer (pH 7.4): **c** absorbance and, **d** fluorescence. (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

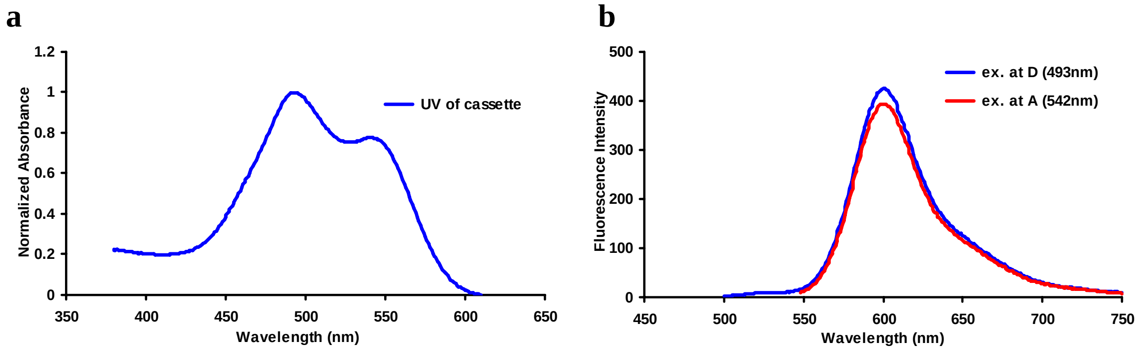


Figure F8. **a** Normalized absorbance and, **b** fluorescence of cassette **32** in MeOH (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

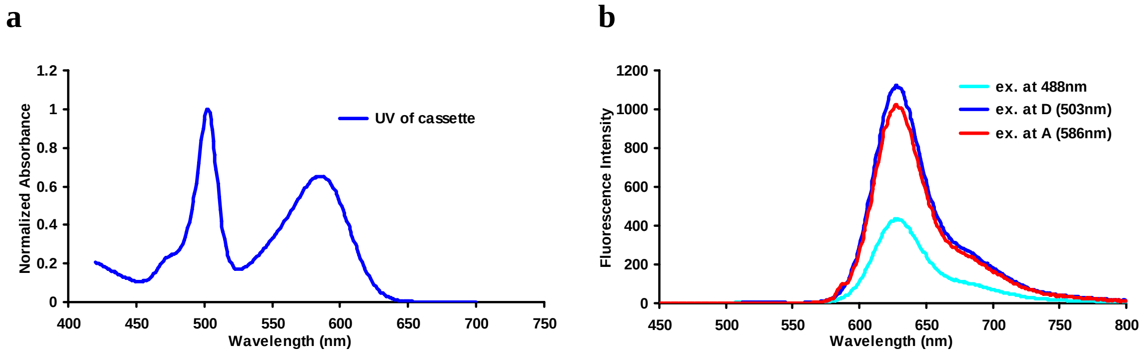


Figure F9. **a** Normalized absorbance and, **b** fluorescence of cassette **33** in MeOH (10^{-6} M for absorbance and 10^{-7} M for fluorescence).

Table F1. Photophysical properties of cassette **33**

solvent	$\lambda_{\text{abs D}}$ (nm)	$\lambda_{\text{abs A}}$ (nm)	$\lambda_{\text{emiss A}}$ (nm)	log($\varepsilon_{\max}$)		Φ of acceptor		**E.T.E** (%)[c]
				D	A	Ex. D[a]	Ex. A[b]	
MeOH	503	586	625	4.97	4.78	0.63	0.67	94

[a] Rhodamine 6G (Φ = 0.94 in EtOH); and [b] Cresyl Violet (Φ = 0.54 in MeOH) were used as standards; [c] energy transfer efficiency (E.T.E) was calculated by dividing the quantum yield of acceptor when excited at donor by the quantum yield of acceptor when excited at acceptor.

Protein-protein interaction studies:

(1) Syntheses of fluorescently labeled proteins:

*Cassette **1** refers to BODIPY cassette **31b**; Cassette **2** refers to BODIPY cassette **31e***

Synthesis of 1-Streptavidin

Cassette **1** (1.0 mg, 0.6 μmol) was dissolved in 100 μL dry DMF and NHS (0.4 mg, 3.5 μmol), DIC (0.5 μL, 3.2 μmol) were added. The reaction mixture was stirred at room temperature for 24 h.

The activated cassette **1** solution (35 μL, 5 eq.) was added to the solution of streptavidin (2.2 mg, 1 eq.) in 0.5 mL sodium bicarbonate buffer (0.1 M, pH 8.3). The solution was stirred at room temperature in the dark for 30 min. The unreacted cassette **1** was removed by Sephadex^{TM} G-25 (PD-10) desalting column.

Synthesis of 1-Bovine Serum Albumin (BSA)

The activated cassette **1** solution (63 μL, 5 eq.) was added to the solution of BSA (5 mg, 1 eq.) in 1 mL sodium bicarbonate buffer (0.1 M, pH 8.3). The solution was stirred at room temperature in the dark for 30 min. The unreacted cassette **1** was removed by Sephadex^{TM} G-25 (PD-10) desalting column.

Synthesis of 2-Avidin

Cassette **2** (1.0 mg, 0.9 μmol) was dissolved in 100 μL dry DMF and NHS (0.5 mg, 4.3 μmol), DIC (0.7 μL, 4.5 μmol) were added. The reaction mixture was stirred at room temperature for 24 h.

The activated cassette **2** solution (100 μL, 5 eq.) was added to the solution of avidin (12 mg, 1 eq.) in 1.0 mL sodium bicarbonate buffer (0.1 M, pH 8.3). The solution was stirred at room temperature in the dark for 30 min. The unreacted cassette **2** was removed by Sephadex^{TM} G-25 (PD-10) desalting column.

Synthesis of Atto-425-BSA

A solution of Atto 425-NHS ester (1.0 mg, 6 eq.) in 100 μL DMF was added to the solution of BSA (22 mg, 1 eq.) in 2.0 mL sodium bicarbonate buffer (0.1 M, pH 8.3).

The solution was stirred at room temperature in the dark for 30 min. The unreacted Atto 425- NHS was removed by Sephadex^{TM} G-25 (PD-10) desalting column.

Synthesis of Atto-425-BSA-Biotin

A solution of biotin NHS ester (0.7 mg, 10 eq.) in 100 μL DMF was added to the solution of Atto-425-BSA (13 mg, 1 eq.) in 1.0 mL sodium bicarbonate buffer (0.1 M, pH 8.3). The solution was stirred at room temperature in the dark for 30 min. The unreacted biotin-NHS was removed by Sephadex^{TM} G-25 (PD-10) desalting column.

Determination of Dye/Protein Ratios

(i) measure the absorbance of the protein-dye conjugate at 280 nm (A_{280}) and at the λ_{max} for the dye (A_{max})

(ii) correct for the contribution of the dye to the absorbance at A_{280}:

$$A_{protein} = A_{280} - A_{max} * (A_{280\ free\ dye}/A_{max\ free\ dye})$$

(iii) Dye/protein ration = $(A_{max}/\varepsilon_{dye}) / (A_{protein}/\varepsilon_{protein})$

Determination of Biotin/Protein Ratios

(i) HABA/avidin reagent (Sigma, H2153) was reconstituted with 10 mL deionized water. In a 1 mL cuvette, pipette 0.9 mL HABA/avidin reagent solution, read $A_{500\ HABA/avidin}$. Add 0.1 mL sample, mix by inverting the cuvette, wait for 5 min then read $A_{500\ HABA/avidin+samples}$.

(ii) sample blank: dilute 0.1 mL sample with 0.9 mL water and read $A_{500\ sample\ blank}$.

(iii) $\Delta A_{500} = 0.9A_{500\ HABA/avidin} + A_{500\ sample\ blank} - A_{500\ HABA/avidin+samples}$

(ΔA_{500} should be in the range of 0.1-0.4. If a change greater than this is observed, the sample should be diluted and measure again, as this indicates there is more biotin present than avidin)

(iv) μmol biotin/ml sample = $10 * (\Delta A_{500}/34)$

(v) biotin/protein ratio = (μmol biotin/ml sample) / (μmol protein/ml sample)

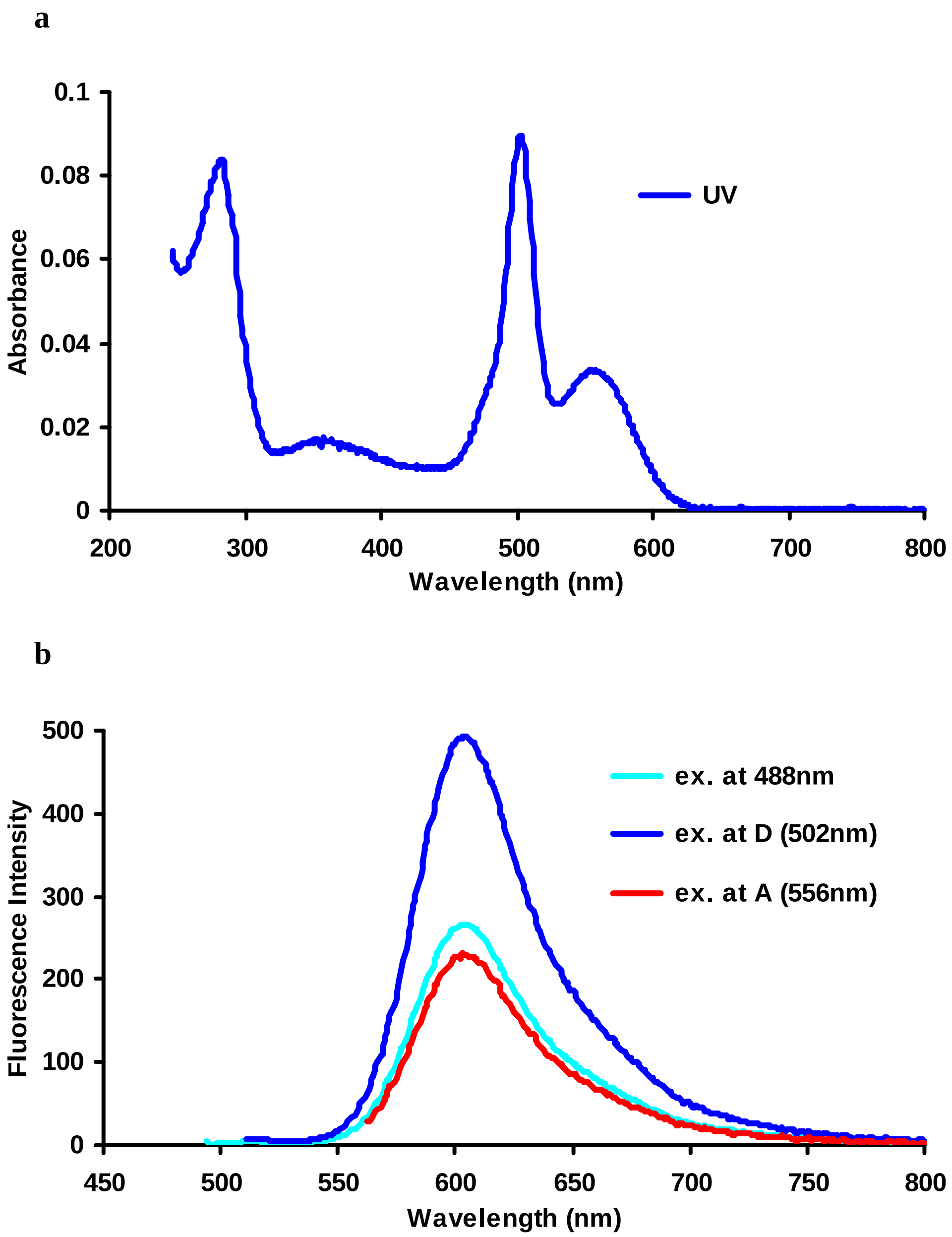


Figure F10. Absorbance and fluorescence spectra for **1**-streptavidin in 0.1 M phosphate buffer (pH 7.4). $\lambda_{\text{max abs}}$(D): 502 nm; $\lambda_{\text{max abs}}$(A): 556 nm; λ_{emiss}(A): 604 nm;

