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 throughout 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

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF SCHEMES</td>
<td>xv</td>
</tr>
<tr>
<td>CHAPTER I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Fluorescence</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Multiplexing with FRET and TBET</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Labeling of proteins</td>
<td>8</td>
</tr>
<tr>
<td>1.3.1 Covalent labeling techniques</td>
<td>8</td>
</tr>
<tr>
<td>1.3.2 Non-covalent labeling methods</td>
<td>17</td>
</tr>
<tr>
<td>1.3.3 Summary</td>
<td>20</td>
</tr>
<tr>
<td>1.4 Delivery of protein into cells</td>
<td>20</td>
</tr>
<tr>
<td>1.5 Detection of protein-protein interactions</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER II DESIGN AND SYNTHESSES OF HIGHLY FLUORESCENT</td>
<td>25</td>
</tr>
<tr>
<td>GFP-CHROMOPHORE ANALOGS</td>
<td></td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>25</td>
</tr>
<tr>
<td>2.2 Results and discussions</td>
<td>33</td>
</tr>
<tr>
<td>2.2.1 Syntheses of the highly fluorescent GFP-chromophore analogs</td>
<td>33</td>
</tr>
<tr>
<td>2.2.2 UV and fluorescence properties of the GFP-chromophore analogs</td>
<td>48</td>
</tr>
<tr>
<td>2.3 Conclusions</td>
<td>54</td>
</tr>
<tr>
<td>CHAPTER III A NEW SYNTHESIS OF BORAINDACENE (BODIPY) DYES</td>
<td>56</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>56</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>3.2</td>
<td>A new approach to symmetric BODIPYs</td>
</tr>
<tr>
<td>3.3</td>
<td>Kinetic and mechanistic studies of the new reaction</td>
</tr>
<tr>
<td>3.4</td>
<td>Conclusion</td>
</tr>
<tr>
<td>Chapter IV</td>
<td>ROSAMINES WITH CYCLIC AMINE SUBSTITUENTS: SYNTHESES AND SPECTROSCOPIC PROPERTIES AND APPLICATIONS</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>4.2</td>
<td>Syntheses of aminoxanthones and rosamines</td>
</tr>
<tr>
<td>4.3</td>
<td>Spectroscopic properties of rosamine derivatives</td>
</tr>
<tr>
<td>4.4</td>
<td>Biological activity test</td>
</tr>
<tr>
<td>4.5</td>
<td>Conclusions</td>
</tr>
<tr>
<td>Chapter V</td>
<td>NEW PYRONIN DYSES WITH MESO-HETEROATOM SUBSTITUENTS: SYNTHESIS AND SPECTROSCOPIC PROPERTIES</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>5.2</td>
<td>Syntheses of new pyronin derivatives</td>
</tr>
<tr>
<td>5.3</td>
<td>Photophysical properties of new pyronin dyes</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusion</td>
</tr>
<tr>
<td>Chapter VI</td>
<td>WATER-SOLUBLE THROUGH-BOND ENERGY TRANSFER CASSETTES: SYNTHESES, SPECTROSCOPIC PROPERTIES AND APPLICATIONS</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>6.2</td>
<td>Syntheses of water-soluble through-bond energy transfer cassettes</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Syntheses of donor fragments</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Syntheses of acceptor parts</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Syntheses of cassettes</td>
</tr>
<tr>
<td>6.3</td>
<td>Spectroscopic properties of BODIPY acceptors and cassettes</td>
</tr>
<tr>
<td>6.4</td>
<td>Protein-protein interaction studies</td>
</tr>
<tr>
<td>6.5</td>
<td>Conclusions</td>
</tr>
<tr>
<td>Chapter VII</td>
<td>CONCLUSIONS AND OUTLOOK</td>
</tr>
<tr>
<td>7.1</td>
<td>Conclusions</td>
</tr>
<tr>
<td>7.1.1</td>
<td>New methodologies for the syntheses of BODIPY and rosamine</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Syntheses of novel fluorescent compounds</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>7.1.3 Monitoring of protein-protein interactions <em>in vitro</em> and in living cells</td>
<td>158</td>
</tr>
<tr>
<td>7.2 Outlook</td>
<td>160</td>
</tr>
<tr>
<td>7.2.1 Rosamine derivatives</td>
<td>160</td>
</tr>
<tr>
<td>7.2.2 Water-soluble through-bond energy transfer cassettes</td>
<td>162</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>167</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>182</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>184</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>243</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>269</td>
</tr>
<tr>
<td>APPENDIX E</td>
<td>314</td>
</tr>
<tr>
<td>APPENDIX F</td>
<td>332</td>
</tr>
<tr>
<td>VITA</td>
<td>482</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A typical Jablonski diagram</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Jablonski diagram for two photon excitation</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>(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.</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Dependence of the orientation factor $k^2$ on the direction of the dipoles.</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>A typical diagram for through-bond energy transfer cassette</td>
<td>7</td>
</tr>
<tr>
<td>1.6</td>
<td>Labeling of proteins by modification of amino acid sidechains</td>
<td>9</td>
</tr>
<tr>
<td>1.7</td>
<td>Labeling of proteins by modification of terminal amino acids</td>
<td>11</td>
</tr>
<tr>
<td>1.8</td>
<td>Other direct labeling methods</td>
<td>13</td>
</tr>
<tr>
<td>1.9</td>
<td>Illustrative examples of enzyme catalyzed labeling of proteins</td>
<td>16</td>
</tr>
<tr>
<td>1.10</td>
<td>Illustrative examples of non-covalent labeling of proteins</td>
<td>18</td>
</tr>
<tr>
<td>1.11</td>
<td>Proposed system for detecting protein-protein interactions</td>
<td>24</td>
</tr>
<tr>
<td>2.1</td>
<td>(a) X-ray structure; (b) chromophore of GFP; (c) schematic diagram of the interactions between the chromophore and its surroundings in the S65T mutant.</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Biogenesis of the chromophore in wild-type GFP</td>
<td>28</td>
</tr>
<tr>
<td>2.3</td>
<td>Structure of the red fluorescent protein (DsRed) chromophore</td>
<td>30</td>
</tr>
<tr>
<td>2.4</td>
<td>Some synthetic analogs of chromophore in fluorescent proteins (FP)</td>
<td>31</td>
</tr>
<tr>
<td>2.5</td>
<td>Thought process that led to the design of analogs 1 and the derivatives discussed in this chapter</td>
<td>32</td>
</tr>
<tr>
<td>2.6</td>
<td>Isomerization of $E,Z$-mixtures of compound 4c.</td>
<td>37</td>
</tr>
<tr>
<td>FIGURE</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>2.7.</td>
<td>Trends in chemical shift differences for compounds 5 relative to compounds 1 illustrate that these can be used to differentiate between them.</td>
<td>39</td>
</tr>
<tr>
<td>2.8.</td>
<td>Top and side views of structures based on single molecule X-ray diffraction analyses of (a) compound 1c; and (b) compound 5c.</td>
<td>41</td>
</tr>
<tr>
<td>2.9.</td>
<td>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).</td>
<td>49</td>
</tr>
<tr>
<td>2.10.</td>
<td>UV absorbance and fluorescence for: (a) GFP-chromophore analogs 1 in MeOH; and, (b) GFP-chromophore analogs 5 in MeOH.</td>
<td>51</td>
</tr>
<tr>
<td>2.11.</td>
<td>Normalized UV absorbance and fluorescence of 1h and the labeled streptavidin 1s in 0.1 M lithium phosphate buffer, pH 7.4.</td>
<td>54</td>
</tr>
<tr>
<td>3.1.</td>
<td>Spectra of BODIPYs 9 in CH_2Cl_2.</td>
<td>65</td>
</tr>
<tr>
<td>3.2.</td>
<td>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.</td>
<td>67</td>
</tr>
<tr>
<td>3.3.</td>
<td>(a) and (b) Two possible mechanisms for the formation of dipyrrromethene.</td>
<td>69</td>
</tr>
<tr>
<td>4.1.</td>
<td>Structures of some common rhodamines.</td>
<td>71</td>
</tr>
<tr>
<td>4.2.</td>
<td>Syntheses of rosamine dyes via: (a) the conventional condensation approach; and, (b) nucleophilic additions to 3,6-diamino-xanthones.</td>
<td>75</td>
</tr>
<tr>
<td>4.3.</td>
<td>Reported syntheses of 3,6-diaminesubstituted-xanthones.</td>
<td>76</td>
</tr>
<tr>
<td>4.4.</td>
<td>X-ray structures for rosamines (a) 26 and (b) 27.</td>
<td>85</td>
</tr>
<tr>
<td>4.5.</td>
<td>Normalized (a) UV-Vis absorption (10^{-6} M); and, (b) fluorescence spectra (10^{-7} M, excited at λ_{max \text{abs}}), for representative rosamines in CH_2Cl_2.</td>
<td>91</td>
</tr>
<tr>
<td>4.6.</td>
<td>Fluorescence intensity (excited at 562 nm) of 11i at 1.4x10^{-7} M in a CH_2Cl_2/MeOH mixed solvent system.</td>
<td>93</td>
</tr>
</tbody>
</table>
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) ................................................................. 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$_2$Cl$_2$ ($ca \ 10^{-6}$ 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
<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>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)</td>
<td>141</td>
</tr>
<tr>
<td>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</td>
<td>142</td>
</tr>
<tr>
<td>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)</td>
<td>145</td>
</tr>
<tr>
<td>6.11. FRET from one protein can be used to excite the donors of TBET-cassettes on another protein to reveal interactions</td>
<td>147</td>
</tr>
<tr>
<td>6.12. <em>In vitro</em> two protein interaction studies</td>
<td>149</td>
</tr>
<tr>
<td>6.13. <em>In vitro</em> three protein interactions</td>
<td>151</td>
</tr>
<tr>
<td>6.14. Studies of protein interactions in COS-7 cells</td>
<td>152</td>
</tr>
<tr>
<td>7.1. Highly fluorescent GFP-chromophore analog 1c and its precursor 4c</td>
<td>157</td>
</tr>
<tr>
<td>7.2. Novel pyronin derivatives</td>
<td>157</td>
</tr>
<tr>
<td>7.3. Illustrative examples of water-soluble BODIPY acceptor and through-bond energy transfer cassette</td>
<td>158</td>
</tr>
<tr>
<td>7.4. Scheme for detecting protein-protein interactions</td>
<td>159</td>
</tr>
<tr>
<td>7.5. Rosamine derivatives for anti-tumor activity study</td>
<td>161</td>
</tr>
<tr>
<td>7.6. Water-soluble through-bond energy transfer cassettes</td>
<td>162</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Important spectroscopic characteristics of fluorophores</td>
<td>3</td>
</tr>
<tr>
<td>1.2. Summary of protein labeling methods</td>
<td>19</td>
</tr>
<tr>
<td>2.1. Spectroscopic properties of $Z$- and $E-4c$ isomers</td>
<td>35</td>
</tr>
<tr>
<td>2.2. Spectroscopic properties of $1c$ and $5c$</td>
<td>38</td>
</tr>
<tr>
<td>2.3. Spectroscopic properties of $4c$, $1c$ and $5c$ in MeOH</td>
<td>49</td>
</tr>
<tr>
<td>2.4. Spectroscopic properties of GFP-chromophore analogs</td>
<td>53</td>
</tr>
<tr>
<td>3.1. Synthesis of symmetric BODIPYs 9</td>
<td>63</td>
</tr>
<tr>
<td>3.2. Photophysical properties of BODIPYs 9 in CH$_2$Cl$_2$</td>
<td>64</td>
</tr>
<tr>
<td>4.1. Amination of the ditriflyl xanthone 16</td>
<td>78</td>
</tr>
<tr>
<td>4.2. Synthesis of rosamine derivatives</td>
<td>80</td>
</tr>
<tr>
<td>4.3. Spectral properties of rosamines</td>
<td>89</td>
</tr>
<tr>
<td>4.4. Spectral properties of rosamine 11i in different solvents</td>
<td>92</td>
</tr>
<tr>
<td>4.5. UV absorption and fluorescence maxima of 15h in phosphate solutions</td>
<td>95</td>
</tr>
<tr>
<td>4.6. The structure-activity relationship (SAR) and <em>in vitro</em> cytotoxicity of rosamine analogues in HL60 cells</td>
<td>99</td>
</tr>
<tr>
<td>4.7. Cytotoxic effects of rhodamine analogues on carcinoma and immortalized normal human epithelial cell types</td>
<td>103</td>
</tr>
<tr>
<td>5.1. Synthesis of new pyronin dyes</td>
<td>115</td>
</tr>
<tr>
<td>5.2. Spectral properties of 30 measured in CH$_2$Cl$_2$</td>
<td>117</td>
</tr>
<tr>
<td>6.1. Summary of the yields for preparation of 47, 48, 37 and cassettes 31</td>
<td>140</td>
</tr>
<tr>
<td>TABLE</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.2. Photophysical properties of BODIPY acceptors</td>
<td>141</td>
</tr>
<tr>
<td>6.3. Photophysical properties of BODIPY cassettes</td>
<td>144</td>
</tr>
<tr>
<td>7.1. Optical properties of cassette and its protein conjugates</td>
<td>159</td>
</tr>
</tbody>
</table>
### LIST OF SCHEMES

<table>
<thead>
<tr>
<th>SCHEME</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Chemical synthesis of model chromophore of GFP</td>
<td>29</td>
</tr>
<tr>
<td>2.2. Routes to target compounds 1</td>
<td>33</td>
</tr>
<tr>
<td>2.3. Revised procedure for synthesis of compounds 1f and 5f</td>
<td>42</td>
</tr>
<tr>
<td>2.4. Method to recycle the undesired compounds 5 into the target material 1</td>
<td>44</td>
</tr>
<tr>
<td>2.5. Syntheses of the water-soluble probes 1g and 1h</td>
<td>46</td>
</tr>
<tr>
<td>2.6. Synthesis of water-soluble thiol-reactive analog 1i</td>
<td>47</td>
</tr>
<tr>
<td>3.1. Typical syntheses of symmetric BODIPYs with meso-substituents from (a) acyl chloride; (b) anhydride; (c) aldehyde</td>
<td>57</td>
</tr>
<tr>
<td>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</td>
<td>59</td>
</tr>
<tr>
<td>3.3. (a) Conventional synthesis of tetramethyl-BODIPY 9a; and, (b) the new approach described here</td>
<td>61</td>
</tr>
<tr>
<td>3.4. Synthesis of pyrrole-2-aldehyde 8</td>
<td>62</td>
</tr>
<tr>
<td>4.1. Synthesis of water-soluble rosamine 25</td>
<td>83</td>
</tr>
<tr>
<td>5.1. Synthesis of 3,6-diamino-xanthones 17 and 29</td>
<td>113</td>
</tr>
<tr>
<td>5.2. Failed attempt to synthesize new pyronin dyes</td>
<td>114</td>
</tr>
<tr>
<td>6.1. Synthesis of donor fragment 34</td>
<td>130</td>
</tr>
<tr>
<td>6.2. Synthesis of donor 35</td>
<td>131</td>
</tr>
<tr>
<td>6.3. Syntheses of pyrrole fragments 38-40</td>
<td>133</td>
</tr>
<tr>
<td>6.4. Syntheses of iodo-benzaldehyde fragments 44-46</td>
<td>135</td>
</tr>
<tr>
<td>SCHEME</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>6.5.</td>
<td>Illustrative synthesis of BODIPY acceptor fragment <strong>37a</strong></td>
</tr>
<tr>
<td>6.6.</td>
<td>Syntheses of acceptors <strong>37h-i</strong></td>
</tr>
<tr>
<td>6.7.</td>
<td>Illustrative synthesis of BODIPY cassettes</td>
</tr>
<tr>
<td>6.8.</td>
<td>Synthesis of cassette <strong>32</strong></td>
</tr>
<tr>
<td>7.1.</td>
<td>Illustrative syntheses of symmetric BODIPYs without meso-substituents</td>
</tr>
<tr>
<td>7.2.</td>
<td>Syntheses of rosamine derivatives</td>
</tr>
<tr>
<td>7.3.</td>
<td>New synthesis of rosamine derivatives</td>
</tr>
<tr>
<td>7.4.</td>
<td>Synthesis of water-soluble rosamines</td>
</tr>
<tr>
<td>7.5.</td>
<td>Synthesis of new water-soluble cassette <strong>49</strong></td>
</tr>
<tr>
<td>7.6.</td>
<td>Synthesis of cassette <strong>50</strong></td>
</tr>
<tr>
<td>7.7.</td>
<td>Syntheses of NIR water-soluble BODIPY-BODIPY cassettes <strong>51-52</strong></td>
</tr>
</tbody>
</table>
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. 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).

![Figure 1.1. A typical Jablonski diagram.](image)

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.](image)

A fluorophore can also be excited by absorbing two photons simultaneously (Figure 1.2.). This process is two-photon excitation. 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. The physical origin of the two-photon cross-section can be understood by some simple comparison. 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 $\sigma_1$ 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:

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

where $I$ is the light intensity and $\sigma_1$ is the cross-section for one-photon absorption. The units are given within the parentheses. Similarly, for two-photon absorption:

$$NA_2 \text{ (photon/s)} = \sigma_2 \cdot I^2\text{(photon/cm}^2\text{s})^2$$
In order for the units to match, the units of two-photon cross-section $\sigma_2$ must be cm$^4$s/photon. The values of $\sigma_2$ are usually reported in terms of GM (Gopper-Mayer) units, where 1 GM = 10$^{-50}$ cm$^4$s/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**

<table>
<thead>
<tr>
<th>terms</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molar extinction coefficient ($\varepsilon$: M$^{-1}$cm$^{-1}$)</td>
<td>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</td>
</tr>
<tr>
<td>Stokes shift ($\text{cm}^{-1}$)</td>
<td>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</td>
</tr>
<tr>
<td>Fluorescence quantum yield ($\Phi_f$)</td>
<td>The number of emitted photons relative to the number of absorbed photons; typically determined relative to a fluorophore of known fluorescence quantum yield</td>
</tr>
<tr>
<td>Fwhm: nm</td>
<td>Fluorescence full width at half maximum height (fwhm): a measurement of fluorescence sharpness</td>
</tr>
<tr>
<td>Brightness ($\varepsilon\Phi_f$)</td>
<td>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</td>
</tr>
<tr>
<td>Fluorescence lifetime ($\tau$: ns)</td>
<td>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 $\log I(t)$ versus $t$</td>
</tr>
<tr>
<td>TPA cross-section ($\sigma_2$: GM)</td>
<td>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</td>
</tr>
<tr>
<td>Photobleaching</td>
<td>Photochemical decomposition of fluorophores; complex mechanisms</td>
</tr>
<tr>
<td>Blinking</td>
<td>The fluorophore randomly enter transient, non-fluorescent states; reversible</td>
</tr>
<tr>
<td>Fluorescence anisotropy</td>
<td>Measure for the polarization of the emitted light upon excitation with linearly polarized light; reflects the rotational freedom of fluorophores</td>
</tr>
</tbody>
</table>
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.).

![Diagram](image)

**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.\textsuperscript{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\textsuperscript{6,7}

\[
k_T(r) = \frac{Q_D \kappa^2}{\tau_D \rho^6} \left[ \frac{9000 (\ln 10)}{128 \pi^3 N n^4} \right] J(\lambda)
\]

where \(Q_D\) is the quantum yield of donor in the absence of the acceptor. \(\tau_D\) 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 x 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 \(\lambda\) in the corrected fluorescence spectrum. \(\varepsilon_A(\lambda)\) is the extinction coefficient of the acceptor at wavelength \(\lambda\).

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

\[
\kappa^2 = (\cos \theta_T - 3 \cos \theta_D \cos \theta_A)^2 \quad \text{or} \quad \kappa^2 = (\sin \theta_T \sin \theta_A \cos \phi - 2 \cos \theta_D \cos \theta_A)^2
\]

where \(\theta_T\) is the angle between the emission transition dipole of the donor and the acceptor absorption dipole, \(\theta_D\) and \(\theta_A\) are the angles between these dipoles and the vector joining the donor and the acceptor, and \(\phi\) is the angle between the planes (Figure 1.4.). \(\kappa^2\) 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, \(\kappa^2 = 0\). \(\kappa^2\) 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 $\kappa^2$ 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

$$R_0^6 = \left[ \frac{9000(\ln10)Q_0k^2}{128\pi^5Nn^4} \right] \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda$$

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

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

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

$$E = \frac{k_T}{\tau_D^4 + k_T}$$

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 through 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.](image)

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.\textsuperscript{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 \textit{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.\textsuperscript{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.). 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.

![Figure 1.6.](image)

**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.\textsuperscript{16} Oxidation of N-terminal serine and threonine residues can form reactive aldehydes/ketones allowing further selective derivatization.\textsuperscript{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.\textsuperscript{18} Since endogenous N-terminal cysteine-containing proteins are rare, protein ligation has been used for site-specific labeling of proteins.\textsuperscript{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).\textsuperscript{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.
Figure 1.7. Labeling of proteins by modification of terminal amino acids. Modifications 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 \( S_N^2 \)-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.

**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.). Recognition 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 \(\gamma\)-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.\textsuperscript{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 intracellularly relies on the self-alkylation reaction of human O\textsuperscript{6}-alkylguanine transferase (hAGT or SNAP-tag) to label hAGT fusion proteins with a number of cell permeable, fluorescent O\textsuperscript{6}-benzylguanine (BG) substrates.\textsuperscript{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.\textsuperscript{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 \textit{in vitro} and in living cells.
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.
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. Similar strategy was exploited to label avidin fusion proteins with fluorescent biotin derivatives using the biotin-avidin interactions. More recent approaches have employed the interaction between an FKB12 mutant (F36V) and a synthetic ligand (SLF'). 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.

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.
The membrane-permeant fluorogenic biarsenical (FlAsH) compounds become strongly fluorescent upon high-affinity binding ($K_d : 2-4 \text{ pM}$) 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.\(^{49}\)

![Figure 1.10. Illustrative examples of non-covalent labeling of proteins. a DHFR/methotrexate; b tetracysteine/biarsenical (FlAsH).](image)
Table 1.2. Summary of protein labeling methods

<table>
<thead>
<tr>
<th>labeling method</th>
<th>tag size (aa)</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Covalently labeling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) direct attachment to proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>modification of amino acid side chains (Lys, Cys, Glu, Asp, Trp, Tyr)</td>
<td>0</td>
<td>selective but not specific, only for <em>in vitro</em> labeling</td>
</tr>
<tr>
<td>terminal amino acid modification (Ser, Thr, Cys, Trp)</td>
<td>0</td>
<td>specific but only for <em>in vitro</em> labeling</td>
</tr>
<tr>
<td>unnatural amino acids</td>
<td>0</td>
<td>excellent specificity, not yet broadly applicable in all cell types</td>
</tr>
<tr>
<td>intein mediated protein splicing (temporary)</td>
<td>~150</td>
<td>specific, tag size problem even though temporary</td>
</tr>
<tr>
<td>ligand-directed labeling (GFP)</td>
<td>0</td>
<td>high affinity ligand is needed</td>
</tr>
<tr>
<td>(2) enzyme catalyzed labeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACP (PCP)/PPTase</td>
<td>77 (80)</td>
<td>specific but tags are big and only for cell surface protein labeling</td>
</tr>
<tr>
<td>AP/biotin ligase</td>
<td>12</td>
<td>cell surface protein labeling</td>
</tr>
<tr>
<td>Q-tag/TGase</td>
<td>6</td>
<td>moderate specificity, background labeling</td>
</tr>
<tr>
<td>CAAX/FTase</td>
<td>4</td>
<td>specificity needs improvement</td>
</tr>
<tr>
<td>Sulfatase motif/FGE</td>
<td>6</td>
<td>specificity needs improvement</td>
</tr>
<tr>
<td>BG/hAGT</td>
<td>182</td>
<td>specific and intracellular labeling, but tag is big</td>
</tr>
<tr>
<td>Alkylhalide/HaloTag</td>
<td>296</td>
<td>tag size problem</td>
</tr>
<tr>
<td><strong>Non-covalently labeling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hapten (phOx)/antibody (sFv)</td>
<td>25 KDa</td>
<td>tag size problem</td>
</tr>
<tr>
<td>biotin/avidin</td>
<td>516</td>
<td>tag is big</td>
</tr>
<tr>
<td>SLF’/FKBP12 (FK36V)</td>
<td>108</td>
<td>protein-ligand dissociation, tag size problem</td>
</tr>
<tr>
<td>Methotrexate/DHFR</td>
<td>157</td>
<td>protein-ligand dissociation, tag size problem</td>
</tr>
<tr>
<td>Biarsenicals/tetracystein</td>
<td>4</td>
<td>background labeling, only for reducing environment</td>
</tr>
<tr>
<td>Texas Red/fluorettte</td>
<td>38</td>
<td>Texas Red only</td>
</tr>
<tr>
<td>Ni-NTA/His6</td>
<td>6</td>
<td>low affinity</td>
</tr>
<tr>
<td>Zn labels/polyaspartate</td>
<td>12</td>
<td>low affinity</td>
</tr>
</tbody>
</table>
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. 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. 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), oligo-arginine and polyethyleneimine (PEI). 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 (KETTWETWTEWSQPKKPKKV), 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.\textsuperscript{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 \textit{et al.} postulated that the membrane crossing process involved formation of a transient transmembrane pore-like structure.\textsuperscript{57} In contrast, Henriques \textit{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.\textsuperscript{58} Zerbe \textit{et al.} proposed a model which was compatible both with the active endocytic mechanism and with the passive inverted micelle pathway.\textsuperscript{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.\textsuperscript{52,53}

\subsection*{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.\textsuperscript{60,61}

The two-hybrid system is a method that uses transcriptional activity as a measure of protein-protein interactions.\textsuperscript{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.\textsuperscript{64} Interaction-mediated reconstitution has also been described for ubiquitin, mouse dihydrofolate reductase (DHFR) and so on.\textsuperscript{60}

Immunoprecipitation is another commonly used method to detect PPIs.\textsuperscript{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.\textsuperscript{65} Several approaches, such as mathematical correction, internal calibration, external calibration, donor lifetime measurement and acceptor bleaching, have been employed to quantify 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 (e.g., 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.
What Partner(s) Does Protein 1 Interact With?

FRET donor and cassette donor through-bond energy transfer cassettes

FRET transfer through-bond transfer

multiplexing with the through bond energy transfer cassettes excited by FRET transfer from the donor on protein1, enables multiple protein-protein 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.\(^6^6,6^7\) The phenomenon of green fluorescence from jellyfish has been known for years.\(^6^8\) 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.\(^6^9\) 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.\(^7^0\) Two years later in 1994, Chalfie successfully expressed GFP in *E. coli* so that they fluoresced green when irradiated with blue light.\(^2^4\) 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.\(^7^1-7^3\) 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.

---

The primary sequence of GFP was deduced from the nucleotide sequence of its cDNA by Prasher in 1992. 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 and was further confirmed and corrected by Prasher in 1993. 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. 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 (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.
Figure 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. 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.

![Diagram of GFP biogenesis](image)

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

Chemical synthesis of GFP model chromophore is straightforward. The model compound was prepared via Erlenmeyer azlactone synthesis\(^7^9\) 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.\(^7^8\) 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. 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. 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 ($\lambda_{ex}$ 558 nm, $\lambda_{em}$ 583 nm, $\Phi$ 0.7). 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.](image)

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 and B), 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\(^7\) or naturally occurring analogs (F for red fluorescent protein RFP)\(^8\) 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.\(^8\) 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 \(\beta\)-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

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 all the syntheses of GFP-chromophore analogs are based on condensation of hippuric or aceturic acid derivatives with aldehydes (the “Erlenmeyer azlactone synthesis”. However, at least one report indicates these conditions do not work well for pyrrole-2-carbaldehyde 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, 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, b the strategy developed here involving condensation of imidazolinones.

a

![Scheme 2.2](image-url)
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. Table 2.1. shows
selected spectroscopic parameters for the two isomers of 4c. Coupling between the C⁶-H proton and the C⁵=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 H⁶, H¹', H⁴' and the carbon at C⁵ were noted as possible indicators of E- or Z- stereochemistry.