dye/protein ratio = 2.5 : 1.0

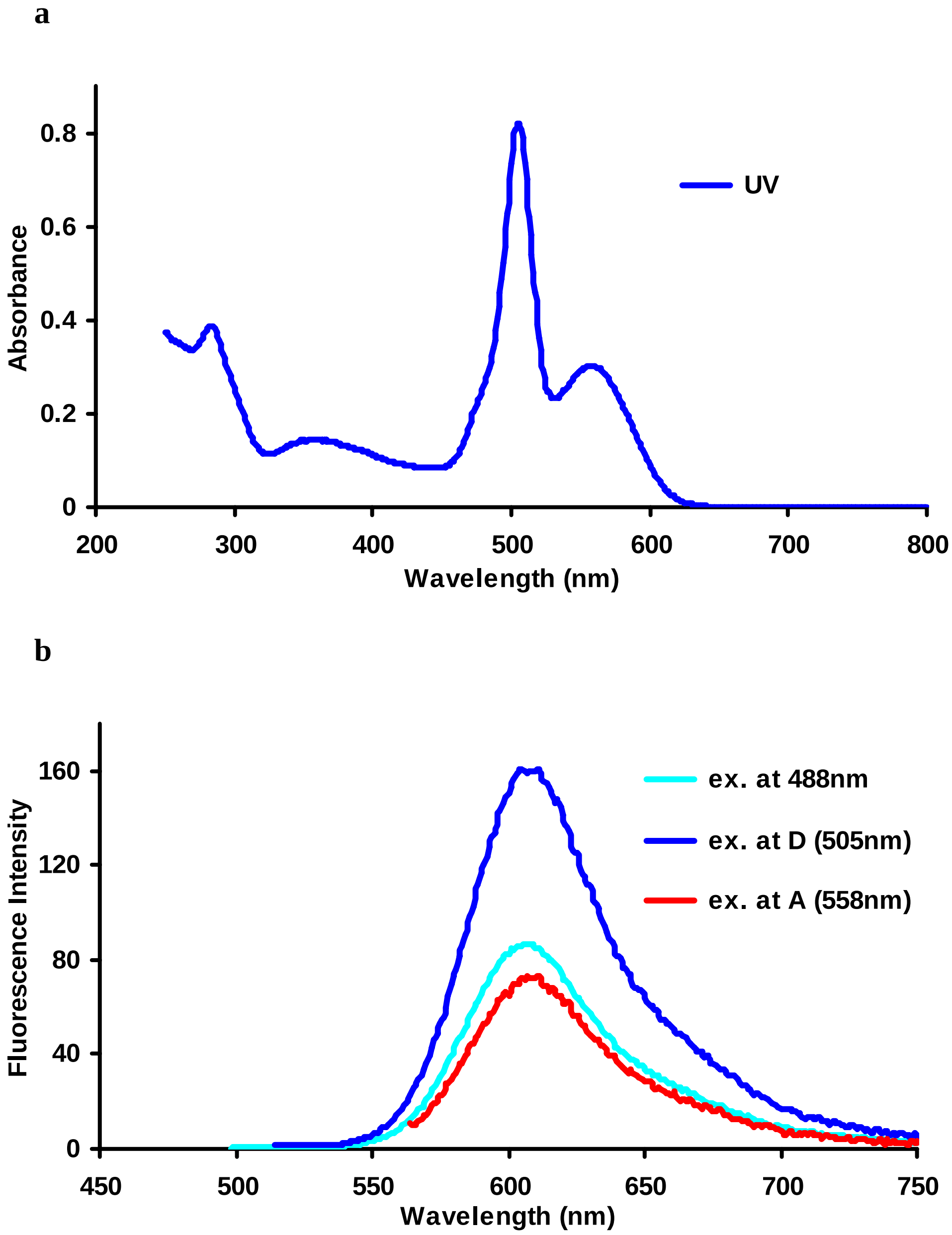


Figure F11. Absorbance and fluorescence spectra for **1**-BSA in 0.1 M phosphate buffer (pH 7.4). $\lambda_{max\ abs}$(D): 505 nm; $\lambda_{max\ abs}$(A): 558 nm; λ_{emiss}(A): 604 nm;

dye/protein ratio = 4.0 : 1.0

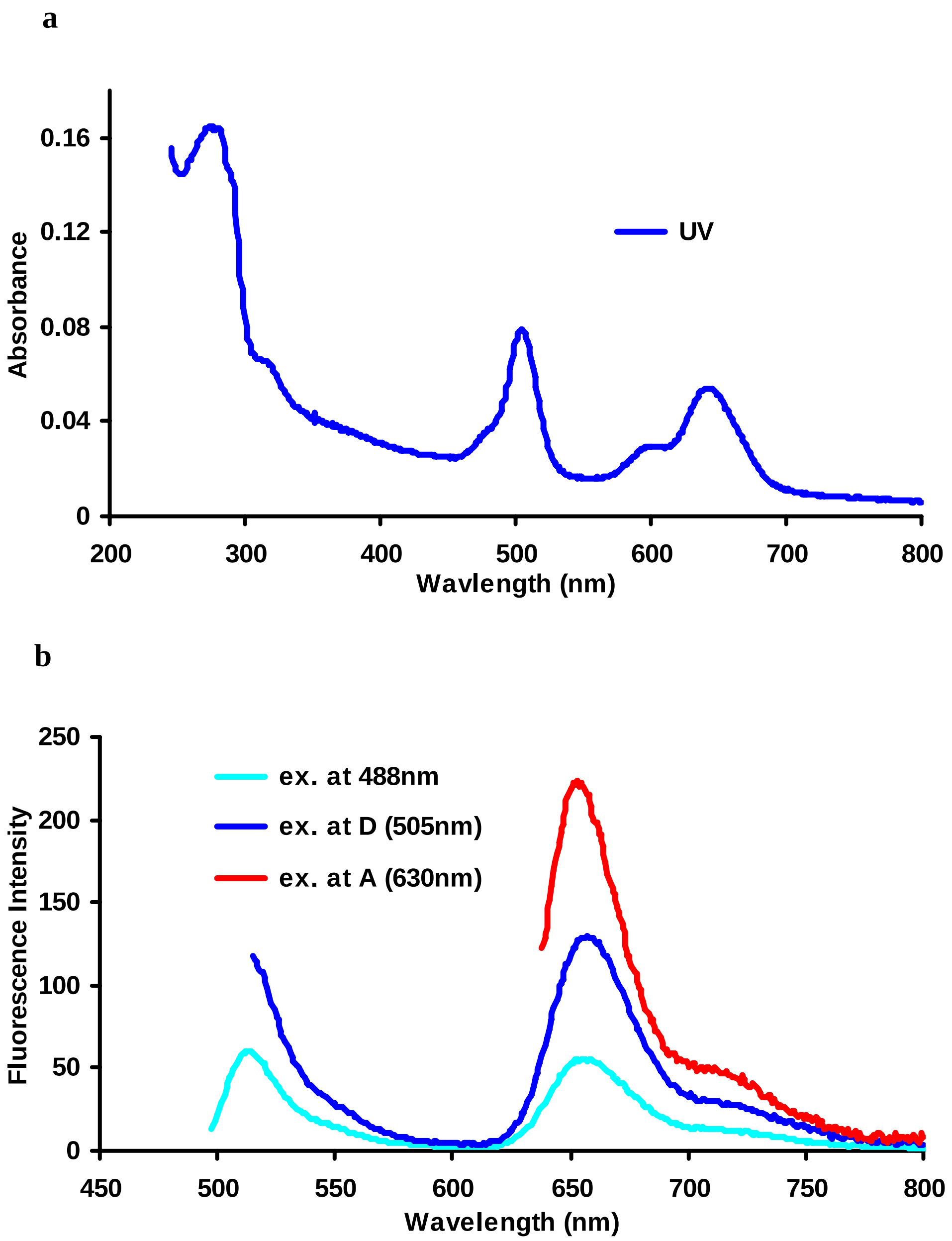


Figure F12. Absorbance and fluorescence spectra for **2**-avidin in 0.1 M phosphate buffer (pH 7.4). $\lambda_{\text{max abs}}$(D): 505 nm; $\lambda_{\text{max abs}}$(A): 642 nm; λ_{emiss}(A): 655 nm;

dye/protein ratio = 1.6 : 1.0

(Some energy leakage from the donor was observed when cassette **2** was attached to avidin.)

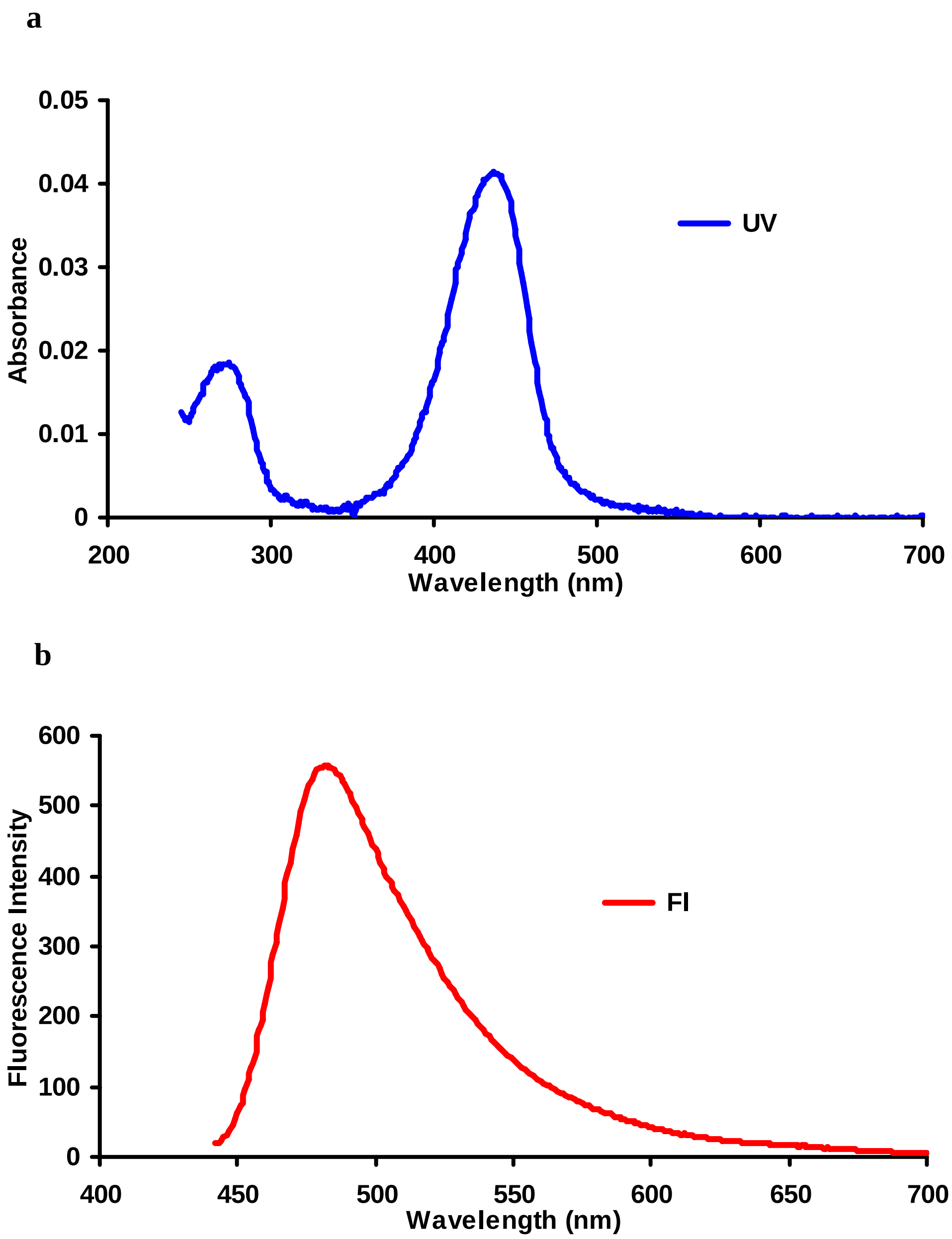


Figure S13. Absorbance and fluorescence spectra for Atto-425-BSA in 0.1 M phosphate buffer (pH 7.4). $\lambda_{\text{max abs}}$: 437 nm; λ_{emiss}(A): 482 nm;