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

<table>
<thead>
<tr>
<th></th>
<th>²J₆₅ (Hz)</th>
<th>νC=O (cm⁻¹)</th>
<th>δ₆ (ppm)</th>
<th>δ₅ (ppm)</th>
<th>δ¹' (ppm)</th>
<th>δ⁴' (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z-4c</td>
<td>3.8</td>
<td>1686</td>
<td>7.17</td>
<td>169.8</td>
<td>10.62</td>
<td>5.87</td>
</tr>
<tr>
<td>E-4c</td>
<td>8.9</td>
<td>1663</td>
<td>7.33</td>
<td>168.6</td>
<td>13.01</td>
<td>5.99</td>
</tr>
</tbody>
</table>
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\textsuperscript{78} and G\textsuperscript{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\textper{}mol\textsuperscript{-1} via experimental measurements.\textsuperscript{91} Similarly, B3LYP calculations for molecule G indicated that the Z-isomer was thermodynamically preferred by 1.66 kcal\textper{}mol\textsuperscript{-1}.\textsuperscript{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 \textit{in vacuo} and in a medium of the same dielectric constant as DMSO (by 1.33 and 0.61 kcal\textper{}mol\textsuperscript{-1} at 25 °C, and 0.82 and 0.60 kcal\textper{}mol\textsuperscript{-1} at 100 °C, respectively), whereas NMR experiments in \textit{d}\textsubscript{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 \textit{d}\textsubscript{6}-DMSO at 100 °C was quite slow (the reaction approaches equilibrium after 51 h indicative of a very high energy barrier to \textit{E,Z}-isomerization). The \textit{E,Z}-ratio at equilibrium was \textit{ca} 0.6:1.0 indicating the Z-isomer was more stable by 0.41 kcal\textper{}mol\textsuperscript{-1} at 100 °C. Photoisomerization of two different \textit{E:Z} mixtures of 4c in CDCl\textsubscript{3} radiated at 360 nm was also studied, the data (Figure 2.6.b) indicates the \textit{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$_6$-H proton and the 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 $^{13}$C=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

<table>
<thead>
<tr>
<th></th>
<th>$\nu_{\text{C=O}}$ (cm$^{-1}$)</th>
<th>$\delta_{\text{C5}}$ (ppm)</th>
<th>$\delta_{\text{H6}}$ (ppm)</th>
<th>$\delta_{\text{H4'}}$ (ppm)</th>
<th>$\delta_{\text{F}}$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1c</td>
<td>1705</td>
<td>160.7</td>
<td>7.41</td>
<td>6.01</td>
<td>37.02</td>
</tr>
<tr>
<td>5c</td>
<td>1610</td>
<td>158.4</td>
<td>7.64</td>
<td>6.28</td>
<td>34.20</td>
</tr>
</tbody>
</table>

weaker C=O bond
strong $^{19}$F-$^{13}$C coupling
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.\(^96,97\)
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; 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. Finally, the BF2-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.
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 BF2-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 BF2-group to regenerate intermediate 4c. If this
intermediate is not heated above room temperature or radiated with relatively intense and energetic light (e.g., 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.
Review of the literature on BODIPY dyes leads to the conclusion that there are an abundance of lipophilic probes of that general type, but relatively few water-soluble ones. 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 NaHCO₃ (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 CH₂Cl₂/iPrOH). 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 CH₂Cl₂. 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. Again, the sulfonated product could be isolated via column chromatography on silica (20% MeOH in CH₂Cl₂). 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 indicated this was almost exactly 3.0:1.0.
Scheme 2.5. Syntheses of the water-soluble probes 1g and 1h.

1c  
(i) 4 eq. ClSO$_3$H  
CH$_2$Cl$_2$  
0 °C, 10 min  
(ii) NaHCO$_3$(aq.)

1g  
92 %

1h  
71 %

(i) NHS, DIC, DMF  
25 °C, 12 h  
(ii) Streptavidin  
0.1M pH 8.3 NaHCO$_3$  
25 °C, 1 h

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.
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.

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

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{\text{max abs}}$ (nm)</th>
<th>$\varepsilon$ (M$^{-1}$cm$^{-1}$)</th>
<th>$\lambda_{\text{max emiss}}$ (nm)</th>
<th>fwhm$^a$ (nm)</th>
<th>$\Phi^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z-4c$</td>
<td>458</td>
<td>40200</td>
<td>515</td>
<td>69</td>
<td>0.0005</td>
</tr>
<tr>
<td>$1c$</td>
<td>492</td>
<td>57300</td>
<td>532</td>
<td>54</td>
<td>0.86 ± 0.02</td>
</tr>
<tr>
<td>$E-4c$</td>
<td>458</td>
<td>48000</td>
<td>515</td>
<td>73</td>
<td>0.0003</td>
</tr>
<tr>
<td>$5c$</td>
<td>463</td>
<td>42700</td>
<td>530</td>
<td>54</td>
<td>0.01</td>
</tr>
</tbody>
</table>

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

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 lipophilic 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 (λ_{max abs} = 494 ± 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 (λ_{max emis} = 526 ± 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.
Figure 2.10. Continued.
Table 2.4. Spectroscopic properties of GFP-chromophore analogs

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

\text{a} Fluorescein in 0.1M NaOH as standard (Φ=0.92); \text{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.

![Graph of normalized UV absorbance and fluorescence of 1h and the labeled streptavidin 1s in 0.1 M lithium phosphate buffer, pH 7.4.](image)

**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.

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.\textsuperscript{98,104}

8-Substituted BODIPY dyes (\textit{ie} ones with substituents in the \textit{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.\textsuperscript{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 \textit{p}-chloranil or DDQ and eventually the byproducts from this had to be removed (in fact, this was done after complexation with the boron).

\textbf{Scheme 3.1.} Typical syntheses of symmetric BODIPYs with meso-substituent from (a) acyl chloride; (b) anhydride; (c) aldehyde.
Scheme 3.1. Continued.

b

\[
\begin{align*}
\text{N} & \quad \text{BF}_2 \quad \text{N} \\
\text{O} & \quad \text{O} \\
\text{HO}_2C \\
\end{align*}
\]

(i) BF$_3$OEt$_2$ reflux, 5 h
(ii) Et$_3$N BF$_3$OEt$_2$ 25 °C, 12 h

16 %

\[
\begin{align*}
\text{NH} & \quad \text{R}_3 \\
\text{R} & \quad \text{R}_1 \\
\text{R} & \quad \text{R}_2 \\
\text{Ar} & \quad \text{R}_3 \\
\text{R} & \quad \text{R}_1 \\
\text{R} & \quad \text{R}_2 \\
\text{ArCHO} & \quad \text{cat. TFA} \\
\text{DDQ or} & \quad \text{chloranil} \\
\text{base} & \quad \text{BF}_3\text{OEt}_2
\end{align*}
\]

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. 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.
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.\textsuperscript{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.
The syntheses of BODIPYs without *meso*-substituents (e.g. $R^4 = H$ 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.\textsuperscript{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
3.2 A new approach to symmetric BODIPYs

When 3,5-dimethylpyrrole-2-carboxaldehyde 8a in CH₂Cl₂ was treated with POCl₃ as indicated in Scheme 3.3.b, a yellow solution resulted, and evolution of gas was observed. ¹H, ¹³C, ³¹P NMR and MS analyses of the crude reaction mixture showed that it contained the known¹¹⁰-¹¹² dipyrromethenium cation 10 and the dichlorophosphate counterion. This intermediate 10 was converted to the BODIPY product 9a without isolation by adding BF₃•OEt₂ to the reaction mixture.

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

The reaction occurred efficiently with 3,5-dimethylpyrrole-2-carboxaldehyde 8a (0.2 M in CH₂Cl₂) using only 1.2 equivalents of POCl₃ 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.

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 CH₂Cl₂, 1.2 equivalents of POCl₃, room temperature, overnight; then addition of BF₃OEt₂) 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 C₃-
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 CH2Cl2

<table>
<thead>
<tr>
<th></th>
<th>( \lambda_{\text{max abs}} ) (nm)</th>
<th>( \varepsilon ) (M(^{-1})cm(^{-1}))</th>
<th>( \lambda_{\text{max emiss}} ) (nm)</th>
<th>( \text{fwhm}^a ) (nm)</th>
<th>( \Phi_{\text{fl}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9a</td>
<td>505</td>
<td>83200</td>
<td>516</td>
<td>-</td>
<td>0.80(^b)</td>
</tr>
<tr>
<td>9b</td>
<td>629</td>
<td>161500</td>
<td>637</td>
<td>23</td>
<td>0.78 ± 0.03(^c)</td>
</tr>
<tr>
<td>9c</td>
<td>651</td>
<td>149100</td>
<td>662</td>
<td>27</td>
<td>0.71 ± 0.04(^c)</td>
</tr>
<tr>
<td>9d</td>
<td>564</td>
<td>78000</td>
<td>593</td>
<td>-</td>
<td>1.00(^d)</td>
</tr>
<tr>
<td>9e</td>
<td>531</td>
<td>87000</td>
<td>538</td>
<td>21</td>
<td>0.86 ± 0.03(^e)</td>
</tr>
<tr>
<td>9f</td>
<td>512</td>
<td>110200</td>
<td>516</td>
<td>17</td>
<td>0.96 ± 0.02(^e)</td>
</tr>
<tr>
<td>9g</td>
<td>547</td>
<td>92000</td>
<td>553</td>
<td>20</td>
<td>0.90 ± 0.02(^e)</td>
</tr>
</tbody>
</table>

\(^a\) Full width at half maximum height of fluorescence (fwhm); \(^b\) data were obtained from reference\(^{115}\); \(^c\) Nile blue (\(\Phi = 0.27\) in EtOH)\(^{116}\) was used as a standard; \(^d\) data were obtained from reference\(^{117}\); \(^e\) Rhodamine 6G (\(\Phi = 0.94\) 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$_2$Cl$_2$. 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 CDCl₃, at 0.2 M of substrate) was followed by ¹H, ¹³C, and ³¹P 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 x 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•mol⁻¹•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.
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 Vilsmeier-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 possibility is that in the presence of POCl$_3$, pyrrole-2-aldehyde deformylated to generate the free pyrrole which condensed 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.
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.
4.1 Introduction

Rhodamine and rosamine dyes are highly fluorescent xanthylum 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.\textsuperscript{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.\textsuperscript{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 \textit{meso}-substituents are not significantly conjugated to the xanthylum core so that they do not have much effect on

\begin{center}
\textbf{CHAPTER IV}
\end{center}

\begin{center}
ROSAMINES WITH CYCLIC AMINE SUBSTITUENTS: SYNTHESSES, SPECTROSCOPIC PROPERTIES AND APPLICATIONS*
\end{center}

\begin{center}
4.1 Introduction
\end{center}

Rhodamine and rosamine dyes are highly fluorescent xanthylum 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.\textsuperscript{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.\textsuperscript{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 \textit{meso}-substituents are not significantly conjugated to the xanthylum core so that they do not have much effect on

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 \( \text{ca 550 nm} \) and emits \( \text{ca 570 nm} \).\(^{119}\) The additional conjugation of the xanthylium \( \pi \)-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.\(^{122}\)

**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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{122}

Rhodamine derivatives can also be used as substrates for the multidrug resistance (MDR) transporter assays.\textsuperscript{125,126} Multidrug resistance (MDR), mediated by the plasma membrane protein P-glycoprotein\textsuperscript{127} (Pgp), has been identified as a major impediment to successful chemotherapy for a variety of cancers.\textsuperscript{128} Pgp pumps a variety of structurally and mechanistically unrelated chemicals and chemotherapeutic agents from tumor cells, resulting in treatment failures.\textsuperscript{129,130} The mechanism by which Pgp overexpression is induced during exposure to chemotherapeutics is still not fully understood.\textsuperscript{131,132} Attempts have been made to alleviate the function of Pgp using modulators such as verapamil, quinidine, cyclosporin A, PSC833, and others.\textsuperscript{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 xanthylum dyes rhodamine 123 (the methyl ester of rhodamine 110), rhodamine 6G and tetramethylrosamine.\textsuperscript{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.\textsuperscript{125,131} These dyes have also been used as fluorescent markers to determine the efficacy of Pgp modulators.\textsuperscript{139}

Rhodamine 123 has also been extensively used as a mitochondrial stain for studying cellular functions.\textsuperscript{140} Mitochondria are the main energy generators that maintain cell life and essential cell functions.\textsuperscript{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.\textsuperscript{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,\textsuperscript{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 organolithium or -magnesium compound to the carbonyl group of a xanthone, then protonate to cause loss of water and generate the product\textsuperscript{158} (Figure 4.2.b). That approach has been applied extensively by Detty and co-workers to make tetramethylrosamine derivatives (and especially \textit{S}-, \textit{Se}-, and \textit{Te}-analogs)\textsuperscript{159-162} for their photodynamic and other medicinal properties.\textsuperscript{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.\textsuperscript{166} However, the preparation of rosamine derivatives using the addition approach has been limited by the paucity of synthetic routes to 3,6-diamino-substituted-xanthones. The reported approaches to 3,6-diamino-xanthone usually require many step synthesis and only lead to low yield of the desired products.\textsuperscript{166-172} Some routes only work for certain substrates and are not general\textsuperscript{173-178} (Figure 4.3.).

\textbf{Figure 4.2.} Syntheses of rosamine dyes via: (a) the conventional condensation approach; and, (b) nucleophilic additions to 3,6-diamino-xanthones.
Figure 4.3. Reported syntheses of 3,6-diaminesubstituted-xanthes.
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 aminoxanthonones  
(iii) functionalization to prepare water-soluble derivatives  
(iv) photophysical property studies and applications  
(v) biological activity test for the new rosamines
### 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.\textsuperscript{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

<table>
<thead>
<tr>
<th>product</th>
<th>NR₂</th>
<th>yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td></td>
<td>89</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>93</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>
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

![Chemical structure]

(i) $R^3M$, THF
(ii) 2M HCl
$M = Li$ or $MgX$

<table>
<thead>
<tr>
<th>entry</th>
<th>NR$_1$</th>
<th>NR$_2$</th>
<th>$R^3M$</th>
<th>product</th>
<th>yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>CH$_3$MgBr</td>
<td>11a</td>
<td>89</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>PhCH$_2$MgCl</td>
<td>11b</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>PhMgBr</td>
<td>11c</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>I$_2$PhMgCl</td>
<td>11d</td>
<td>90</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>MeO$_2$PhMgBr</td>
<td>11e</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>Ph$_2$MgBr</td>
<td>11f</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>F$_2$MgBr</td>
<td>11g</td>
<td>74</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>Ph$_2$Li</td>
<td>11h</td>
<td>96</td>
</tr>
</tbody>
</table>
Table 4.2. Continued

<table>
<thead>
<tr>
<th>entry</th>
<th>NR$^{1}_2$</th>
<th>NR$^{2}_2$</th>
<th>R$^{3}$M</th>
<th>product</th>
<th>yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>(N)$^{\frac{1}{2}}$</td>
<td>(N)$^{\frac{1}{2}}$</td>
<td>Me-MgCl</td>
<td>11i</td>
<td>99</td>
</tr>
<tr>
<td>10</td>
<td>O-N$^{\frac{1}{2}}$</td>
<td>O-N$^{\frac{1}{2}}$</td>
<td>Me-MgCl</td>
<td>12i</td>
<td>98</td>
</tr>
<tr>
<td>11</td>
<td>N$^{\frac{1}{2}}$</td>
<td>N$^{\frac{1}{2}}$</td>
<td>Me-MgCl</td>
<td>13i</td>
<td>53</td>
</tr>
<tr>
<td>12</td>
<td>N$^{\frac{1}{2}}$</td>
<td>O-N$^{\frac{1}{2}}$</td>
<td>Me-MgCl</td>
<td>22i</td>
<td>98</td>
</tr>
<tr>
<td>13</td>
<td>Boc-N-N$^{\frac{1}{2}}$</td>
<td>Boc-N-N$^{\frac{1}{2}}$</td>
<td>Me-MgCl</td>
<td>14i</td>
<td>90</td>
</tr>
<tr>
<td>14</td>
<td>Boc-N-N$^{\frac{1}{2}}$</td>
<td>Boc-N-N$^{\frac{1}{2}}$</td>
<td>I-MgCl</td>
<td>14d</td>
<td>68</td>
</tr>
<tr>
<td>15</td>
<td>Boc-N-N$^{\frac{1}{2}}$</td>
<td>Boc-N-N$^{\frac{1}{2}}$</td>
<td>OMe-Li</td>
<td>14h</td>
<td>94</td>
</tr>
</tbody>
</table>

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.

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

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).


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.).

\[
\text{[Chemical Structure]}
\]

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.
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 \textit{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).

\begin{equation}
\text{\textit{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).}
\end{equation}
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$_2$Cl$_2$); these values are comparable with a typical rhodamine, rhodamine B (36 nm in CH$_2$Cl$_2$).

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 CH₂Cl₂ to 0.28 in EtOH, while pyrrolidine derivative 13i has very high quantum yields in both solvents (1.00 in CH₂Cl₂ 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-napthalimide¹⁸² 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

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{\text{max , abs}}$ (nm)</th>
<th>$\varepsilon_{\text{max}}$</th>
<th>$\lambda_{\text{fluor}}$ (nm)</th>
<th>fwhm (nm)$^a$</th>
<th>$\Phi_f$ $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH$_2$Cl$_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11a</td>
<td>555</td>
<td>101900</td>
<td>570</td>
<td>35</td>
<td>0.91 ± 0.05</td>
</tr>
<tr>
<td>11b</td>
<td>566</td>
<td>95900</td>
<td>582</td>
<td>34</td>
<td>0.87 ± 0.01</td>
</tr>
<tr>
<td>11c</td>
<td>565</td>
<td>106000</td>
<td>585</td>
<td>37</td>
<td>0.89 ± 0.01</td>
</tr>
<tr>
<td>11d</td>
<td>570</td>
<td>115400</td>
<td>590</td>
<td>38</td>
<td>0.85 ± 0.01</td>
</tr>
<tr>
<td>11e</td>
<td>562</td>
<td>116700</td>
<td>582</td>
<td>35</td>
<td>0.91 ± 0.02</td>
</tr>
<tr>
<td>11f</td>
<td>566</td>
<td>118000</td>
<td>585</td>
<td>37</td>
<td>0.92 ± 0.03</td>
</tr>
<tr>
<td>11g</td>
<td>582</td>
<td>107600</td>
<td>603</td>
<td>41</td>
<td>0.33 ± 0.01$^c$</td>
</tr>
<tr>
<td>11h</td>
<td>565</td>
<td>87300</td>
<td>586</td>
<td>36</td>
<td>0.86 ± 0.04</td>
</tr>
<tr>
<td>11i</td>
<td>565</td>
<td>123600</td>
<td>582</td>
<td>34</td>
<td>0.89 ± 0.01</td>
</tr>
<tr>
<td>12i</td>
<td>554</td>
<td>106900</td>
<td>575</td>
<td>37</td>
<td>0.80 ± 0.01</td>
</tr>
<tr>
<td>13i</td>
<td>558</td>
<td>127400</td>
<td>573</td>
<td>31</td>
<td>1.00 ± 0.09</td>
</tr>
<tr>
<td>22i</td>
<td>559</td>
<td>111000</td>
<td>579</td>
<td>36</td>
<td>0.59 ± 0.01</td>
</tr>
<tr>
<td>14i</td>
<td>555</td>
<td>120800</td>
<td>577</td>
<td>40</td>
<td>0.47 ± 0.01</td>
</tr>
<tr>
<td>14d</td>
<td>561</td>
<td>77900</td>
<td>585</td>
<td>42</td>
<td>0.41 ± 0.01</td>
</tr>
<tr>
<td>14h</td>
<td>555</td>
<td>109300</td>
<td>580</td>
<td>40</td>
<td>0.49 ± 0.02</td>
</tr>
<tr>
<td>23</td>
<td>564</td>
<td>62800</td>
<td>586</td>
<td>44</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EtOH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11c</td>
<td>564</td>
<td>86700</td>
<td>588</td>
<td>39</td>
<td>0.30 ± 0.01</td>
</tr>
<tr>
<td>11d</td>
<td>568</td>
<td>70700</td>
<td>592</td>
<td>40</td>
<td>0.28 ± 0.01</td>
</tr>
<tr>
<td>11e</td>
<td>561</td>
<td>100400</td>
<td>582</td>
<td>37</td>
<td>0.31 ± 0.01</td>
</tr>
<tr>
<td>11f</td>
<td>564</td>
<td>115600</td>
<td>586</td>
<td>37</td>
<td>0.36 ± 0.01</td>
</tr>
<tr>
<td>11g</td>
<td>581</td>
<td>56500</td>
<td>606</td>
<td>43</td>
<td>0.10 ± 0.01$^c$</td>
</tr>
<tr>
<td>11i</td>
<td>564</td>
<td>119500</td>
<td>584</td>
<td>36</td>
<td>0.28 ± 0.01</td>
</tr>
<tr>
<td>12i</td>
<td>556</td>
<td>98600</td>
<td>582</td>
<td>38</td>
<td>0.16 ± 0.01</td>
</tr>
</tbody>
</table>
### Table 4.3. Continued

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{\text{max abs}}$ (nm)</th>
<th>$\varepsilon_{\text{max}}$</th>
<th>$\lambda_{\text{fluor}}$ (nm)</th>
<th>fwhm (nm)$^a$</th>
<th>$\Phi_f^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EtOH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13i</td>
<td>558</td>
<td>120500</td>
<td>576</td>
<td>33</td>
<td>0.95 ± 0.01</td>
</tr>
<tr>
<td>22i</td>
<td>560</td>
<td>97000</td>
<td>582</td>
<td>38</td>
<td>0.14 ± 0.01</td>
</tr>
<tr>
<td>14i</td>
<td>556</td>
<td>108000</td>
<td>580</td>
<td>39</td>
<td>0.19 ± 0.01</td>
</tr>
<tr>
<td>24</td>
<td>557</td>
<td>71000</td>
<td>580</td>
<td>37</td>
<td>0.26 ± 0.01</td>
</tr>
<tr>
<td>25</td>
<td>557</td>
<td>84800</td>
<td>580</td>
<td>38</td>
<td>0.49 ± 0.02</td>
</tr>
<tr>
<td>15h</td>
<td>561</td>
<td>81000</td>
<td>582</td>
<td>45</td>
<td>0.008</td>
</tr>
<tr>
<td>26</td>
<td>562</td>
<td>82400</td>
<td>589</td>
<td>37</td>
<td>0.28 ± 0.01</td>
</tr>
<tr>
<td>27</td>
<td>553</td>
<td>117200</td>
<td>592</td>
<td>38</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0.1 M Phosphate buffer, pH = 7.4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15h</td>
<td>542</td>
<td>71500</td>
<td>568</td>
<td>41</td>
<td>0.14 ± 0.02</td>
</tr>
<tr>
<td>25</td>
<td>567</td>
<td>76500</td>
<td>594</td>
<td>42</td>
<td>0.13 ± 0.01</td>
</tr>
<tr>
<td>26</td>
<td>564</td>
<td>73600</td>
<td>602</td>
<td>44</td>
<td>0.009</td>
</tr>
</tbody>
</table>

$^a$ 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 ($\Phi = 1.0$ in EtOH)$^{187}$; $^c$ Cresyl Violet ($\Phi = 0.54$ 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.
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 CH₂Cl₂.

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$_2$Cl$_2$ 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

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Dielectric Constant</th>
<th>Dipole Moment</th>
<th>$\lambda_{\text{max abs}}$ (nm)</th>
<th>$\lambda_{\text{fluor}}$ (nm)</th>
<th>$\Phi_f$\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHCl$_3$</td>
<td>4.8</td>
<td>1.04</td>
<td>566</td>
<td>580</td>
<td>0.95 ± 0.02</td>
</tr>
<tr>
<td>EtOAc</td>
<td>6.1</td>
<td>1.78</td>
<td>567</td>
<td>586</td>
<td>0.19 ± 0.01</td>
</tr>
<tr>
<td>THF</td>
<td>7.5</td>
<td>1.75</td>
<td>567</td>
<td>588</td>
<td>0.32 ± 0.01</td>
</tr>
<tr>
<td>CH$_2$Cl$_2$</td>
<td>8.9</td>
<td>1.60</td>
<td>565</td>
<td>582</td>
<td>0.89 ± 0.01</td>
</tr>
<tr>
<td>Acetone</td>
<td>21.0</td>
<td>2.88</td>
<td>564</td>
<td>587</td>
<td>0.15 ± 0.01</td>
</tr>
<tr>
<td>EtOH</td>
<td>24.3</td>
<td>1.69</td>
<td>564</td>
<td>584</td>
<td>0.28 ± 0.01</td>
</tr>
<tr>
<td>MeOH</td>
<td>33.0</td>
<td>1.70</td>
<td>562</td>
<td>582</td>
<td>0.19 ± 0.01</td>
</tr>
<tr>
<td>CH$_3$CN</td>
<td>36.6</td>
<td>3.92</td>
<td>562</td>
<td>586</td>
<td>0.14 ± 0.01</td>
</tr>
<tr>
<td>DMF</td>
<td>38.3</td>
<td>3.82</td>
<td>570</td>
<td>594</td>
<td>0.20 ± 0.01</td>
</tr>
</tbody>
</table>

\textsuperscript{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](image_url)  
**Figure 4.6.** Fluorescence intensity (excited at 562 nm) of 11i at 1.4 x 10^{-7} M in a CH$_2$Cl$_2$/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.

![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).](image)

**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-butyloxy carbonyl 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. 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>
<thead>
<tr>
<th>pH</th>
<th>$\lambda_{\text{max abs}}$ (nm)</th>
<th>$\lambda_{\text{max emiss}}$ (nm)</th>
<th>pH</th>
<th>$\lambda_{\text{max abs}}$ (nm)</th>
<th>$\lambda_{\text{max emiss}}$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>538</td>
<td>565</td>
<td>8.12</td>
<td>551</td>
<td>571</td>
</tr>
<tr>
<td>3.00</td>
<td>538</td>
<td>565</td>
<td>8.43</td>
<td>555</td>
<td>574</td>
</tr>
<tr>
<td>4.08</td>
<td>538</td>
<td>565</td>
<td>9.09</td>
<td>559</td>
<td>582</td>
</tr>
<tr>
<td>5.18</td>
<td>539</td>
<td>565</td>
<td>10.22</td>
<td>560</td>
<td>584</td>
</tr>
<tr>
<td>6.04</td>
<td>539</td>
<td>565</td>
<td>10.67</td>
<td>560</td>
<td>584</td>
</tr>
<tr>
<td>7.08</td>
<td>541</td>
<td>567</td>
<td>11.13</td>
<td>560</td>
<td>584</td>
</tr>
<tr>
<td>7.55</td>
<td>544</td>
<td>568</td>
<td>11.98</td>
<td>560</td>
<td>584</td>
</tr>
</tbody>
</table>

Table 4.5. UV absorption and fluorescence maxima of 15h in phosphate solutions
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; excited 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$_3$PO$_4$+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 $\lambda_{\text{max abs}}$ 560 nm 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).
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 NaHCO₃ 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](image-url)  
**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. Results were expressed as IC$_{50}$: the concentration of compound (in $\mu$M) 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 $\mu$M 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 $\mu$M.

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

<table>
<thead>
<tr>
<th>Compd</th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$R_3$</th>
<th>IC$_{50}$ ($\mu$M) ± S.D.$^a$ in HL60</th>
</tr>
</thead>
<tbody>
<tr>
<td>13i</td>
<td>(\text{N})</td>
<td>(\text{N})</td>
<td>(\text{Cl}^-)</td>
<td>0.72 ± 0.09</td>
</tr>
<tr>
<td>11i</td>
<td>(\text{N})</td>
<td>(\text{N})</td>
<td></td>
<td>0.76 ± 0.03</td>
</tr>
<tr>
<td>22i</td>
<td>(\text{N})</td>
<td>(\text{N})</td>
<td>(\text{O} )</td>
<td>0.35 ± 0.01</td>
</tr>
<tr>
<td>Compd</td>
<td>R₁</td>
<td>R₂</td>
<td>R₃</td>
<td>IC₅₀ (μM) ± S.D.⁴ in HL60</td>
</tr>
<tr>
<td>-------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>--------------------------</td>
</tr>
<tr>
<td>12i</td>
<td></td>
<td></td>
<td></td>
<td>8.27 ± 2.16</td>
</tr>
<tr>
<td>14i</td>
<td></td>
<td></td>
<td></td>
<td>0.62 ± 0.07</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td>53.3 ± 0.30</td>
</tr>
<tr>
<td>11a</td>
<td></td>
<td></td>
<td>CH₃</td>
<td>3.86 ± 1.46</td>
</tr>
<tr>
<td>11b</td>
<td></td>
<td></td>
<td>Ph</td>
<td>0.82 ± 0.04</td>
</tr>
<tr>
<td>11c</td>
<td></td>
<td></td>
<td></td>
<td>0.47 ± 0.13</td>
</tr>
<tr>
<td>11g</td>
<td></td>
<td></td>
<td></td>
<td>0.10 ± 0.04</td>
</tr>
<tr>
<td>11d</td>
<td></td>
<td></td>
<td></td>
<td>0.09 ± 0.01</td>
</tr>
<tr>
<td>11e</td>
<td></td>
<td></td>
<td></td>
<td>0.66 ± 0.17</td>
</tr>
<tr>
<td>11f</td>
<td></td>
<td></td>
<td></td>
<td>0.95 ± 0.01</td>
</tr>
<tr>
<td>11h</td>
<td></td>
<td></td>
<td></td>
<td>0.25 ± 0.22</td>
</tr>
</tbody>
</table>
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 IC₅₀ 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 IC₅₀ 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 IC₅₀ 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 IC₅₀ values than the unsubstituted meso-aryl 11c. Having an

<table>
<thead>
<tr>
<th>Compd</th>
<th>R₁</th>
<th>R₂</th>
<th>R₃</th>
<th>IC₅₀ (μM) ± S.D.ᵃ in HL60</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>( \text{HN} )</td>
<td>( \text{HN} )</td>
<td>( \text{HO} )</td>
<td>2.91 ± 1.86</td>
</tr>
<tr>
<td>25</td>
<td>( \text{HN} )</td>
<td>( \text{HN} )</td>
<td>( \text{HOCH₂CO} )</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

ᵃMean IC₅₀ and standard deviation of triplicate determination assessed *in vitro* at 24 h post-treatment using MTT assay.
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 IC₅₀ 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 IC50 values compared to Rh123. In addition, the thiofuran-substituted compound (11g) consistently showed between 1.5-fold to 4-fold lower IC50 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

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Tissue Origin</th>
<th>IC50 (μM) ± S.D.(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11d</td>
<td>11g</td>
</tr>
<tr>
<td>HCT116</td>
<td>Colon</td>
<td>0.39 ± 0.11</td>
</tr>
<tr>
<td>MCF-7</td>
<td>Breast</td>
<td>0.39 ± 0.22</td>
</tr>
<tr>
<td>HSC2</td>
<td>Oral</td>
<td>0.25 ± 0.12</td>
</tr>
<tr>
<td>OKF6(^b)</td>
<td>Oral</td>
<td>0.41 ± 0.07</td>
</tr>
<tr>
<td>HK1</td>
<td>Nasopharyngeal</td>
<td>0.42 ± 0.06</td>
</tr>
<tr>
<td>NP69(^b)</td>
<td>Nasopharyngeal</td>
<td>0.51 ± 0.16</td>
</tr>
</tbody>
</table>

\(^a\)Mean IC50 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\textsubscript{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\textsubscript{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.
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 $\mu$M) 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 $\mu$M compound 11d at different duration.](image-url)
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 IC₅₀ 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.