dye/protein ratio = 5.2 : 1.0

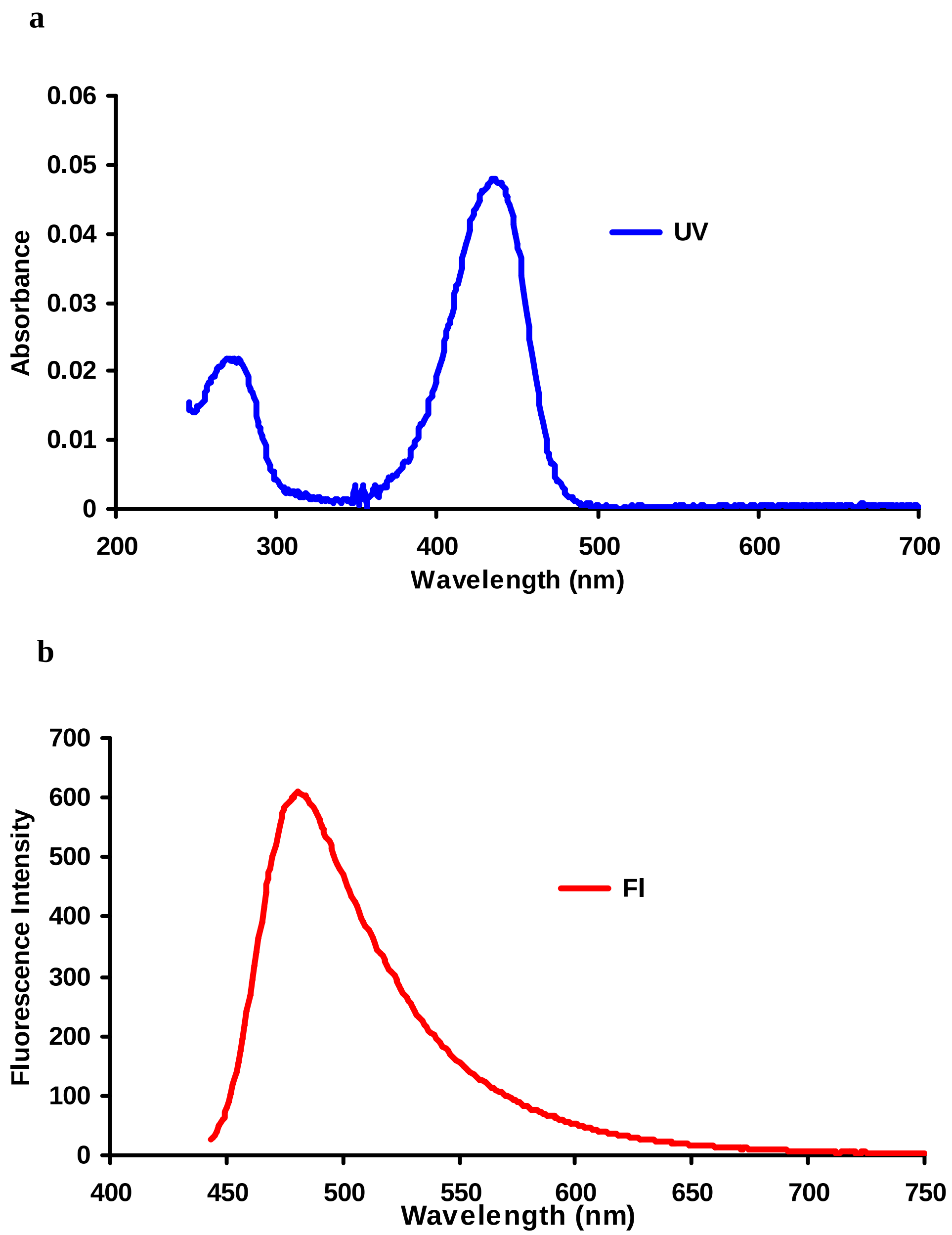


Figure F14. Absorbance and fluorescence spectra for Atto-425-BSA-biotin in 0.1 M phosphate buffer (pH 7.4). $\lambda_{\text{max abs}}$: 435 nm; λ_{emiss}(A): 481 nm;

biotin/protein/dye ratio = 6.6 : 1.0 : 5.2

(2) *In-vitro* three protein interactions

Experiment (ii) (the key experiment):

In a fluorescence cuvette, **1**-streptavidin (3.0 x 10^{-7} M) and **2**-avidin (9.0 x 10^{-7} M) were pretreated with BSA (1.8 x 10^{-5} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto-425-BSA-biotin (2.0 x 10^{-7} M) was added. These relative concentrations were used to give similar fluorescence intensities. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The solution was also excited at 488 nm to obtain the fluorescence after FRET ($F_{488\ after}$).

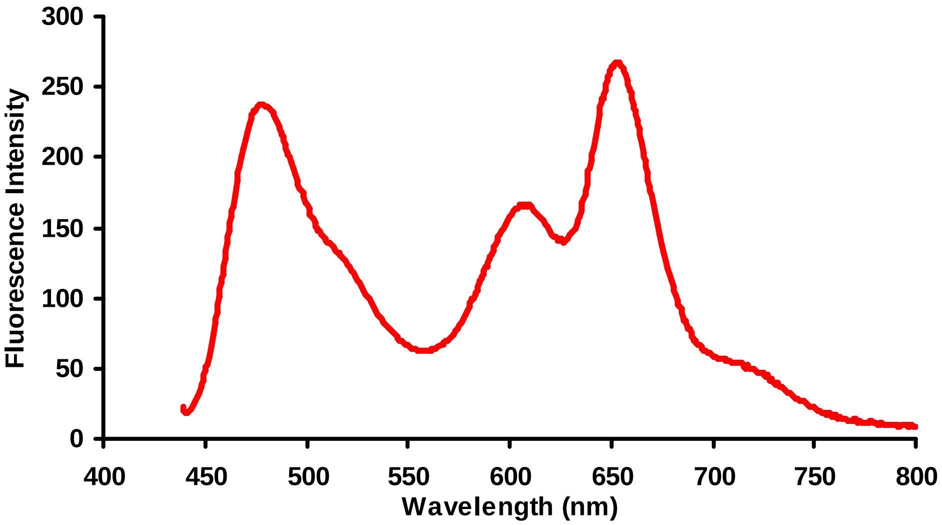


Figure F15. Fluorescence spectra for exp. (ii) after FRET (ex. 430 nm)

The following strategies were used to deduce the approximate fluorescence of the reaction mixture before FRET:

(i) Streptavidin (3.0 x 10^{-7} M), avidin (9.0 x 10^{-7} M) and BSA (1.8 x 10^{-5} M) were mixed in 0.1 M phosphate buffer (pH 7.4) for 20 min, then Atto-425-BSA-biotin (2.0 x 10^{-7} M) was added. The mixture was incubated for 10 min then excited at 430 nm to obtain the emission spectrum before FRET for Atto-425-BSA-biotin ($F_{baseline-1}$). This gives a baseline for one component (Atto-425-BSA-biotin).

(ii) Practically it is impossible to prepare BSA-biotin which has the same biotin/protein ratio and biotin distribution as the Atto-425-BSA-biotin used. However, it is necessary to correct the fluorescence of the cassettes involved in the protein-protein interaction for the presence of the BSA-biotin. Consequently, we made the assumption that the percentage correction for the presence of BSA-biotin would be almost the same when the sample was irradiated at 430 nm and at 488 nm. Thus, the following steps were used to correct the baseline for the **1**-streptavidin and **2**-avidin:

(a) **1**-streptavidin (3.0×10^{-7} M), **2**-avidin (9.0×10^{-7} M) were pretreated with BSA (1.8×10^{-5} M) in 0.1 M phosphate buffer (pH 7.4) for 20 min. Fluorescence of this solution was taken when excited at 430 nm (F_{430}) and 488 nm ($F_{488\ \mathrm{before}}$).

(b) Since Atto-425-BSA-biotin is not excited at 488 nm, the fluorescence changes observed between $F_{488\ \mathrm{after}}$ and $F_{488\ \mathrm{before}}$ ("before" and "after" refer to FRET) should be caused by environmental changes and/or interactions between **1**-streptavidin and **2**-avidin under the experimental conditions. This then can be used to correct the F_{430} for the same change and obtain the baseline for **1**-streptavidin and **2**-avidin according to:

$$F_{\mathrm{baseline\text{-}2}} = F_{430} \times (F_{488\ \mathrm{after}} / F_{488\ \mathrm{before}})$$

(iii) The approximate fluorescence of the reaction mixture before FRET was then the sum of the two components defined above:

$$F_{\mathrm{baseline}} = F_{\mathrm{baseline\text{-}1}} + F_{\mathrm{baseline\text{-}2}}$$

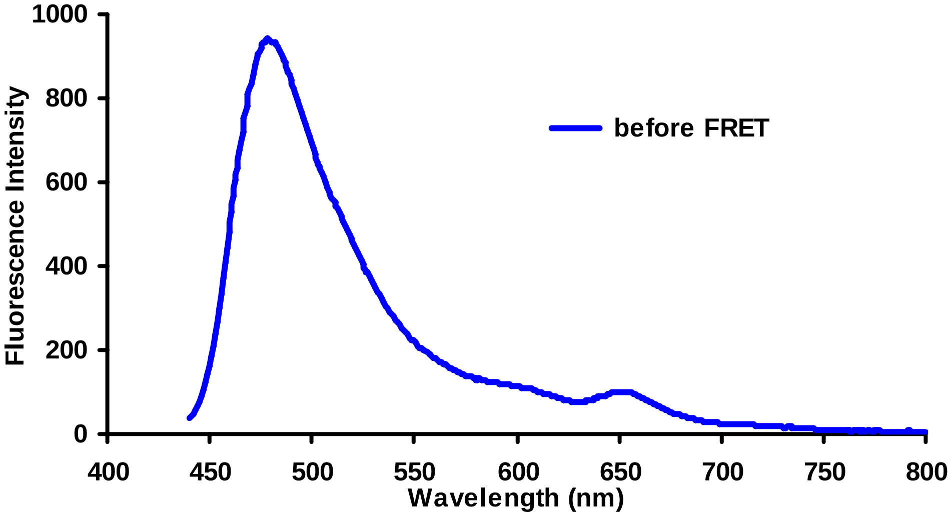


Figure F16. Fluorescence spectra for exp. (ii) before FRET (ex. 430 nm)

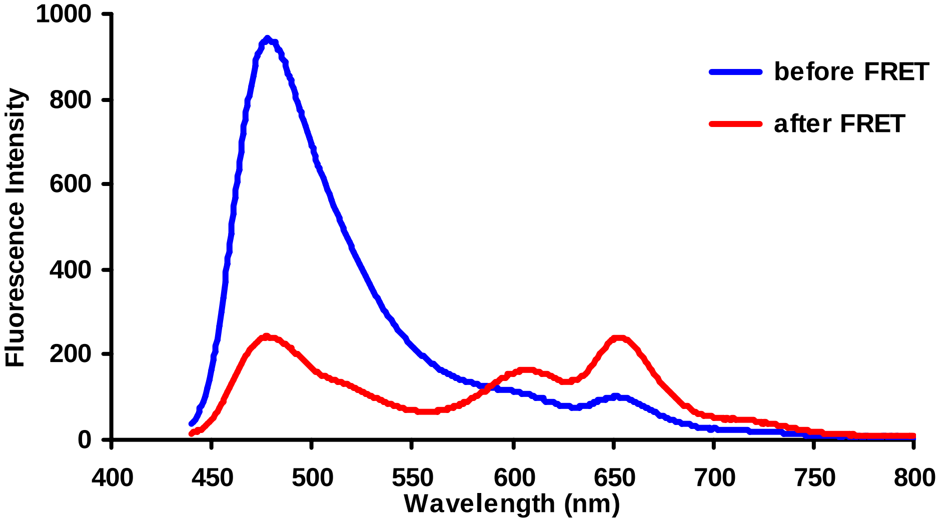


Figure F17. Fluorescence spectra for exp. (ii) before *and after* FRET (ex. 430 nm)

Experiment (i) (control):

In a fluorescence cuvette, **1**-streptavidin (3.0×10^{-7} M) and **2**-avidin (9.0×10^{-7} M) were pretreated with BSA (1.8×10^{-5} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto-425-BSA (2.0×10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described for experiment (ii).

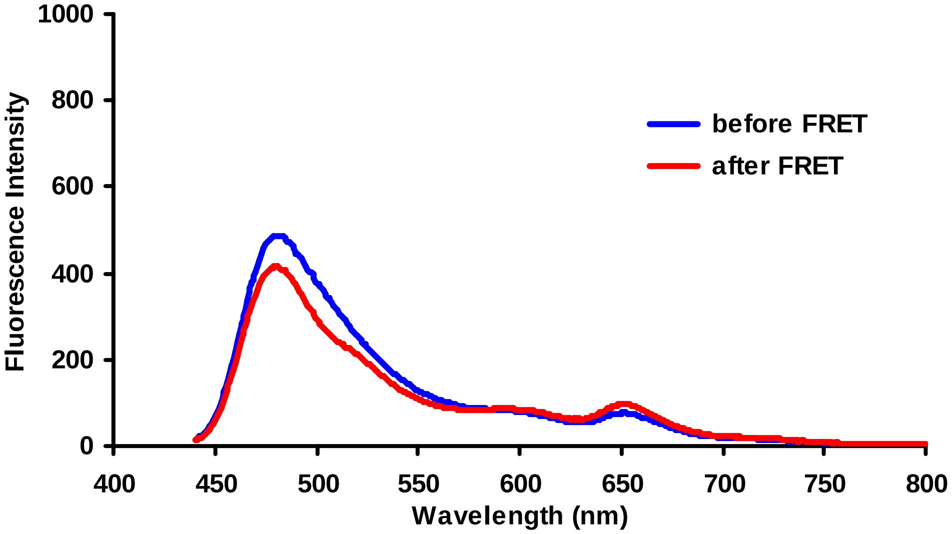


Figure F18. Fluorescence spectra for exp. (i), before and after FRET (ex. 430 nm)

Experiment (iii) (biotin block):

In a fluorescence cuvette, **1**-streptavidin (3.0 x 10^{-7} M) and **2**-avidin (9.0 x 10^{-7} M) were pretreated with BSA (1.8 x 10^{-5} M) and biotin (1.0 x 10^{-4} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto-425-BSA-biotin (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described for experiment (ii).

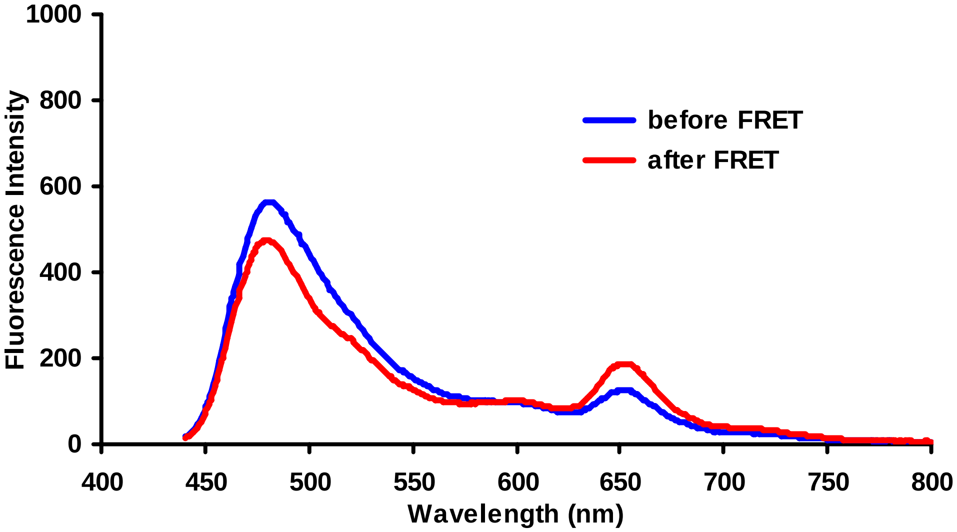


Figure F19. Fluorescence spectra for exp. (iii), before and after FRET (ex. 430 nm)

Experiment (iv) (selective binding):

In a fluorescence cuvette, **1**-BSA (3.0 x 10^{-7} M) and **2**-avidin (9.0 x 10^{-7} M) were pretreated with BSA (1.8 x 10^{-5} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto-425-BSA-biotin (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described for experiment (ii).

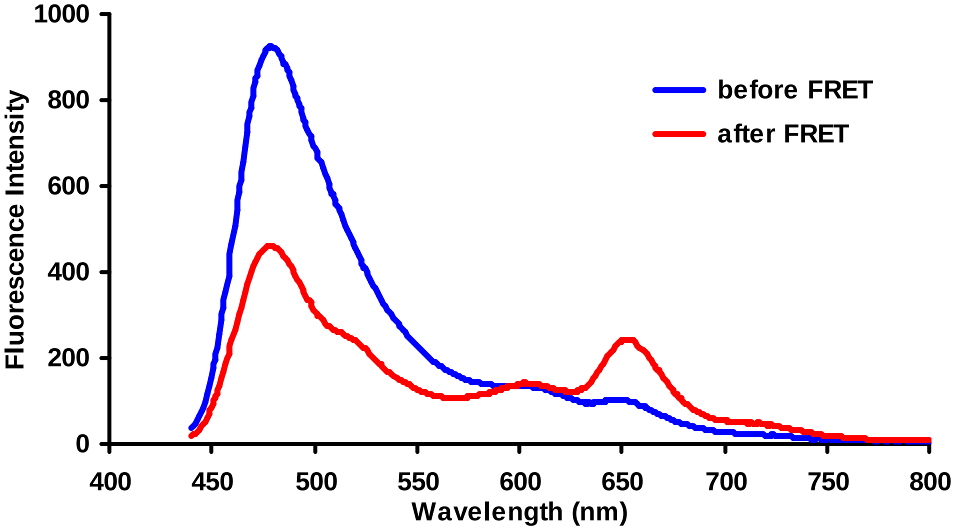


Figure F20. Fluorescence spectra for exp. (iv), before and after FRET (ex. 430 nm)

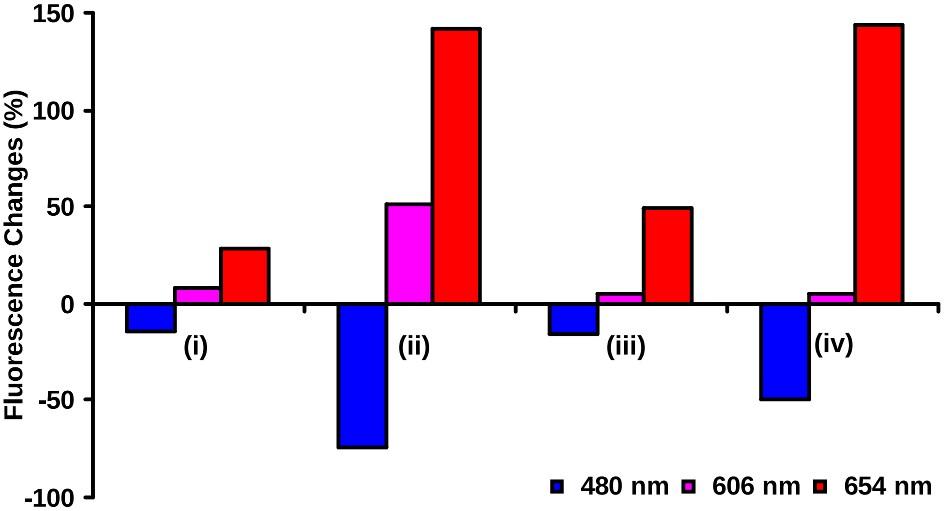


Figure F21. Overall fluorescence changes at 480 nm, 606 nm and 654 nm for exp (i)-(iv). (excited at 430 nm)

Data presented in Figure F21 represents the overall fluorescence changes. The next step in the process is to deconvolute the components of these changes that are due to interaction of Atto 425 with cassette **1** and cassette **2**. To do this it is important to understand that fluorescence at 480 nm should be due to the FRET-donor (Atto 425), since cassette **1** and cassette **2** do not fluorescence at 480 nm. However, fluorescence observed at 606 nm for the protein-protein interactions had components from cassette **1** *and* from Atto 425, but not from cassette **2** which does not emit at this wavelength. Similarly, fluorescence observed at 654 nm had components from *both* cassettes *and* from Atto 425. Thus the following steps were taken for deconvolution (throughout: excited at 430 nm):

(i) The emission at 480 nm was measured to reveal the fluorescence from Atto 425.

(ii) The value from step (i) was used to deduce the fluorescence from Atto 425 at 606 nm.

(iii) The value from step (ii) was subtracted from the observed value at 606 nm to reveal the fluorescence intensity attributed to cassette **1** alone.

(iv) The values from (i) and (iii) were used to deduce the fluorescence of Atto 425 and of cassette **1** at 654 nm.

(v) The values from step (iv) were subtracted from the observed value at 654 nm to reveal the fluorescence intensity attributed to cassette **2** alone.