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). The color of their acidic and neutral solution is red (xanthylum 3,6-bis(dimethylamino) chloride PY: $\lambda_{\text{max}}^{\text{abs}} = 547 \text{ nm}$; $\lambda_{\text{max}}^{\text{emiss}} = 566 \text{ nm}$), while their alkaline solutions are colorless (PY: $\lambda_{\text{max}}^{\text{abs}} = 390 \text{ nm}$; $\lambda_{\text{max}}^{\text{emiss}} = 463 \text{ nm}$); 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.

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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{[210]} The interactions of pyronin Y with nucleic acids are complex process.\textsuperscript{[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.\textsuperscript{[208]} In the presence of these dyes, PY can be used as a specific RNA fluorophore.\textsuperscript{[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.\textsuperscript{[210]}

Pyronins have also been used as laser dyes and to detect trace proteins and determine species such as As(III)\textsuperscript{[212]}, Cr(VI)\textsuperscript{[213]} and tannins.\textsuperscript{[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\textsubscript{2} or SR. In fact, to the best of our knowledge, the only reported compound in this category is “Pyronin 12”, \textit{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,\textsuperscript{[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.

\[ \text{A} \quad \quad \text{B} \]

\[ X = \text{aryl or alkyl; well known} \]

\[ X = \text{NR}_2 \text{ or SR; unexplored} \]

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.
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$_2$Cl$_2$ 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

<table>
<thead>
<tr>
<th>entry</th>
<th>Y</th>
<th>NuH</th>
<th>yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30a</td>
<td>H</td>
<td>(\text{NH} )</td>
<td>81</td>
</tr>
<tr>
<td>30b</td>
<td>H</td>
<td>(\text{Et}_2\text{NH} )</td>
<td>61</td>
</tr>
<tr>
<td>30c</td>
<td>H</td>
<td>(\text{H}_2\text{N} )</td>
<td>98</td>
</tr>
<tr>
<td>30d</td>
<td>H</td>
<td>(\text{H}_2\text{N}\text{CO}_2\text{Bu} )</td>
<td>80</td>
</tr>
<tr>
<td>30e</td>
<td>H</td>
<td>(\text{BnSH} )</td>
<td>23</td>
</tr>
<tr>
<td>30f</td>
<td>MeO\text{C}</td>
<td>(\text{NH} )</td>
<td>86</td>
</tr>
</tbody>
</table>
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$_2$Cl$_2$ 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>
<thead>
<tr>
<th></th>
<th>$\lambda_{\text{max}}$ (nm)</th>
<th>$\varepsilon_{\text{max}}$ (L·mol$^{-1}$·cm$^{-1}$)</th>
<th>$\lambda_f$ (nm)</th>
<th>fwhm (nm)$^a$</th>
<th>$\Phi_f$ $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30a</td>
<td>491</td>
<td>60400</td>
<td>562</td>
<td>55</td>
<td>0.14 ± 0.01</td>
</tr>
<tr>
<td>30b</td>
<td>499</td>
<td>51000</td>
<td>558</td>
<td>41</td>
<td>0.11 ± 0.02</td>
</tr>
<tr>
<td>30c</td>
<td>456</td>
<td>52400</td>
<td>537</td>
<td>52</td>
<td>0.85 ± 0.01</td>
</tr>
<tr>
<td>30d</td>
<td>462</td>
<td>62000</td>
<td>544</td>
<td>52</td>
<td>0.90 ± 0.02</td>
</tr>
<tr>
<td>30e</td>
<td>598</td>
<td>82600</td>
<td>619</td>
<td>38</td>
<td>0.79 ± 0.02</td>
</tr>
<tr>
<td>30f</td>
<td>482</td>
<td>54900</td>
<td>552</td>
<td>47</td>
<td>0.02 ± 0.01</td>
</tr>
</tbody>
</table>

$^a$ 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$_2$Cl$_2$ (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 NEt₃ (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 NEt₃ 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 NEt₃ caused decomposition (see Appendix Information).

![Graph](image.png)

**Figure 5.3.** Dependence of electronic spectra of 30c on pH. UV absorbance for 30c under: a acidic; and, b basic conditions (CH₂Cl₂, 1.1 x 10⁻⁵ M). Fluorescence pH dependence of 30c under: c acidic; and, d basic conditions (CH₂Cl₂, 2.5 x 10⁻⁷ M, excited at 456 nm). e Fluorescence intensity of 30c under alternating basic (Et₃N) and acidic (TFA) conditions (CH₂Cl₂, 2.5 x 10⁻⁷ M, excited at 456 nm).
Figure 5.3. Continued.
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.).

![Figure 5.4. Likely acid/base equilibrium for 30c.](image)

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\textsuperscript{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.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.1.png}
\caption{(a) Through-space and (b) through-bond energy transfer cassettes.}
\end{figure}

Our research focuses on energy transfer cassettes in which the donor is connected to acceptors via electronically conjugated linkers (Figure 6.1.b)\textsuperscript{8,9,218-220} and energy transfer may occur via through-bond pathways such as Dexter energy transfer\textsuperscript{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.

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. 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. 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.

![Diagram of energy transfer cassettes]

**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. 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. 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.](image)

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
31a $R^1 = \text{OMe}, R^2 = \text{H}, n = 0$;
31b $R^1 = \text{O(CH}_2\text{CH}_2\text{O)}_3\text{Me}, R^2 = \text{H}, n = 6$;
31c $R^1 = \text{H}, R^2 = \text{O(CH}_2\text{CH}_2\text{O)}_3\text{Me}, n=6$

31e $R = \text{H}, n = 0$;
31f $R = \text{OMe}, n = 0$;
31g $R = \text{O(CH}_2\text{CH}_2\text{O)}_6\text{Me}, n = 6$

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.

![Donor fragments for water-soluble cassettes](image)

*Figure 6.6. Donor fragments for water-soluble cassettes.*

The synthesis and properties of compound 1g were described in Chapter II.\textsuperscript{223} This compound is soluble in water and highly fluorescent. BODIPY 34 was prepared and studied by a previous group member.\textsuperscript{224} The synthesis of this compound is summarized in Scheme 6.1.
Scheme 6.1. Synthesis of donor fragment 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.

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. 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.
Figure 6.7. BODIPY acceptor fragments for cassettes.
Scheme 6.3. Syntheses of pyrrole fragments 38-40.

a

\[
\text{N} \quad \text{Boc} \quad \text{Br} \quad \xrightarrow{\text{ArB(OH)}_2, \text{Pd(PPh}_3)_4, \text{aq. Na}_2\text{CO}_3} \quad \text{N} \quad \text{Boc} \quad \text{R}_1 \quad \text{R}_2 \quad \xrightarrow{\text{NaOMe, MeOH}, \text{THF, 25 }^\circ\text{C, 24 h}} \quad \text{N} \quad \text{Boc} \quad \text{R}_1 \quad \text{R}_2
\]

5:1 PhCH\text{Me}/MeOH 80 °C, 12 h

41

\text{a 99 %; b 94 %; c 94 %}

\[
\text{N} \quad \text{Boc} \quad \text{R}_1 \quad \text{R}_2 \quad \xrightarrow{\text{NaOMe, MeOH}} \quad \text{N} \quad \text{Boc} \quad \text{R}_1 \quad \text{R}_2
\]

\text{a: R}_1 = \text{OMe}, R_2 = \text{H}
\text{b: R}_1 = \text{O(\text{CH}_2\text{CH}_2\text{O})}_3\text{Me}, R_2 = \text{H}
\text{c: R}_1 = \text{H}, R_2 = \text{O(\text{CH}_2\text{CH}_2\text{O})}_3\text{Me}

38

\text{a 92 %; b 87 %; c 82 %}

b

\[
\text{R} \quad \text{N} \quad \text{OH} \quad \xrightarrow{\text{acetylene}} \quad \text{R} \quad \text{N} \quad \text{HN}
\]

LiOH, DMSO 140 °C, 5 h

39

\text{a 71 %; b 65 %; c 42 %}

c

\[
\text{Br} \quad \text{O} \quad \text{CHO} \quad \xrightarrow{(i) \text{ethyl azidoacetate}} \quad \text{Br} \quad \text{O} \quad \text{CO}_2\text{Et}
\]

NaOEt, EtOH

\text{(ii) PhCH}_3, \text{reflux, 30 min}

\text{(iii) Et}_3\text{N, Boc}_2\text{O, DMAP}

CH\text{C}_2\text{Cl}_2, 25 °C, 12 h

42 73 %
Scheme 6.3. Continued.

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. Compounds 38b-c were prepared using similar conditions reported for 38a (Scheme 6.3.a). Pyrrole 39c was synthesized according to a known procedure developed for 39a-b (Scheme 6.3.b). Analogs of pyrrole 40a-b have also been reported. 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 and then protected by tert-butyloxycarbonyl (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

\[
\begin{align*}
\text{I} & \quad \text{CHO} \\
\text{O} & \quad \text{CO}_2\text{Me} \\
\text{CHO} & \quad \text{O} \quad \text{CO}_2\text{Me} \\
\text{K}_2\text{CO}_3, \text{CH}_3\text{CN}, \text{reflux} & \\
\end{align*}
\]

44 \( X = \text{Br}, n = 0, 88 \% \)

45 \( X = \text{OTs}, n = 6, 92 \% \)

b

\[
\begin{align*}
\text{I} & \quad \text{CHO} \\
\text{Br} & \quad \text{CO}_2\text{Me} \\
\text{CHO} & \quad \text{CO}_2\text{Me} \\
\text{K}_2\text{CO}_3, \text{CH}_3\text{CN} & \quad \text{reflux, 30 min} \\
\end{align*}
\]

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\(_2\)O) 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.

\[
\begin{align*}
44 & \quad + \quad 38a \\
\text{(i) TFA, CH}_2\text{Cl}_2 & \quad 25 \text{ °C, 1 h} \\
\text{(ii) chloranil} & \\
\text{(iii) Et}_3\text{N, BF}_3\text{OEt}_2 & \quad \text{ClCH}_2\text{CH}_2\text{Cl} \\
& \quad \text{reflux, 1 h}
\end{align*}
\]

47a 69%

\[
\begin{align*}
\text{TMS} & \quad \text{PdCl}_2(\text{PPh}_3)_2 \\
\text{Cul, Et}_3\text{N, CH}_2\text{Cl}_2 & \quad 25 \text{ °C, 24 h}
\end{align*}
\]

48a 92%

\[
\begin{align*}
\text{K}_2\text{CO}_3 & \\
\text{MeOH/THF/H}_2\text{O} & \quad 25 \text{ °C, 24 h}
\end{align*}
\]

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 K₂CO₃ in MeOH/THF/H₂O (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.
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 synthesis 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.
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.
Table 6.1. summarizes the yields for the syntheses of cassettes 31, corresponding BODIPY acceptors 37 and the intermediates 47-48.

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

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

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

$^a$ Cresyl Violet ($\Phi=0.54$ in MeOH); $^b$ Nile Blue ($\Phi=0.27$ 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 strong absorption band around 500 nm corresponds to the $\lambda_{\text{max}}$ 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.

![Absorption spectra](image1)

**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 cassettes 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

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{\text{abs D}}$ (nm)</th>
<th>$\lambda_{\text{abs A}}$ (nm)</th>
<th>$\lambda_{\text{emiss A}}$ (nm)</th>
<th>log($\varepsilon_{\text{max}}$)</th>
<th>$\Phi$ of acceptor</th>
<th>E.T.E (%)$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_{\text{abs D}}$</td>
<td>$\lambda_{\text{abs A}}$</td>
<td>$\lambda_{\text{emiss A}}$</td>
<td>$\log(\varepsilon_{\text{max}})$</td>
<td>$\Phi$ of acceptor</td>
<td>E.T.E (%)$^d$</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>A</td>
<td>Ex. D$^a$</td>
<td>Ex. A</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>31a</td>
<td>506</td>
<td>542</td>
<td>595</td>
<td>4.96</td>
<td>4.52</td>
<td>0.50</td>
</tr>
<tr>
<td>31b</td>
<td>506</td>
<td>552</td>
<td>601</td>
<td>5.15</td>
<td>4.74</td>
<td>0.46</td>
</tr>
<tr>
<td>31c</td>
<td>506</td>
<td>584</td>
<td>623</td>
<td>5.02</td>
<td>4.73</td>
<td>0.69</td>
</tr>
<tr>
<td>31d</td>
<td>506</td>
<td>542</td>
<td>594</td>
<td>5.02</td>
<td>4.58</td>
<td>0.77</td>
</tr>
<tr>
<td>31e</td>
<td>506</td>
<td>626</td>
<td>640</td>
<td>4.67</td>
<td>4.72</td>
<td>0.69</td>
</tr>
<tr>
<td>31f</td>
<td>506</td>
<td>649</td>
<td>666</td>
<td>4.81</td>
<td>4.86</td>
<td>0.47</td>
</tr>
<tr>
<td>31g</td>
<td>506</td>
<td>652</td>
<td>668</td>
<td>4.90</td>
<td>4.95</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>0.1M PBS buffer, pH 7.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31a</td>
<td>500</td>
<td>535</td>
<td>594</td>
<td>4.94</td>
<td>4.41</td>
<td>0.25</td>
</tr>
<tr>
<td>31b</td>
<td>501</td>
<td>556</td>
<td>599</td>
<td>5.06</td>
<td>4.58</td>
<td>0.30</td>
</tr>
<tr>
<td>31c</td>
<td>503</td>
<td>589</td>
<td>628</td>
<td>4.94</td>
<td>4.53</td>
<td>0.030</td>
</tr>
<tr>
<td>31d</td>
<td>500</td>
<td>535</td>
<td>592</td>
<td>4.95</td>
<td>4.42</td>
<td>0.069</td>
</tr>
<tr>
<td>31e</td>
<td>499</td>
<td>631</td>
<td>645</td>
<td>4.52</td>
<td>4.39</td>
<td>0.24</td>
</tr>
<tr>
<td>31f</td>
<td>500</td>
<td>655</td>
<td>674</td>
<td>4.65</td>
<td>4.48</td>
<td>0.032</td>
</tr>
<tr>
<td>31g</td>
<td>503</td>
<td>654</td>
<td>673</td>
<td>4.82</td>
<td>4.63</td>
<td>0.008</td>
</tr>
</tbody>
</table>

$^a$ Rhodamine 6G ($\Phi=0.94$ in EtOH); $^b$ Cresyl Violet ($\Phi=0.54$ in MeOH); $^c$ Nile Blue ($\Phi = 0.27$ 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 interactions simultaneously (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.