Mathematically, an illustrative correction procedure is described for experiment (ii):
(1) Fluorescence changes at 480 nm is due to fluorescence changes of FRET donor (Atto-425-BSA-biotin), since **1**-streptavidin and **2**-avidin do not emit at 480 nm.

(2) (a) Fluorescence intensity of **1**-streptavidin at 606 nm before FRET ($F_{\text{before, }\mathbf{1}\text{-streptavidin, }606}$) can be read from the baseline ($F_{\text{baseline-2}}$) described above;

(b) The fluorescence intensity of **1**-streptavidin after FRET is given by:

$$F_{\text{after, }\mathbf{1}\text{-streptavidin, }606} = F_{\text{after, }606} - F_{\text{after, }480} * (F_{606}/F_{480})_{\text{Atto-425-BSA-biotin}}$$

where $F_{\text{after, }606}$, $F_{\text{after, }480}$ are the fluorescence intensity for the reaction mixture after FRET at 606nm and 480nm, respectively; $(F_{606}/F_{480})_{\text{Atto-425-BSA-biotin}}$ is the ratio of fluorescence intensity at 606nm and 480nm for Atto-425-BSA-biotin;

(c) The percentage fluorescence change for **1**-streptavidin at 606nm is given by:

$$(F_{\text{after, }\mathbf{1}\text{-streptavidin, }606} - F_{\text{before, }\mathbf{1}\text{-streptavidin, }606}) * 100 / F_{\text{before, }\mathbf{1}\text{-streptavidin, }606}$$

(3) (a) $F_{\text{before, }654}$ is read from the baseline value $F_{\text{baseline-2}}$ as described above, then that value is applied to the following equation to obtain the fluorescence intensity of **2**-avidin at 654 nm before FRET:

$$F_{\text{before, }\mathbf{2}\text{-avidin, }654} = F_{\text{before, }654} - F_{\text{before, }\mathbf{1}\text{-streptavidin, }606} * (F_{654}/F_{606})_{\mathbf{1}\text{-streptavidin}}$$

where $F_{\text{before, }\mathbf{1}\text{-streptavidin, 606}}$ is the fluorescence intensity of **1**-streptavidin at 606 nm before FRET; $(F_{654}/F_{606})_{\mathbf{1}\text{-streptavidin}}$ is the ratio of fluorescence intensity at 654 nm and 606 nm for **1**-streptavidin;

(b) The fluorescence intensity of **2**-avidin at 654 nm after FRET is calculated from:

$$F_{\text{after, }\mathbf{2}\text{-avidin, 654}} =$$
$$F_{\text{after, 654}} - F_{\text{after, 480}} * (F_{654}/F_{480})_{\text{Atto-425-BSA-biotin}} - F_{\text{after, }\mathbf{1}\text{-streptavidin, 606}} * (F_{654}/F_{606})_{\mathbf{1}\text{-streptavidin}}$$

where $(F_{654}/F_{480})_{\text{Atto-425-BSA-biotin}}$ is the ratio of fluorescence intensity at 654 nm and 480 nm for Atto-425-BSA-biotin;

(c) The percentage fluorescence change for **2**-avidin at 654nm is calculated from:

$$(F_{\text{after, }\mathbf{2}\text{-avidin, 654}} - F_{\text{before, }\mathbf{2}\text{-avidin, 654}}) * 100 / F_{\text{before, }\mathbf{2}\text{-avidin, 654}}$$

Similar corrections were made to experiment (i), (iii) and (iv). The data shown in Figure F22 are that after all these corrections were applied.

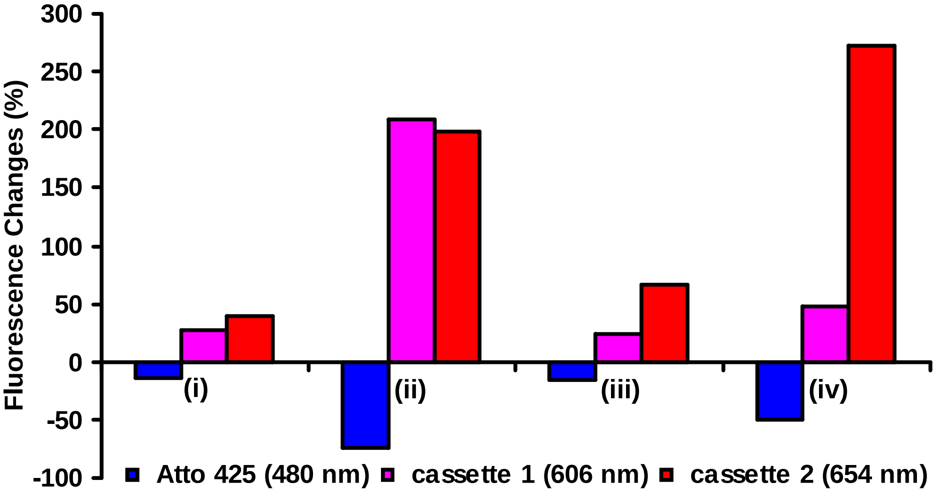


Figure F22. Fluorescence changes for Atto 425 (480 nm), cassette 1 (606 nm) and cassette 2 (654 nm) in exp (i)-(iv)

FRET Efficiency

The stoichiometry of BSA-biotin/avidin/streptavidin complexes formed in the experiments (i)-(iv) was impossible to determine. Consequently, the FRET efficiency for the reaction was calculated for the overall energy transfer, which is the energy transfer from the FRET donor (Atto425) to cassettes **1** and **2**.

FRET efficiency (E) was calculated from:

$$E = 1- F_{after} / F_{before}$$

where F_{after} and F_{before} are the FRET donor (Atto 425) fluorescence intensity after and before FRET, respectively.

Table F2. FRET efficiency for experiments (i)-(iv)

	experiment (**i**)	experiment (**ii**)	experiment (**ii**)	experiment (**iv**)
FRET efficiency (%)	15	74	16	50

(3) *In-vitro* two protein interactions

(i) In a fluorescence cuvette, **1**-streptavidin (3.0 x 10^{-7} M) was pretreated with BSA (9.0 x 10^{-6} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto425-BSA-biotin (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described above.

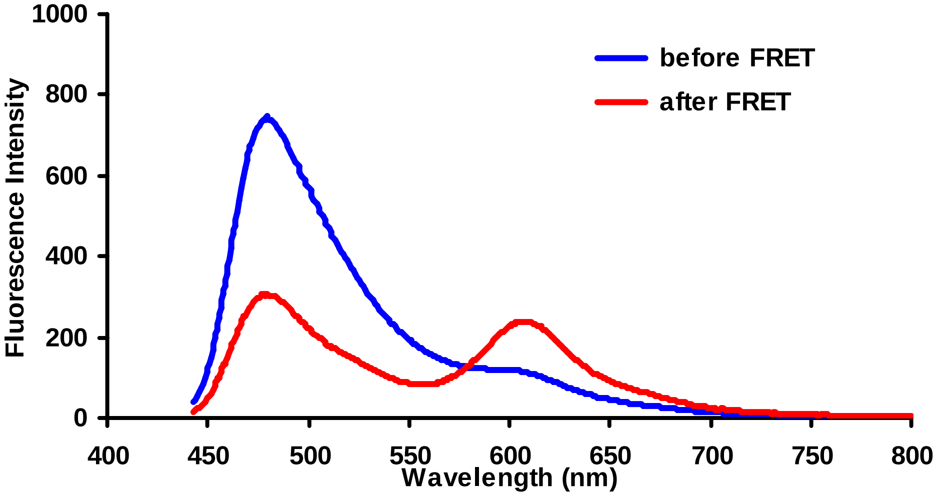


Figure F23. Fluorescence spectra for exp. (i), before and after FRET (ex. 430 nm)

(ii) In a fluorescence cuvette, **1**-streptavidin (3.0 x 10^{-7} M) was pretreated with BSA (9.0 x 10^{-6} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto425-BSA (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described above.

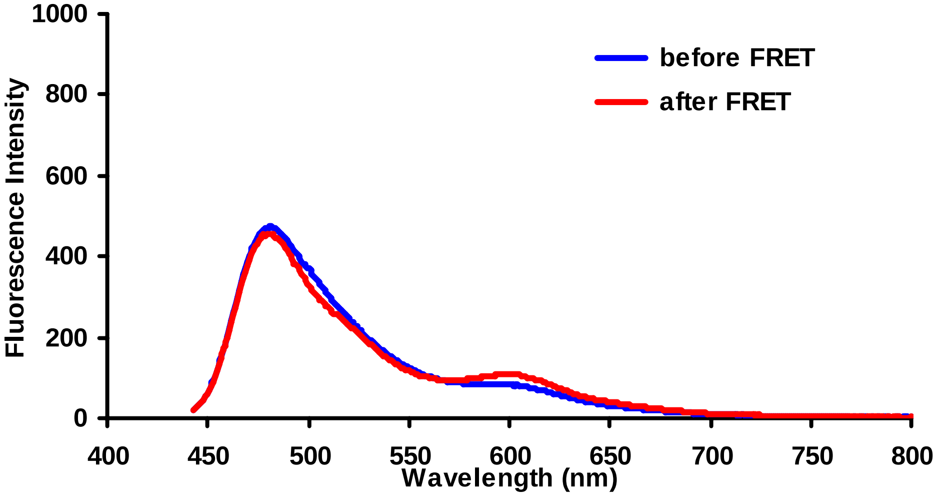


Figure F24. Fluorescence spectra for exp. (ii), before and after FRET (ex. 430 nm)

(iii) In a fluorescence cuvette, **2**-avidin (9.0 x 10^{-7} M) was pretreated with BSA (9.0 x 10^{-6} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto425-BSA-biotin (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described above.

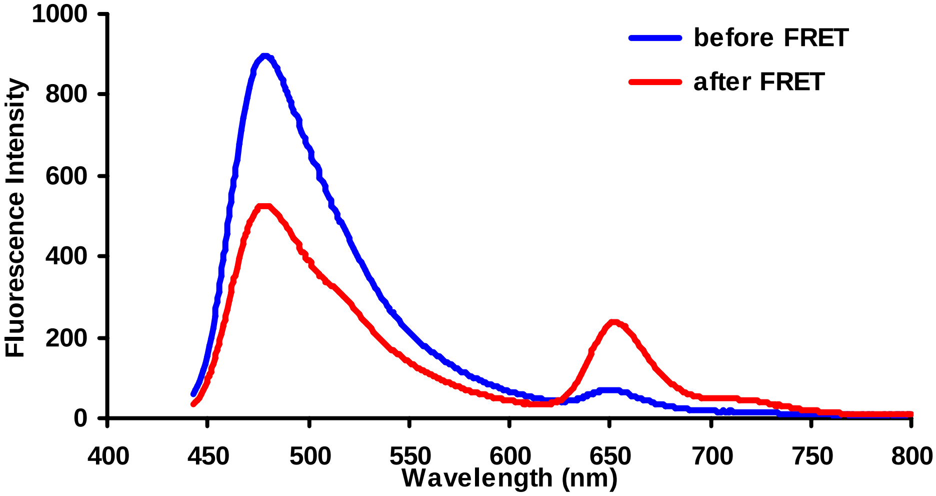


Figure F25. Fluorescence spectra for exp. (iii), before and after FRET (ex. 430 nm)

(iv) In a fluorescence cuvette, **2**-avidin (9.0 x 10^{-7} M) was pretreated with BSA (9.0 x 10^{-6} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto425-BSA (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described above.