![Figure 6.10](image1)

**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).
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. The virtually 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 as 5.2:6.6:1.0. Cassette 1 was activated (di-isopropylcarbodiimide, 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 (λ_{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-streptavidin 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 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-cassette1; (c) avidin-cassette2; (d) biotin-BSA-Atto425 + streptavidin-cassette1; (e) biotin-BSA-Atto425 + avidin-cassette2; (f) biotin-BSA-Atto425 + streptavidin-cassette1 + avidin-cassette2. Throughout, excited at 458 nm.
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.

Figure 6.14. Continued.
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.

a

\[
\begin{align*}
\text{NH} & \quad \text{OHC} \\
& \quad \text{HN} \\
\text{N} & \quad \text{B} & \quad \text{F}_2 \\
& \quad \text{N}
\end{align*}
\]

(i) H\(^+\) or LA

(ii) Et\(_3\)N, BF\(_3\)OEt\(_2\)

70 %

b

\[
\begin{align*}
\text{CHO} & \quad \text{NH} \\
& \quad \text{N} & \quad \text{B} & \quad \text{F}_2 \\
& \quad \text{N}
\end{align*}
\]

(i) POCl\(_3, \text{CH}_2\text{Cl}_2, 25^\circ\text{C, 12 h}

(ii) BF\(_3\)OEt\(_2, \text{Et}_3\)N

25 \degree C, 12 h

92 %

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.

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.

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.

![Figure 7.3. Illustrative examples of water-soluble BODIPY acceptor and through-bond energy transfer cassette.](image)

**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
The quantum yield of the cassette was reduced upon attached to proteins. The absorption and emission wavelength of these conjugates were very similar to those of the cassette (Table 7.1.). The fluorescent conjugates were then introduced back into cells such as COS-7 cells by cell penetrating peptide (e.g. pep-1, Arg₈). These conjugates were easily detected by fluorescence microscope. The intracellular localization was also studied.

<table>
<thead>
<tr>
<th>Table 7.1. Optical properties of cassette and its protein conjugates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cassette 31b</td>
</tr>
<tr>
<td>BSA-cassette</td>
</tr>
<tr>
<td>β-Galactosidase-cassette</td>
</tr>
<tr>
<td>Avidin-cassette</td>
</tr>
<tr>
<td>Streptavidin-cassette</td>
</tr>
</tbody>
</table>

\(^a\) Rhodamine 6G (\(\Phi=0.94\) in EtOH); \(^b\) Cresyl Violet (\(\Phi=0.54\) 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.

![Figure 7.4. Scheme for detecting protein-protein interactions.](image)
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 by 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.

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.
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 IC₅₀ 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.](image-url)
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 \textit{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.

\textbf{Figure 7.6.} Water-soluble through-bond energy transfer cassettes.
Scheme 7.5. Synthesis of new water-soluble cassette 49.

1. Reaction with TsO
2. K2CO3, CH3CN reflux
3. (i) NH2OH
4. (ii) LiOH, acetylene DMSO
5. (iii) BF3OEt2, Et3N
6. (v) K2CO3
7. (i) TFA; (ii) DDQ
8. (iii) TMS-acetylene, "Pd"
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.

(i) BBr$_3$; (ii) triflation
(iii) Sonogashira
(iv) deprotection

donor-N$_3$

"Click"

Em: 690-700 nm

D = NaO$_3$S

51
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


(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.


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


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


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


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$_3$ ($\delta$ 7.24 ppm $^1$H; $\delta$ 77.0 ppm $^{13}$C) CD$_3$OD ($\delta$ 3.31 ppm $^1$H; $\delta$ 49.0 ppm $^{13}$C) or $d^6$-DMSO ($\delta$ 2.49 ppm $^1$H; $\delta$ 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 = ε • C • 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:

\[ \Phi_x = \Phi_{st} \left( \frac{I_x}{I_{st}} \right) \left( \frac{A_{st}}{A_x} \right) \left( \frac{\eta_x^2}{\eta_{st}^2} \right) \]

where \( \Phi_{st} \) 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 \( \eta \) 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

\[
\text{N} \quad \text{O} \\
\text{Cl} \quad \text{O} \\
\text{Et}_3\text{N, CH}_2\text{Cl}_2 \\
25 \degree C, 2 \text{ h} \\
\text{N} \quad \text{O} \\
\text{O} \\
\text{N}_3 \\
2a
\]

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\(_3\)N (15.2 mL, 0.11 mol) in 100 mL CH\(_2\)Cl\(_2\). The solution was stirred at 25\(^\circ\)C for 2h, then washed with water (3 x 30 mL) and brine (1 x 30 mL). The organic layer was dried over Na\(_2\)SO\(_4\) 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\(_2\)O (3 x 30 mL) and brine (1 x 30 mL). The organic layer was dried over Na\(_2\)SO\(_4\) 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. \(^1\)H NMR (300 MHz, CDCl\(_3\)) \(\delta\) 4.34 (s, 2H), 3.23 (s, 3H), 2.31 (s, 3H); \(^13\)C NMR (75 MHz, CDCl\(_3\)) \(\delta\) 173.1, 171.1, 55.5, 31.9, 25.4. MS (ESI) m/z calcd for (M+H)\(^+\) C\(_5\)H\(_9\)N\(_4\)O\(_2\) 157.07; found 157.07.
\[ \text{NMR (CDCl}_3\text{)} \]

\[ \text{1H NMR (CDCl}_3\text{)} \]

\[ \text{13C NMR (CDCl}_3\text{)} \]
General procedure for the synthesis of α-azido substituted imides 2b-d\textsuperscript{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\textsubscript{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\textsubscript{2}O (3 x 20 mL) and brine (1 x 20 mL). The organic layer was dried over Na\textsubscript{2}SO\textsubscript{4} and concentrated under reduced pressure. The crude product was purified by flash chromatography to afford 2.

\[ \text{PhN} \quad \text{O} \quad \text{O} \quad \text{N}_3 \]

\[ 2b \]

\textit{N-(2-Azidoacetyl)-N-methylbenzamide (2b)}

Prepared according to literature procedure,\textsuperscript{90} as a white solid (3.66 g, 78 %). \textsuperscript{1}H NMR (300 MHz, CDCl\textsubscript{3}) \( \delta \) 7.60-7.45 (m, 5H), 4.38 (s, 2H), 3.21 (s, 3H); \textsuperscript{13}C NMR (75 MHz, CDCl\textsubscript{3}) \( \delta \) 173.6, 171.9, 133.9, 132.6, 128.9, 128.3, 54.5, 35.2.
Ph\N\N_3

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

$^{13}\text{C NMR (CDCl}_3\text{)}$
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$_3$) $\delta$ 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$_3$) $\delta$ 172.9, 171.7, 138.2, 133.3, 129.6, 99.9, 54.5, 35.1.
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\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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\(_4\)O\(_3\) 249.10; found 249.10.
$N$-(2-Azidoacetyl)benzamide (2e)

$NaNN_3$ (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. $^1H$ NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 171.4, 165.7, 133.8, 131.5, 129.2, 127.9, 54.3.

$\text{Ph} \ \text{N} \ \text{H} \ \text{O} \ \text{O} \ \text{N}_3$  

$2e$

$\text{Ph} \ \text{N} \ \text{H} \ \text{O} \ \text{O} \ \text{N}_3$

$\text{2e}$

$\text{1H NMR (CDCl}_3\text{)}$

$\text{13C NMR (CDCl}_3\text{)}$
General procedure for the synthesis of imidazolinones 3\textsuperscript{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.

\[
\begin{align*}
&\text{O} \\
&\text{N} \\
&\text{N} \\
\end{align*}
\]

\(3a\)

1,2-Dimethyl-1\textit{H}-imidazol-5(4\textit{H})-one (3a)

Purified by flash chromatography (1:19 MeOH/EtOAc) to afford 3a (2.5 g, 87 %) as a yellow oil. \(^1\text{H}\) NMR (300 MHz, CDCl\(_3\)) \(\delta\) 3.98 (q, 2H, \(J = 2.4\) Hz), 3.00 (s, 3H), 2.15 (t, 3H, \(J = 2.4\) Hz); \(^{13}\text{C}\) NMR (75 MHz, CDCl\(_3\)) \(\delta\) 181.1, 163.1, 58.3, 26.3, 15.8. MS (ESI) m/z calcd for (M+H)\(^+\) \(\text{C}_7\text{H}_9\text{N}_2\text{O}\) 113.07; found 113.07.
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
1-Methyl-2-phenyl-1H-imidazol-5(4H)-one (3b)

Prepared according to literature procedure,\textsuperscript{90} as a red solid (1.0 g, 89 %). $^1$H NMR (300 MHz, CDCl\textsubscript{3}) $\delta$ 7.63-7.48 (m, 5H), 4.27 (2, 2H), 3.13 (s, 3H); $^{13}$C NMR (75 MHz, CDCl\textsubscript{3}) $\delta$ 182.0, 164.9, 131.2, 129.7, 128.8, 128.0, 58.9, 28.6.
2-(4-iodophenyl)-1-methyl-1H-imidazol-5(4H)-one (3c)

Purified by flash chromatography (9:1 CH₂Cl₂/EtOAc) to afford 3c (1.1 g, 83 %) as a yellow solid. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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)⁺ C₁₀H₁₀IN₂O 300.98; found 300.98.
2-(4-Methoxyphenyl)-1-methyl-1\textit{H}-imidazol-5(4\textit{H})-one (3d)

Purified by flash chromatography (3:2 CH\textsubscript{2}Cl\textsubscript{2}/EtOAc) to afford 3d (890 mg, 65\%) as a red oil. \textsuperscript{1}H NMR (300 MHz, CDCl\textsubscript{3}) \(\delta\) 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); \textsuperscript{13}C NMR (75 MHz, CDCl\textsubscript{3}) \(\delta\) 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)\textsuperscript{+} C\textsubscript{11}H\textsubscript{13}N\textsubscript{2}O\textsubscript{2} 205.10; found 205.09.
2-Phenyl-1H-imidazol-5(4H)-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$_9$H$_9$N$_2$O 161.07; found 161.07.
Synthesis of compounds \( Z - 4c \) and \( E - 4c \)

\[
\text{3c}
\]

\[
\text{Z - 4c}
\]

\[
\text{E - 4c}
\]

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\(_2\)Cl\(_2\), washed with water, brine and dried over Na\(_2\)SO\(_4\). 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.

\[
\text{3c}
\]

\[
\text{E - 4c}
\]
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 NaHCO₃(aq.) then taken into EtOAc. The organic layer was washed with NaHCO₃(aq.), water and dried over Na₂SO₄. The solvent was removed under reduced pressure to afford the pure product E-4c as a red solid (80 mg, quant.).

\[ \text{E-4c: } ^1\text{H NMR (500 MHz, CDCl}_3) \delta 10.62 \text{ (br, 1H), 7.85 (d, 2H, } J = 8.6 \text{ Hz), 7.48 (d, 2H, } J = 8.6 \text{ Hz), 7.17 (s, 1H), 5.87 (d, 1H, } J = 2.7 \text{ Hz), 3.32 (s, 3H), 2.31 (s, 3H), 2.21 (s, 3H); } ^{13}\text{C NMR (125 MHz, CDCl}_3) \delta 169.8, 155.3, 138.1, 138.0, 132.9, 130.6, 129.8, 129.3, 127.0, 117.1, 112.1, 97.2, 28.9, 13.9, 11.3. \text{ From proton-coupled } ^{13}\text{C NMR: } J_{\text{H6-C5}} = 3.8 \text{ Hz. IR (neat) } \nu_{\text{C=O}} = 1686 \text{ cm}^{-1}; \lambda_{\text{max abs}} = 458 \text{ nm (} \varepsilon = 40196 \text{ M}^{-1} \text{cm}^{-1}), \lambda_{\text{max emiss}} = 515 \text{ nm, } \Phi = 0.0005 \text{ in MeOH; MS (ESI) m/z calcd for (M+H)}^+ \text{ C}_{17}\text{H}_{17}\text{IN}_3\text{O 406.04; found 406.04.} \]

\[ \text{Z-4c: } ^1\text{H NMR (500 MHz, CDCl}_3) \delta 13.01 \text{ (br, 1H), 7.80 (d, 2H, } J = 8.5 \text{ Hz), 7.44 (d, 2H, } J = 8.5 \text{ Hz), 7.33 (s, 1H), 5.99 (d, 1H, } J = 2.2 \text{ Hz), 3.39 (s, 3H), 2.36 (s, 3H), 2.24 (s, 3H); } ^{13}\text{C NMR (125 MHz, CDCl}_3) \delta 168.6, 151.7, 138.2, 137.9, 134.4, 129.5, 129.3, 129.2, 127.2, 125.6, 113.5, 96.6, 29.1, 14.0, 11.5. \text{ From proton-coupled } ^{13}\text{C NMR: } J_{\text{H6-C5}} = 8.9 \text{ Hz. IR (neat) } \nu_{\text{C=O}} = 1663 \text{ cm}^{-1}; \lambda_{\text{max abs}} = 458 \text{ nm (} \varepsilon = 48034 \text{ M}^{-1} \text{cm}^{-1}), \lambda_{\text{max emiss}} = 515 \text{ nm, } \Phi = 0.0003 \text{ in MeOH; MS (ESI) m/z calcd for (M+H)}^+ \text{ C}_{17}\text{H}_{17}\text{IN}_3\text{O 406.04; found 406.04.} \]
H - coupled $^{13}$C NMR (CDCl$_3$)

*Z* - 4c

$^{169.84}$ ppm

H - coupled $^{13}$C NMR (CDCl$_3$)

*E* - 4c

$^{168.58}$ ppm
Photo and thermal isomerization

\[ \text{Z - 4c} \xrightarrow{hv} \text{E - 4c} \]

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

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 \textit{in}
vacuo to afford a yellow solid. The solid was dissolved in 1,2-dichloroethane (0.2 M) then Et₃N (3 eq.) was added. The reaction mixture was stirred at 25 °C for 15 min then BF₃•OEt₂ (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 Na₂SO₄ and evaporated under reduced pressure to give the crude product, which was purified by flash chromatography to afford the pure products 1 and 5.