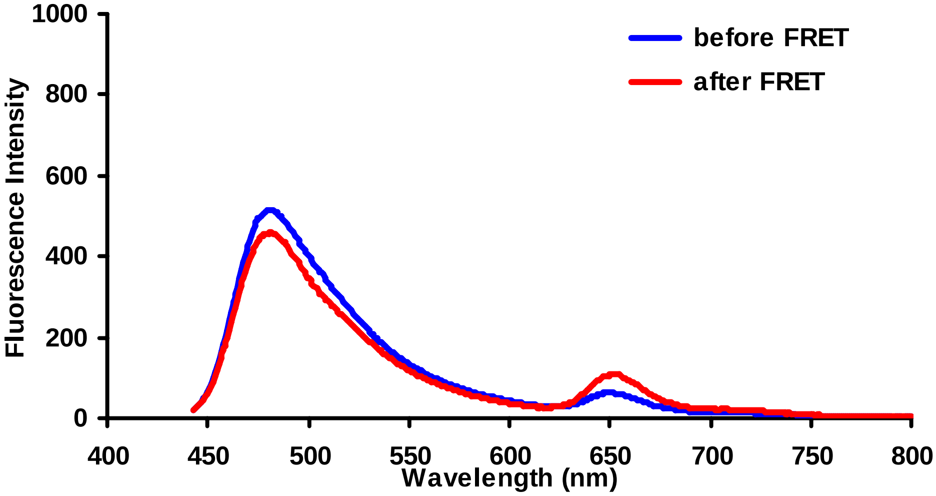


Figure F26. Fluorescence spectra for exp. (iv), before and after FRET (ex. 430 nm)

(v) In a fluorescence cuvette, **1**-BSA (3.0 x 10^{-7} M) was pretreated with BSA (9.0 x 10^{-6} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto425-BSA-biotin (2.0 x 10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described above.

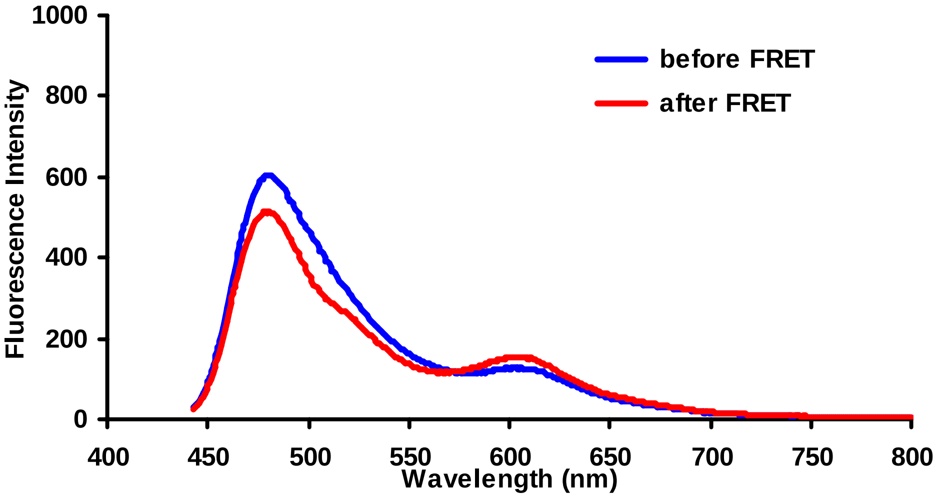


Figure S27. Fluorescence spectra for exp. (v), before and after FRET (ex. 430 nm)

(vi) In a fluorescence cuvette, **1**-BSA (3.0×10^{-7} M) was pretreated with BSA (9.0×10^{-6} M) in 0.1 M phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto425-BSA (2.0×10^{-7} M) was added. Fluorescence (ex. 430 nm) of the solution was taken after 10 min. The baseline for this experiment was obtained using similar methods as described above.

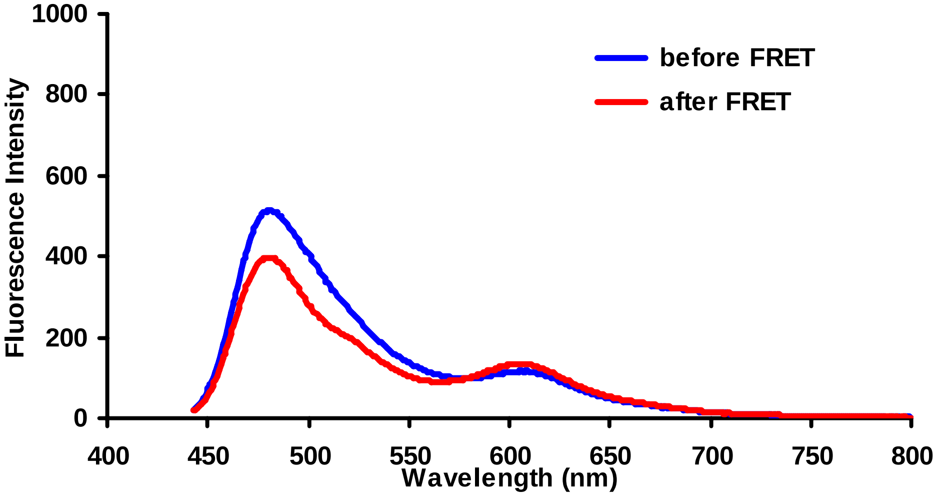


Figure F28. Fluorescence spectra for exp. (vi), before and after FRET (ex. 430 nm)

(4) *Ex vivo* protein-protein interaction studies

a. Cell culture

COS-7 cells (American Type Culture Collection) were cultured as subconfluent monolayers on 75 cm^2 culture flask with vent caps in DMEM supplemented with 10 % fetal bovine serum (FBS) in a humidified incubator at 37 ºC with 5 % CO_2. Cells grown to subconfluence were enzymatically dissociated from the surface with trypsin and plated 2-3 days prior to the experiments in Lab-Tek two well chambered coverglass slides (Nunc).

b. Fluorescence microscopy

Protein interactions were studied on living COS7 cells using a Zeiss 510 META NLO Multiphoton Microscope System consisting of an Axiovert 200 MOT microscope. Digital images of streptavidin-**1** (acceptor 1), avidin-**2** (acceptor 2) and biotin-BSA-**Atto425** (FRET donor) were captured with a 40x / 1.3 oil objective with the following filter sets: Excitation 458 nm; Emission BP 480-520 for BSA-**Atto425** (FRET donor); Emission BP 565-615 for streptavidin-**1** (acceptor 1); and Emission BP 640-704 for avidin-**2** (acceptor 2).

The protein:carrier complexes were pre-formed at room temperature in ACAS medium for 30 min by mixing (in a mol:mol ratio) the labeled protein and the carrier, Pep-1 (Active Motif). The different complexes were formed as follows:

- FRET donor (biotin-BSA-**Atto425**) : Pep-1 : 1:20 mol:mol ratio of protein:carrier
- Streptavidin-**1** (acceptor 1) : Pep-1 : 1:15 mol:mol ratio of protein:carrier
- Avidin-**2** (acceptor 2) : Pep-1 : 1:10 mol:mol ratio of protein:carrier
- Streptavidin-**1** (acceptor 1) : Avidin-**2** (acceptor 2) : Pep-1 : 1:3:50 mol:mol ratio of protein:carrier

c. *Ex-vivo* three protein interactions

First, the FRET donor was imported inside COS-7 cells by treating the cells with the FRET donor (biotin-BSA-**Atto425**) : Pep-1 complex. After a 10 min incubation period at 37 °C, the cells were washed several times with PBS, then the Streptavidin-**1** (acceptor 1) : Avidin-**2** (acceptor 2) : Pep-1 complex was added, and the cells were incubated for another 10 min at 37 °C. The cells were washed several times with PBS buffer and then put on the stage of the Zeiss 510.

Energy Transfer data in COS-7 cells loaded with FRET donor (biotin-BSA-**Atto425**), Streptavidin-**1** and Avidin-**2** were collected using 458 nm excitation wavelength. Emission of FRET donor (biotin-BSA-**Atto425**) (blue channel; donor signal) was collected using a BP 480-520 emission filter whereas emissions of Streptavidin-**1** and Avidin-**2** (FRET channel; acceptor signals) were collected using a BP 565-615 and BP 640-704 filters, respectively. Donor bleed through signal to the FRET channel was calculated by measuring the FRET channel signal resulting from COS cells loaded only with the donor. Acceptor bleed through to the FRET channel was calculated by measuring the FRET channel signal resulting from COS cells loaded with Streptavidin-**1** and Avidin-**2** alone.

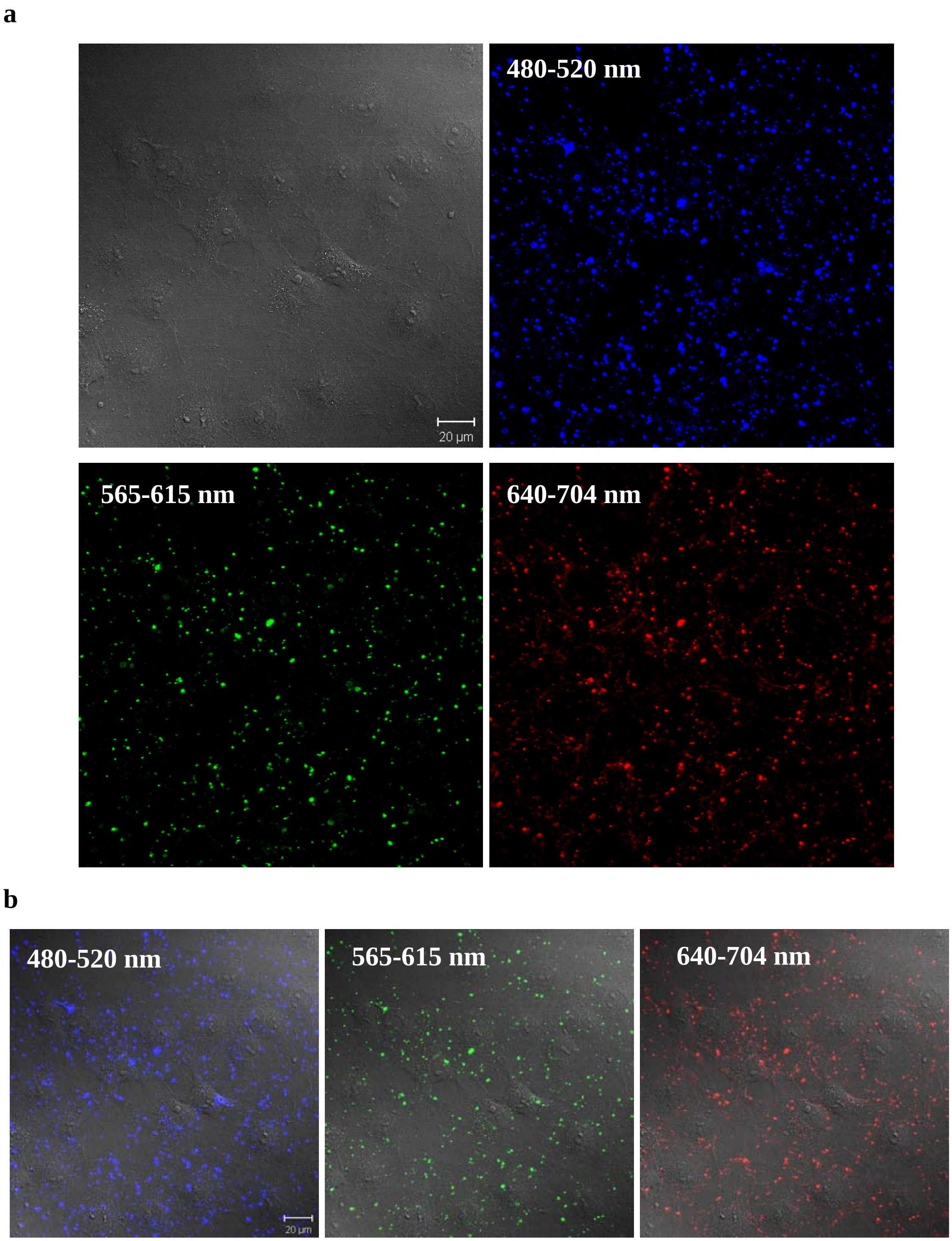

a
480-520 nm
20 µm
565-615 nm
640-704 nm
b
480-520 nm
565-615 nm
640-704 nm
20 µm

Figure F29. Three protein interactions in living cells (biotin-BSA-**Atto425** + Streptavidin-**1** + Avidin-**2**): **a** DIC and fluorescence images of cells; **b** overlay

d. *Ex-vivo* two protein interactions (donor and acceptor 1; donor and acceptor 2)

As described above, cells were first treated with the FRET donor (10 min at 37 °C), followed by the Acceptor 1 or Acceptor 2 (10 min at 37 °C).

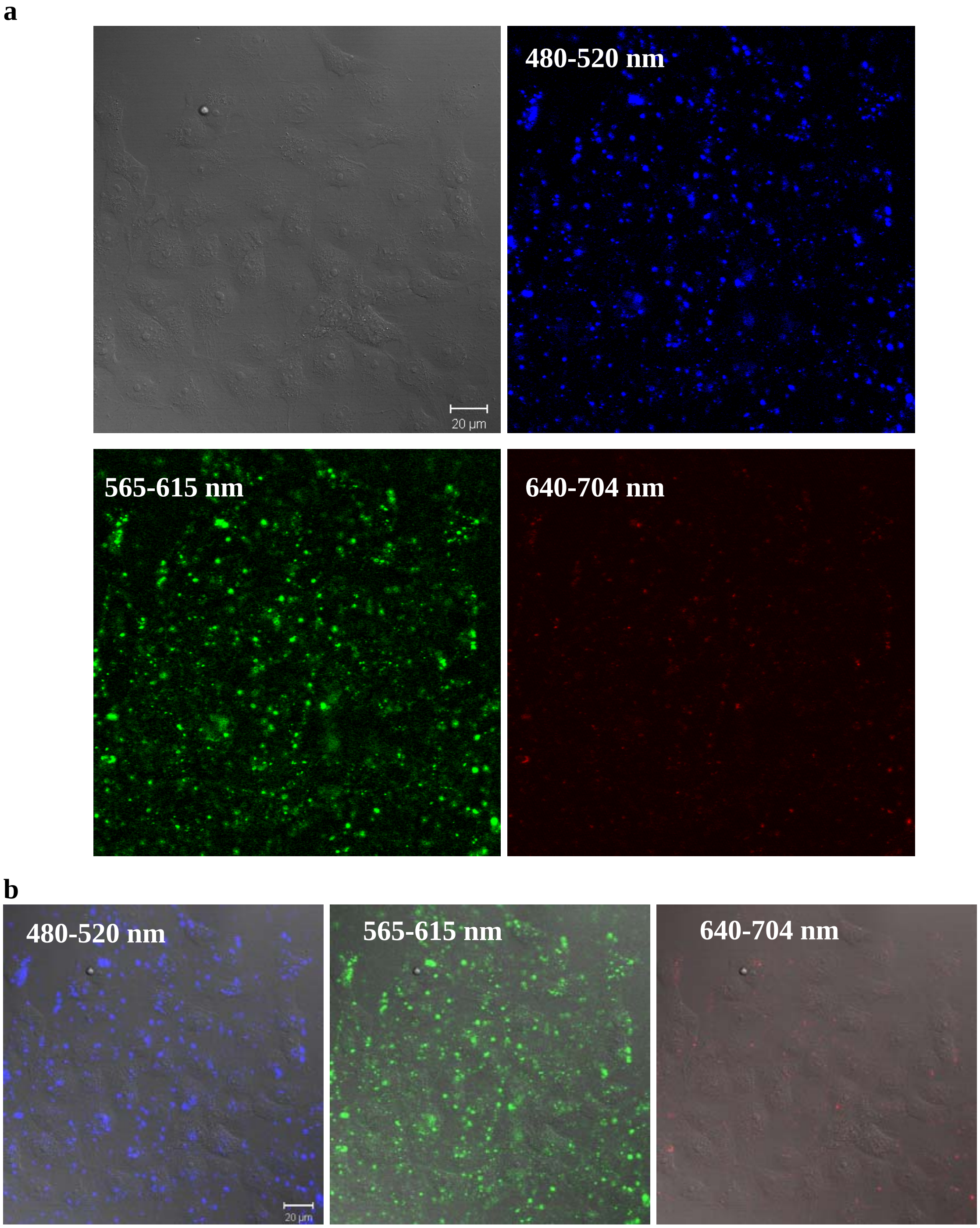

a
480-520 nm
20 µm
565-615 nm
640-704 nm
b
480-520 nm
565-615 nm
640-704 nm
20 µm

Figure F30. Two protein interactions in living cells (biotin-BSA-**Atto425** + Streptavidin-**1**): **a** DIC and fluorescence images of cells; **b** overlay

Donor and Acceptor 2:

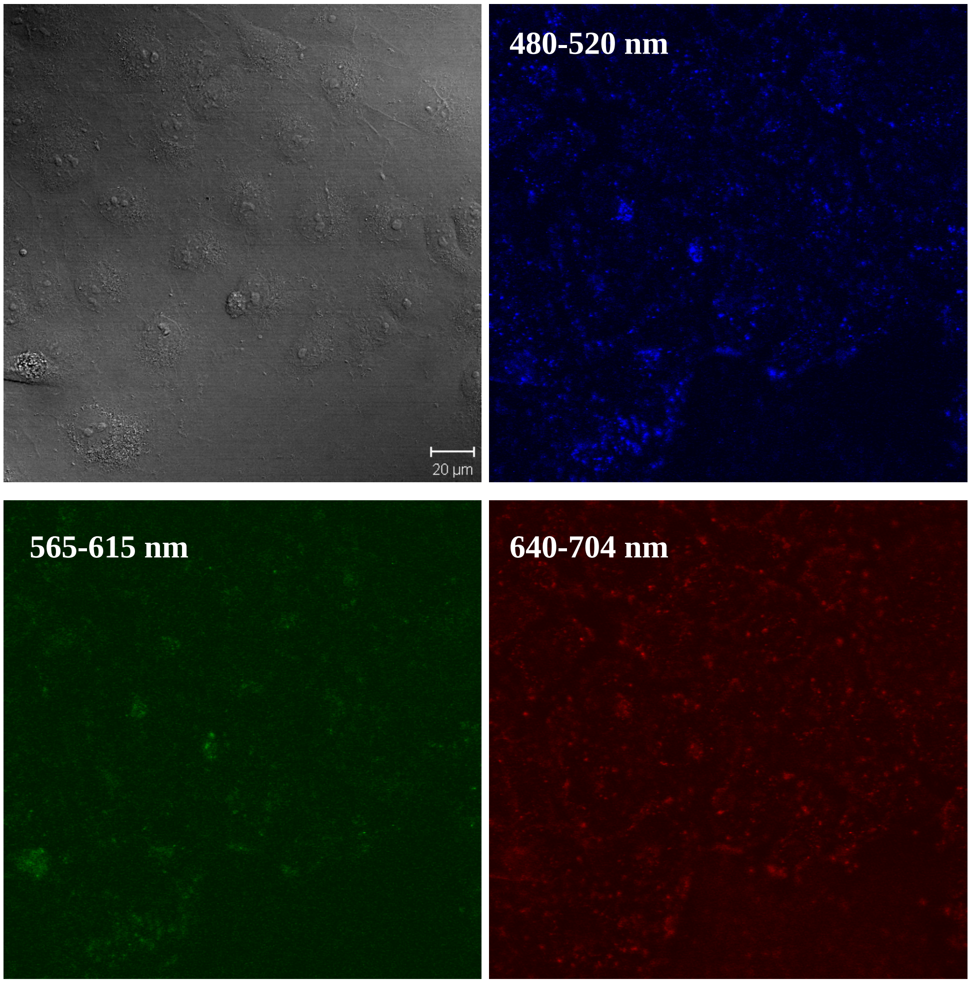

480-520 nm
20 µm
565-615 nm
640-704 nm

Figure F31. Two protein interactions (biotin-BSA-**Atto425** + Avidin-**2**): DIC and fluorescence images of cells.

e. *Ex-vivo* control experiments

Energy Transfer data in untreated COS-7 cells *i.e.* cell autofluorescence and in cells loaded uniquely with FRET donor (biotin-BSA-**Atto425**), Streptavidin-**1** and Avidin-**2** were collected using 458 nm excitation wavelength. Emission of FRET donor (biotin-BSA-**Atto425**) (blue channel; donor signal) was collected using a BP 480-520 emission filter whereas emissions of Streptavidin-**1** and Avidin-**2** (FRET channel; acceptor signals) were collected using BP 565-615 and BP 640-704 emission filters, respectively. Donor bleed through signal to the FRET channel was calculated by measuring the FRET channel signal resulting from COS cells loaded only with the donor. Acceptor bleed through to the FRET channel was calculated by measuring the FRET channel signal resulting from COS cells loaded with Streptavidin-**1** and Avidin-**2** alone.