![Image of 1a](image_url)

**GFP-chromophore analog 1a**

Purified by flash chromatography (3:2 EtOAc/hexanes) to afford 1a (153 mg, 24 %) as a yellow solid. ¹H NMR (300 MHz, d⁶-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); ¹³C NMR (75 MHz, d⁶-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; ¹⁹F NMR (282 MHz, d⁶-DMSO) δ 39.43 (q, J = 31.5 Hz); IR (neat) νC=O 1710 cm⁻¹; λ_max_abs 477 nm (ε = 57286 M⁻¹cm⁻¹), λ_max_emiss 485 nm, Φ = 0.89 in MeOH; MS (ESI) m/z calcd for (M+H)⁺ C₁₂H₁₅BF₂N₃O 266.13; found 266.13.
$^{1}\text{H NMR (d}^6\text{-DMSO)}$

$^{13}\text{C NMR (d}^6\text{-DMSO)}$

$^{19}\text{F NMR (d}^6\text{-DMSO)}$
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) $\delta$ 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) $\delta$ 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) $\delta$ 37.33 (q, $J = 18.4$ Hz); IR (neat) $\nu_{C=O}$ 1622 cm$^{-1}$; $\lambda_{max \text{ abs}}$ 426 nm ($\epsilon = 37009$ M$^{-1}$cm$^{-1}$), $\lambda_{max \text{ emiss}}$ 485 nm, $\Phi = 0.0007$ in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ C$_{12}$H$_{15}$BF$_2$N$_3$O 266.13; found 266.13.
$^1$H NMR ($d^6$-DMSO)

$^{13}$C NMR ($d^6$-DMSO)
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 40.92 (q, $J = 30.6$ Hz); IR (neat) $\nu_{\text{C}=\text{O}}$ 1687 cm$^{-1}$; $\lambda_{\text{max abs}}$ 490 nm ($\varepsilon = 58684$ M$^{-1}$cm$^{-1}$), $\lambda_{\text{max emiss}}$ 521 nm, $\Phi = 0.87$ in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ C$_{17}$H$_{17}$BF$_2$N$_3$O 328.14; found 328.14.
$$\text{1b}$$

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

$${}^{13}\text{C NMR (CDCl}_3)$$
**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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 34.22 (q, $J = 19.3$ Hz); IR (neat) $\nu_{C=O}$ 1640 cm$^{-1}$; $\lambda_{\text{max abs}}$ 463 nm ($\varepsilon = 50452$ M$^{-1}$cm$^{-1}$), $\lambda_{\text{max emiss}}$ 517 nm, $\Phi = 0.006$ in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ C$_{17}$H$_{17}$BF$_2$N$_3$O 328.14; found 328.13.

**$^{19}$F NMR (CDCl$_3$)**
GFP-chromophore analog 1c

Purified by flash chromatography (1:9 EtOAc/hexanes) to afford 1c (672 mg, 24 %) as an orange solid. \( ^1H \) NMR (300 MHz, CDCl\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 37.02 (q, \( J = 27.5 \) Hz); IR (neat) \( \nu_{C=O} \) 1705 cm\(^{-1}\); \( \lambda_{\text{max abs}} \) 492 nm (\( \varepsilon = 57303 \) M\(^{-1}\)cm\(^{-1}\)), \( \lambda_{\text{max emiss}} \) 532 nm, \( \Phi = 0.86 \) in MeOH; MS (ESI) m/z calcd for (M+H\(^+\)) \( C_{17}H_{16}BF_2IN_3O \) 454.04; found 454.04.

\( ^{19}F \) NMR (CDCl\(_3\))
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 34.20 (q, $J = 18.9$ Hz); IR (neat) $\nu_{C=O}$ 1610 cm$^{-1}$; $\lambda_{max \, abs}$ 472 nm ($\varepsilon = 42669 \, M^{-1} cm^{-1}$), $\lambda_{max \, emiss}$ 520 nm, $\Phi = 0.01$ in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ C$_{17}$H$_{16}$BF$_2$IN$_3$O 454.04; found 454.03.
GFP-chromophore analog 1d

Purified by flash chromatography (1:1:3 EtOAc/CH₂Cl₂/hexanes) to afford 1d (48 mg, 13 %) as a red solid. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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; ¹⁹F NMR (282 MHz, CDCl₃) δ 41.17 (q, J = 27.4 Hz); IR (neat) ν_C=O 1664 cm⁻¹; λ_max abs 493 nm (ε = 51269 M⁻¹cm⁻¹), λ_max emiss 521 nm, Φ = 0.85 in MeOH; MS (ESI) m/z calcd for (M+H)⁺ C₁₈H₁₉BF₂N₃O₂ 358.15; found 358.15.
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
GFP-chromophore analog 5d

Purified by flash chromatography (1:1:3 EtOAc/CH₂Cl₂/hexanes) to afford 5d (144 mg, 40 %) as a red solid. ¹H NMR (500 MHz, CDCl₃) 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); ¹³C NMR (125 MHz, CDCl₃) 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; ¹⁹F NMR (282 MHz, CDCl₃) δ 34.29 (q, J = 18.7 Hz); IR (neat) νₐₗ₉ 1616 cm⁻¹; λₚₘₐₓₐbs 455 nm (ε = 38642 M⁻¹cm⁻¹), λₚₘₐₓₑₗ₉iss 528 nm, Φ = 0.03 in MeOH; MS (ESI) m/z calcd for (M+H)+ C₁₈H₁₉BF₂N₃O₂ 358.15; found 358.15.
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$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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$_3$OD) $\delta$ 41.35 (q, $J$ = 29.1 Hz); IR (neat) $\nu$$_{C=O}$ 1670 cm$^{-1}$; $\lambda$$_{\text{max abs}}$ 495 nm ($\epsilon$ = 37867 M$^{-1}$cm$^{-1}$), $\lambda$$_{\text{max emiss}}$ 531 nm, $\Phi$ = 0.80 in MeOH; MS (ESI) m/z calcd for (M+H)$^+$ C$_{16}$H$_{15}$BF$_2$N$_3$O 314.13; found 314.13.

GFP-chromophore analog 5e

Purification was unsuccessful.
A solution of 2-(4-nitrobenzamido) acetic acid (1.48 g, 6.6 mmol) in 10 mL Ac₂O 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 CH₂Cl₂/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.
NMR (500 MHz, CDCl3) **Z-6**: δ 10.30 (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)+ C16H14N3O4 312.10; found 312.09.

CH3NH2 (40 % in H2O, 1.9 mL, 22 mmol) was added to the solution of 6 (685 mg, 2.2 mmol) and K2CO3 (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 (CH2Cl2 >1:19 EtOAc/CH2Cl2) 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. 1H NMR (500 MHz, CDCl3) **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)+ C17H17N4O3 325.13; found 325.13.

Et3N (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 BF3.OEt2 (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/CH2Cl2) to afford compound 1f (84 mg, 46 %) as a red solid and 5f (79 mg, 44 %) as a red solid.

For 1f: 1H NMR (500 MHz, CDCl3) 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); 13C NMR (500 MHz, CDCl3) 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\(_3\)) \( \delta \) 41.58 (q, \( J = 28.1 \) Hz); IR (neat) \( \nu_{C=O} \) 1678 cm\(^{-1}\); \( \lambda_{max \, abs} \) 492 nm (\( \epsilon = 44003 \) M\(^{-1}\)cm\(^{-1}\)), \( \lambda_{max \, emiss} \) 598 nm, \( \Phi = 0.0004 \) 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\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 34.31 (q, \( J = 18.5 \) Hz); IR (neat) \( \nu_{C=O} \) 1612 cm\(^{-1}\); \( \lambda_{max \, abs} \) 472 nm (\( \epsilon = 35502 \) M\(^{-1}\)cm\(^{-1}\)), \( \lambda_{max \, emiss} \) 525 nm, \( \Phi = 0.005 \) in MeOH; MS (ESI) m/z calcd for (M+H\(^+\)) \( C_{17}H_{16}BF_2N_4O_3 \) 373.13; found 373.12.
$^{19}$F NMR (CDCl$_3$)

![Chemical Structure 1f]

$^{19}$F NMR (CDCl$_3$)

![Chemical Structure 5f]
Recycle of undesired GFP-chromophore analog 5

Compound 5c (227 mg, 0.5 mmol) was dissolved in 5 mL CH₂Cl₂ and 5 mL 1:9 H₂SO₄/TFA was added. The solution was stirred at 25 °C for 10 min then poured into 20 mL ice-water and extracted with CH₂Cl₂ (3 x 50 mL). The organic layers were combined, washed carefully with saturated NaHCO₃ (3 x 30 mL), H₂O (1 x 30 mL), brine (1 x 30 mL) and dried over Na₂SO₄. 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 ClCH₂CH₂Cl. After refluxing for 1 h, the solution was cooled to room temperature and Et₃N (0.35 mL, 2.5 mmol) was added, followed by BF₃•OEt₂ (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 CH₂Cl₂ and washed with 30 mL H₂O, 30 mL brine and dried over Na₂SO₄. 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**

ClSO$_3$H (53 $\mu$L, 0.8 mmol) was added dropwise to the solution of 1c (91 mg, 0.2 mmol) in 2 mL CH$_2$Cl$_2$ at 0 $^\circ$C. After stirring at 0 $^\circ$C for 10 min, the reaction mixture was poured into 20 mL ice-water, neutralized with saturated NaHCO$_3$(aq.) and extracted with iPrOH/CH$_2$Cl$_2$ (1:1) (3 x 20 mL). The organic layer was dried over Na$_2$SO$_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography (1:9 MeOH/CH$_2$Cl$_2$) to afford the product 1g (102 mg, 92 \%) as an orange solid. $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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) $\nu_{C=O}$ 1688 cm$^{-1}$; $\lambda_{max \, ab} = 488$ nm ($\varepsilon = 48132 \, M^{-1}cm^{-1}$), $\lambda_{max \, emiss} = 531$ nm, $\Phi = 0.88$ in MeOH; MS (ESI) m/z calcd for (M-Na)$^-\ C_{17}H_{14}BF_{2}IN_{3}O_{4}S$ 531.98; found 531.88.
Synthesis of GFP-chromophore analog 1h

\[
\begin{align*}
\text{NaO}_2\text{S} & \quad \begin{array}{c} \text{N} \\
\text{B} & \quad \begin{array}{c} \text{F}_2 \\
\text{N} & \quad \begin{array}{c} \text{O} \\
\text{I} & \quad \begin{array}{c} \text{NaO}_3\text{S} \\
\text{COOH} & \quad \begin{array}{c} \text{PdCl}_2(\text{PPh}_3)_2 \\
10\text{mol} \% \\
\text{CuI, 10 eq. Et}_3\text{N} \\
\text{DMF, 25} \, ^\circ\text{C, 12 h} \\
\end{array} \\
\end{array} \\
\end{array} \\
\end{array} \\
\end{align*}
\[
\begin{align*}
\text{N} & \quad \begin{array}{c} \text{B} \\
\text{F}_2 & \quad \begin{array}{c} \text{N} \\
\text{O} & \quad \begin{array}{c} \text{HO}_3\text{S} \\
\text{COOH} & \quad \begin{array}{c} \text{1h} \\
\text{Compound} 1\text{g (56 mg, 0.1 mmol), PdCl}_2(\text{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}_3\text{N (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}_2\text{O and acidified to pH} \approx 2.0 \text{ with 1N HCl. The solution was extracted with} \text{PrOH/CH}_2\text{Cl}_2 (1:1) (3 x 30 mL). The organic layer was dried over Na}_2\text{SO}_4 \text{ and concentrated in vacuo. The crude product was purified by flash chromatography (1:4 MeOH/CH}_2\text{Cl}_2) \text{ to give the desired product 1h as an orange solid (37 mg, 71 %).} \text{ 1H NMR (500 MHz, CD}_3\text{OD}) \delta 7.74 (d, 2H, J \approx 8.4 Hz), 7.65 (s, 1H), 7.61 (d, 2H, J \approx 8.4 Hz), 3.22 (s, 3H), 2.57-2.47 (m, 10H), 1.96-1.90 (m, 2H); \text{ 13C NMR (125 MHz, d}^6\text{-DMSO}) \delta 178.0, 160.7, 155.3, 146.2, 134.2, 133.2, 131.1, 130.1, 127.1, 127.0, 123.4, 123.1, 117.9, 117.8, 94.3, 79.9, 35.7, 28.2, 24.9, 18.6, 13.0, 10.4. \text{ IR (neat) ν}_c=\text{o} \quad 1702, 1593\text{cm}^{-1}; \nu_{\text{max abs}} 488 \text{ nm (ε = 35024}}
\end{align*}
M⁻¹cm⁻¹), $\lambda_{\text{max} \ emiss} = 538$ nm, $\Phi = 0.88$ in MeOH; MS (ESI) m/z caled for (M-H)⁻ $C_{23}H_{21}BF_2N_3O_6S$ 516.12; found 516.05.
Labeling of streptavidin with 1h

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 NaHCO₃ 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

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\(_2\)Cl\(_2\) 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\(_3\)) \(\delta\) 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\(_2\)N\(_4\)O\(_2\) 501.09; found 501.10 (X = Br); calcd for (M+H)+ C\(_{22}\)H\(_{21}\)BF\(_2\)IN\(_4\)O\(_2\) 549.08; found 549.08 (X = I).
A solution of the above product 7 (70 mg, 0.14 mmol) in 10 mL CH₂Cl₂ was cooled to 0 °C and ClSO₃H (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 H₂O. The mixture was concentrated under reduced pressure and the residue was purified by flash chromatography (20 % MeOH/CH₂Cl₂) to give a red solid. This intermediate was dissolved in 6 mL CH₃CN and 4 mL H₂O 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/CH₂Cl₂) to afford the product 1i (86 mg, 98 %) as a red solid. ¹H NMR (500 MHz, CD₃OD) 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); ¹³C NMR (125 MHz, CD₃OD) 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)⁻ C₂₂H₁₀BF₂I₄N₄O₅S 627.02; found 626.93.
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