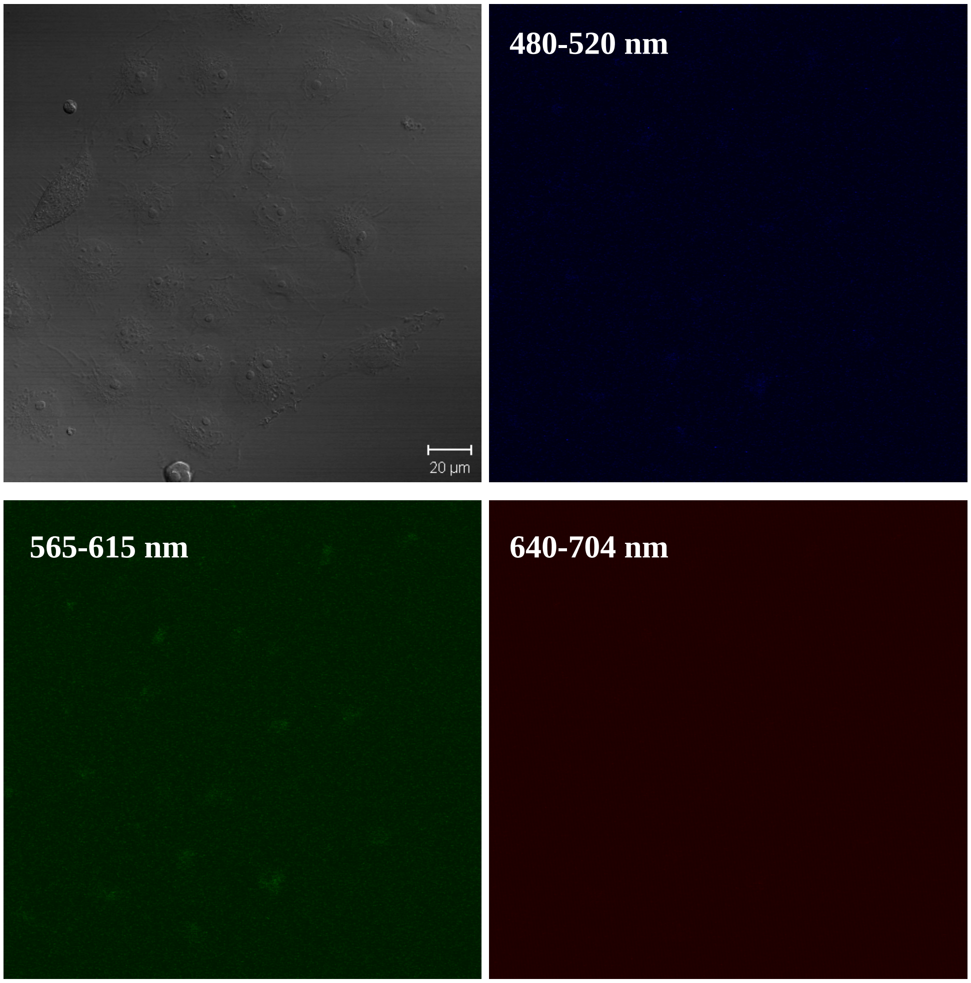


Figure F32. Cell autofluorescence

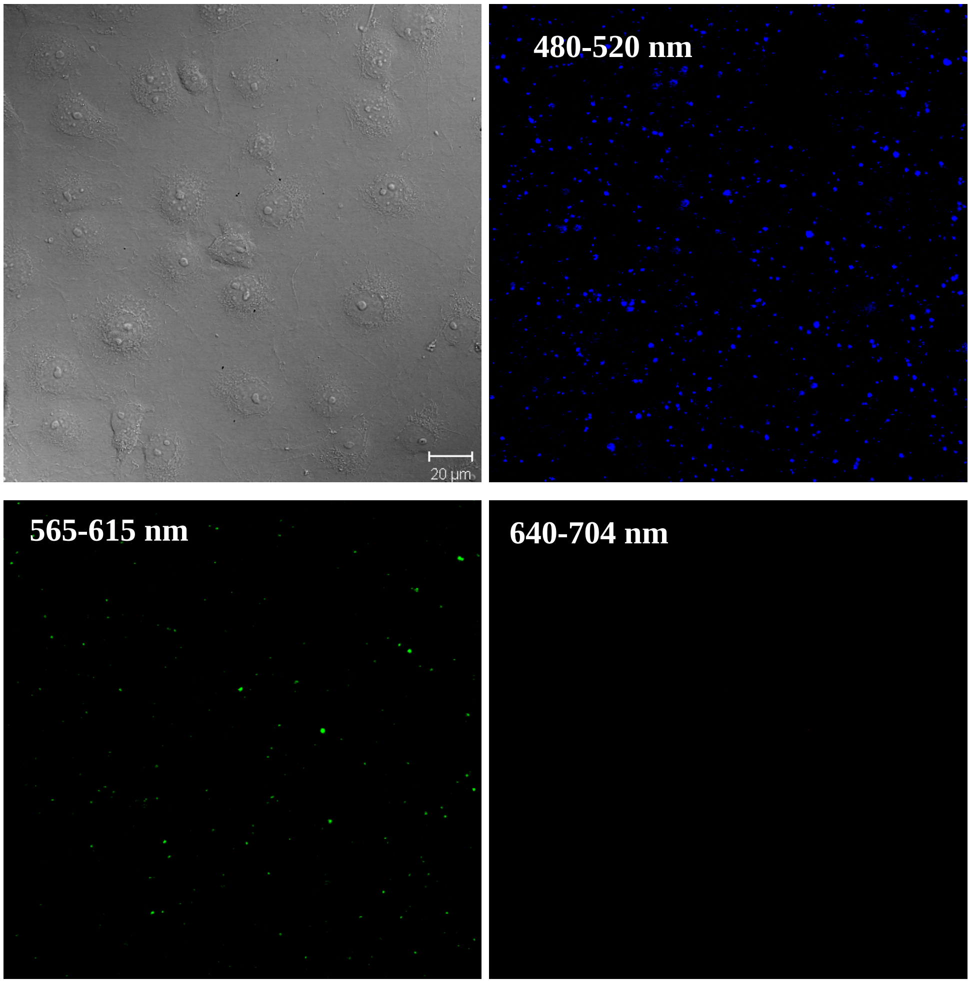

480-520 nm
20 µm
565-615 nm
640-704 nm

Figure F33. Biotin-BSA-**Atto425** (FRET Donor) only

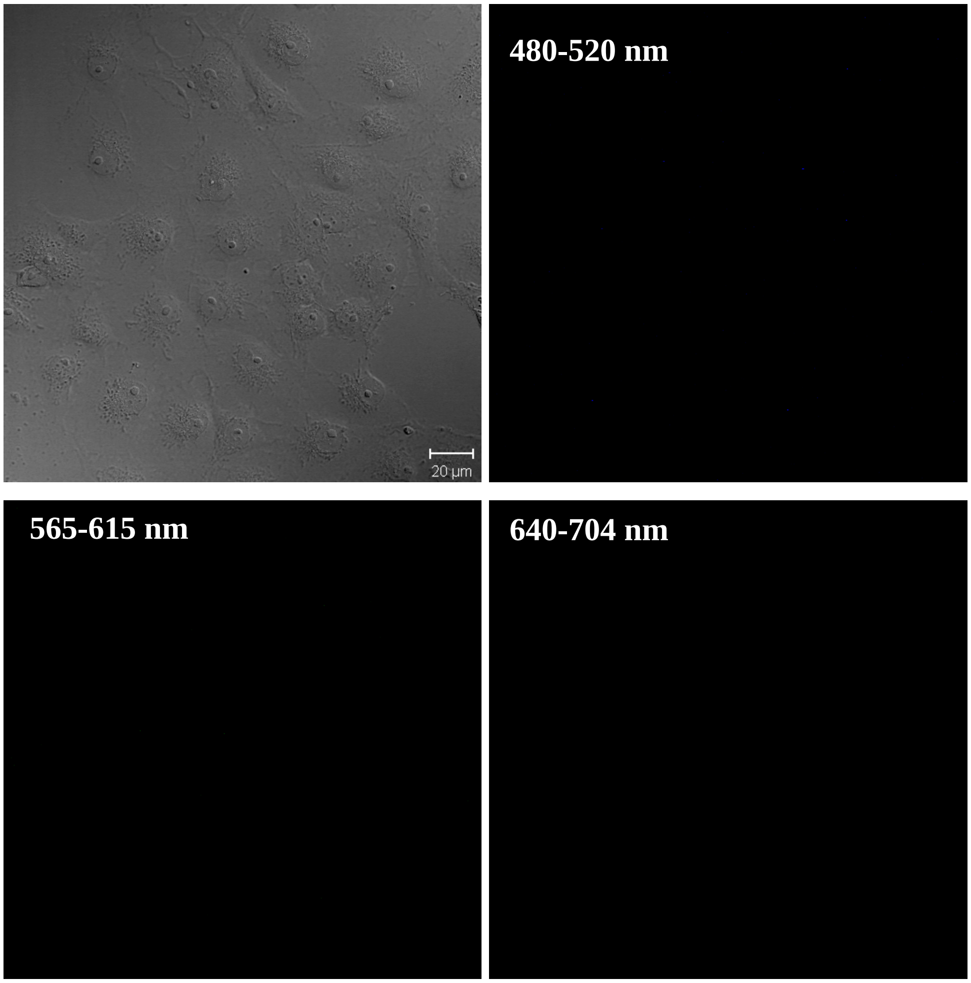

20 µm
480-520 nm
565-615 nm
640-704 nm

Figure F34. Streptavidin-**1 (**Acceptor 1) only

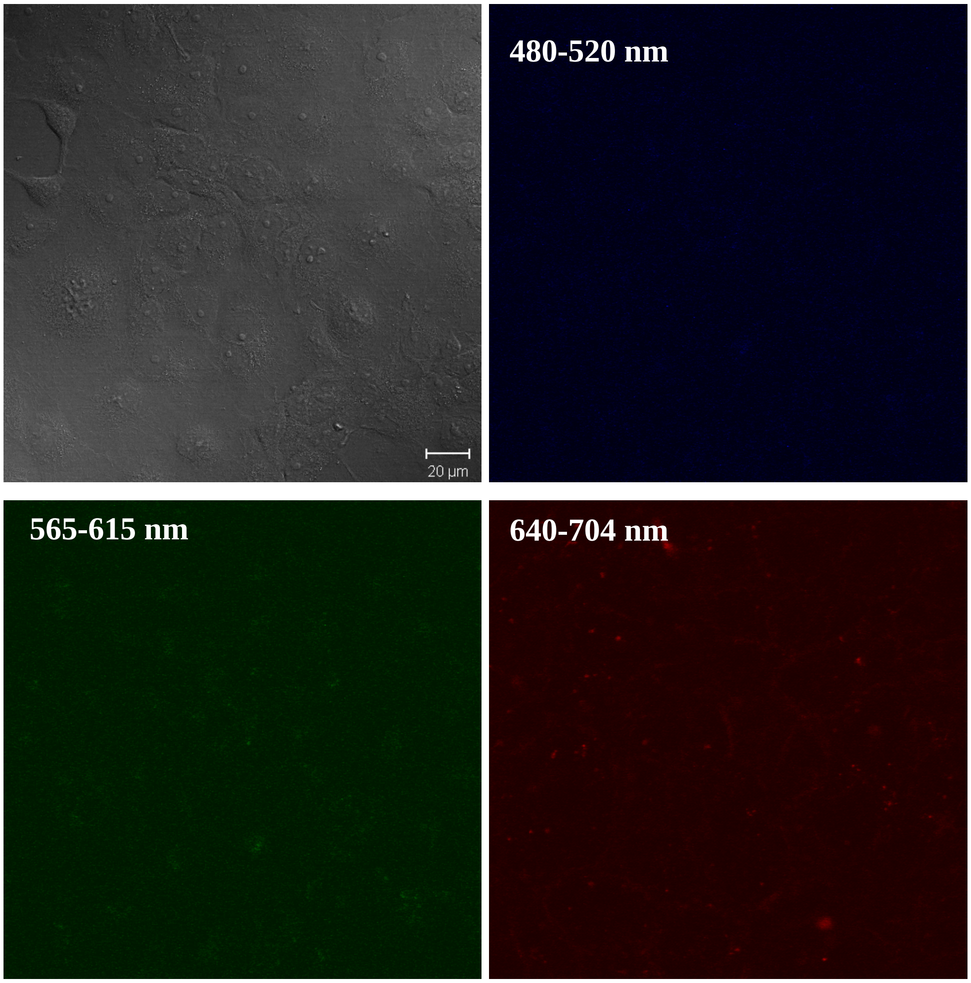

20 µm
480-520 nm
565-615 nm
640-704 nm

Figure F35. Avidin-**2** (Acceptor 2) only

VITA

Name: Liangxing Wu

Permanent Address:
Xinjie, Erlang, Dawu
Xiaogan, Hubei, China

Education:
2003-2009, Ph.D., Chemistry
Texas A&M University
College Station, TX 77842

1999-2003, B.S., Chemistry
University of Science and Technology of China
Hefei, Anhui, China