<table>
<thead>
<tr>
<th></th>
<th>$A$</th>
<th>$G_{b,88}$</th>
<th>$4c^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aqueous solution</td>
<td>Gas phase</td>
<td>Aqueous solution</td>
</tr>
<tr>
<td>$Z$ (Hartree)</td>
<td>$-$</td>
<td>-837.84379</td>
<td>-837.84976</td>
</tr>
<tr>
<td>$E$ (Hartree)</td>
<td>$-$</td>
<td>-837.84114</td>
<td>-837.84715</td>
</tr>
<tr>
<td>Energy of $E$ vs $Z$ (Kcal/mol)</td>
<td>2.1</td>
<td>1.66</td>
<td>1.64</td>
</tr>
</tbody>
</table>

$^a$ Experimental value; $^b$ calculated with B3LYP/6-31+G** (SDD for I in 4c)

Table B2. Calculated spectroscopic properties of $4c^a$

<table>
<thead>
<tr>
<th></th>
<th>$\Delta E$, eV</th>
<th>$\lambda_{\text{max}}$, nm</th>
<th>Oscillator strength ($f$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Gas phase</td>
<td>DMSO</td>
</tr>
<tr>
<td>$Z - 4c$</td>
<td>2.83</td>
<td>438</td>
<td>0.70</td>
</tr>
<tr>
<td>$E - 4c$</td>
<td>2.89</td>
<td>430</td>
<td>0.99</td>
</tr>
<tr>
<td>$Z - 4c$</td>
<td>2.78</td>
<td>446 ($458^b$)</td>
<td>0.89</td>
</tr>
<tr>
<td>$E - 4c$</td>
<td>2.81</td>
<td>442 ($464^b$)</td>
<td>1.09</td>
</tr>
</tbody>
</table>

$^a$ Calculated with B3LYP/6-31+G** (SDD for I); $^b$ data in bold are experimental values
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

<table>
<thead>
<tr>
<th></th>
<th>1H shift to lower field</th>
<th>19F shift to higher field</th>
<th>weaker C=O bond</th>
<th>strong 19F-13C coupling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 19F signals shown here are 0.1 x their actual values.
<table>
<thead>
<tr>
<th></th>
<th>$\nu_{C=O}$ (cm$^{-1}$)</th>
<th>$\delta_{C5}$ (ppm)</th>
<th>$\delta_F$ (ppm)</th>
<th>$\delta_{H6}$ (ppm)</th>
<th>$\delta_{H4'}$ (ppm)</th>
<th>$\delta_{H6'}$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>1710</td>
<td>160.6</td>
<td>39.43</td>
<td>7.49</td>
<td>6.04</td>
<td>2.31</td>
</tr>
<tr>
<td>5a</td>
<td>1622 (t, $J = 5.3$Hz)</td>
<td>158.7</td>
<td>37.33</td>
<td>7.68</td>
<td>6.37</td>
<td>2.44</td>
</tr>
<tr>
<td>1b</td>
<td>1687</td>
<td>161.0</td>
<td>40.92</td>
<td>7.45</td>
<td>6.01</td>
<td>2.34</td>
</tr>
<tr>
<td>5b</td>
<td>1640 (t, $J = 5.3$Hz)</td>
<td>158.6</td>
<td>34.22</td>
<td>7.66</td>
<td>6.26</td>
<td>2.58</td>
</tr>
<tr>
<td>1c</td>
<td>1705</td>
<td>160.7</td>
<td>37.02</td>
<td>7.41</td>
<td>6.01</td>
<td>2.34</td>
</tr>
<tr>
<td>5c</td>
<td>1610 (t, $J = 6.1$Hz)</td>
<td>158.4</td>
<td>34.20</td>
<td>7.64</td>
<td>6.28</td>
<td>2.58</td>
</tr>
<tr>
<td>1d</td>
<td>1664</td>
<td>161.3</td>
<td>41.17</td>
<td>7.41</td>
<td>5.99</td>
<td>2.35</td>
</tr>
<tr>
<td>5d</td>
<td>1616 (t, $J = 5.5$Hz)</td>
<td>158.7</td>
<td>34.29</td>
<td>7.60</td>
<td>6.22</td>
<td>2.54</td>
</tr>
<tr>
<td>1e</td>
<td>1670</td>
<td>163.6</td>
<td>41.35</td>
<td>7.41</td>
<td>6.05</td>
<td>2.40</td>
</tr>
<tr>
<td>5e</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1f</td>
<td>1678</td>
<td>160.3</td>
<td>41.58</td>
<td>7.44</td>
<td>6.03</td>
<td>2.33</td>
</tr>
<tr>
<td>5f</td>
<td>1612 (t, $J = 5.0$Hz)</td>
<td>158.1</td>
<td>34.31</td>
<td>7.65</td>
<td>6.32</td>
<td>2.61</td>
</tr>
</tbody>
</table>
Table B4. Spectroscopic properties of GFP-chromophore analogs

<table>
<thead>
<tr>
<th></th>
<th>λ\textsubscript{max abs} (nm)</th>
<th>ε (M\textsuperscript{-1} cm\textsuperscript{-1})</th>
<th>λ\textsubscript{max emiss} (nm)</th>
<th>fwhm (nm)</th>
<th>Φ\textsuperscript{a}</th>
<th>solvent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>477</td>
<td>57286</td>
<td>485</td>
<td>19</td>
<td>0.89 ± 0.01</td>
<td>MeOH</td>
</tr>
<tr>
<td>1a</td>
<td>480</td>
<td>52809</td>
<td>491</td>
<td>21</td>
<td>n.d.</td>
<td>CH\textsubscript{2}Cl\textsubscript{2}</td>
</tr>
<tr>
<td>5a</td>
<td>426</td>
<td>37009</td>
<td>485</td>
<td>25</td>
<td>0.0007</td>
<td>MeOH</td>
</tr>
<tr>
<td>1b</td>
<td>490</td>
<td>58684</td>
<td>521</td>
<td>46</td>
<td>0.87 ± 0.01</td>
<td>MeOH</td>
</tr>
<tr>
<td>1b</td>
<td>494</td>
<td>42641</td>
<td>530</td>
<td>48</td>
<td>n.d.</td>
<td>CH\textsubscript{2}Cl\textsubscript{2}</td>
</tr>
<tr>
<td>5b</td>
<td>463</td>
<td>50452</td>
<td>517</td>
<td>64</td>
<td>0.006</td>
<td>MeOH</td>
</tr>
<tr>
<td>1c</td>
<td>492</td>
<td>57303</td>
<td>532</td>
<td>54</td>
<td>0.86 ± 0.02</td>
<td>MeOH</td>
</tr>
<tr>
<td>1c</td>
<td>498</td>
<td>44990</td>
<td>546</td>
<td>53</td>
<td>n.d.</td>
<td>CH\textsubscript{2}Cl\textsubscript{2}</td>
</tr>
<tr>
<td>5c</td>
<td>463</td>
<td>42669</td>
<td>530</td>
<td>54</td>
<td>0.01</td>
<td>MeOH</td>
</tr>
<tr>
<td>1d</td>
<td>493</td>
<td>51269</td>
<td>521</td>
<td>38</td>
<td>0.85 ± 0.03</td>
<td>MeOH</td>
</tr>
<tr>
<td>1d</td>
<td>496</td>
<td>47448</td>
<td>530</td>
<td>43</td>
<td>n.d.</td>
<td>CH\textsubscript{2}Cl\textsubscript{2}</td>
</tr>
<tr>
<td>5d</td>
<td>455</td>
<td>38642</td>
<td>528</td>
<td>69</td>
<td>0.03</td>
<td>MeOH</td>
</tr>
<tr>
<td>1e</td>
<td>495</td>
<td>37867</td>
<td>531</td>
<td>48</td>
<td>0.80 ± 0.01</td>
<td>MeOH</td>
</tr>
<tr>
<td>1e</td>
<td>498</td>
<td>44990</td>
<td>546</td>
<td>46</td>
<td>n.d.</td>
<td>CH\textsubscript{2}Cl\textsubscript{2}</td>
</tr>
<tr>
<td>1f</td>
<td>492</td>
<td>44003</td>
<td>598</td>
<td>120</td>
<td>0.0004</td>
<td>MeOH</td>
</tr>
<tr>
<td>1f</td>
<td>498</td>
<td>n.d.</td>
<td>587</td>
<td>53</td>
<td>0.53 ± 0.01</td>
<td>Hexanes</td>
</tr>
<tr>
<td>1f</td>
<td>499</td>
<td>38658</td>
<td>642</td>
<td>150</td>
<td>n.d.</td>
<td>CH\textsubscript{2}Cl\textsubscript{2}</td>
</tr>
<tr>
<td>5f</td>
<td>473</td>
<td>35502</td>
<td>527</td>
<td>73</td>
<td>0.005</td>
<td>MeOH</td>
</tr>
<tr>
<td>5f</td>
<td>486</td>
<td>n.d.</td>
<td>545</td>
<td>97</td>
<td>0.05</td>
<td>Hexanes</td>
</tr>
<tr>
<td>1g</td>
<td>488</td>
<td>48132</td>
<td>531</td>
<td>53</td>
<td>0.88 ± 0.01</td>
<td>MeOH</td>
</tr>
<tr>
<td>1g</td>
<td>481</td>
<td>46461</td>
<td>518</td>
<td>50</td>
<td>0.87 ± 0.01</td>
<td>Phos 7.4\textsuperscript{b}</td>
</tr>
<tr>
<td>1h</td>
<td>488</td>
<td>35024</td>
<td>538</td>
<td>58</td>
<td>0.84 ± 0.01</td>
<td>MeOH</td>
</tr>
<tr>
<td>1h</td>
<td>482</td>
<td>34766</td>
<td>526</td>
<td>56</td>
<td>0.82 ± 0.01</td>
<td>Phos 7.4\textsuperscript{b}</td>
</tr>
<tr>
<td>1s</td>
<td>482</td>
<td>n.d.</td>
<td>529</td>
<td>56</td>
<td>n.d.</td>
<td>Phos 7.4\textsuperscript{b}</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Fluorescein in 0.1M NaOH as standard (Φ = 0.92); \textsuperscript{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

POCl₃ (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 Na₂CO₃ solution (2 x 50 mL) and brine (1 x 50 mL). The organic layer was dried over Na₂SO₄ and the solvents were removed under reduced pressure. The residue was purified by flash chromatography (SiO₂, 20 % EtOAc/Hexanes) to afford the pure product 8a as a light yellow solid (5.8 g, 89 %). ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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.
4,5-Dihydro-1H-benzo[g]indole-2-carbaldehyde (8b)

Yellow solid (910 mg, 85%). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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.
7-Methoxy-4,5-dihydro-1H-benzo[g]indole-2-carbaldehyde (8c)

Light green solid (345 mg, 100 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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.
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.
4-Ethyl-3,5-dimethyl-1H-pyrrole-2-carbaldehyde (8e)

Brown solid (1.1 g, 85 %). \(^1\)H NMR (300 MHz, CDCl\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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.
5-Ethyl-1H-pyrrole-2-carbaldehyde (8f)

Red solid (1.0 g, 78 %). \(^1\)H NMR (300 MHz, CDCl\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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.
4,5,6,7-Tetrahydro-1H-indole-2-carbaldehyde (8g)

Yellow solid (960 mg, 78 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_9$H$_{12}$NO 150.09; found 150.09.
4-Acetyl-3,5-dimethyl-1H-pyrrole-2-carbaldehyde (8h)

Light yellow solid (481 mg, 40 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_9$H$_{12}$NO$_2$ 166.09; found 166.09.
Typical procedure for the synthesis of symmetric BODIPYs

\[
\begin{align*}
\text{8a} & \xrightarrow{\text{(i) POCl}_3, \text{CH}_2\text{Cl}_2, \text{0-25} \degree \text{C, 12 h}}} \\
\text{9a} & \xrightarrow{\text{(ii) Et}_3\text{N, BF}_3\text{OEt}_2, \text{25} \degree \text{C, 12 h}}} 
\end{align*}
\]

3,5-Dimethyl-1H-pyrrole-2-carbaldehyde **8a** (246 mg, 2 mmol) was dissolved in 10 mL CH\textsubscript{2}Cl\textsubscript{2} and POCl\textsubscript{3} (0.22 mL, 2.4 mmol) was added dropwise over 1 min at 0 \degree \text{C}. The solution was warmed to room temperature slowly and stirred for 12 h. The mixture was cooled to 0 \degree \text{C} and Et\textsubscript{3}N (1.4 mL, 10 mmol) was added dropwise over 5 min. After stirring for 15 min, BF\textsubscript{3}OEt\textsubscript{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\textsubscript{2}Cl\textsubscript{2} to remove the polar impurities. The solvents were removed under reduced pressure. The residue was dissolved in CH\textsubscript{2}Cl\textsubscript{2} and water was added, and the mixture was stirred at room temperature overnight. (to decompose excess BF\textsubscript{3}OEt\textsubscript{2} and other impurities). The organic layer was washed with water, brine and dried over Na\textsubscript{2}SO\textsubscript{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. \textsuperscript{1}H NMR (CDCl\textsubscript{3}, 500 MHz) \(\delta\) 7.01 (s, 1H), 6.02 (s, 2H), 2.51 (s, 6H), 2.22 (s, 6H); \textsuperscript{13}C NMR (CDCl\textsubscript{3}, 125 MHz) \(\delta\) 156.7, 141.2, 133.4, 120.0, 119.0, 14.6, 11.2; \textsuperscript{19}F NMR (CDCl\textsubscript{3}, 282 MHz) \(\delta\) 30.68 (q, \(J = 33.5\) Hz). HRMS (ESI) m/z calcd for (M+H)\textsuperscript{+} C\textsubscript{13}H\textsubscript{16}BF\textsubscript{2}N\textsubscript{2} 249.1375; found 249.1373.

**9b-9h** were prepared using similar methods as described for **9a**.
BODIPY 9b
Green solid (360 mg, 91%). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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) $\delta$ 39.67 (q, $J = 33.6$ Hz). MS (MALDI) m/z calcd for M$^+$ C$_{25}$H$_{19}$BF$_2$N$_2$ 396.16; found 395.88.
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
BODIPY 9c

Blue solid (194 mg, 85 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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) $\delta$ 39.15 (q, $J = 33.6$ Hz). MS (MALDI) m/z calcd for M$^+$ C$_{27}$H$_{23}$BF$_2$N$_2$O$_2$ 456.18; found 455.88.
BODIPY 9d

Green solid (52 mg, 21%). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 7.98-7.95 (m, 4H), 7.54-7.42 (m, 17H), 6.73 (s, 2H); $^{13}$C NMR (75 MHz, CDCl$_3$) $\delta$ 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) $\delta$ 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.
$^{1}H$ NMR (CDCl$_3$)

$^{13}C$ NMR (CDCl$_3$)
BODIPY 9e
Organge solid (227 mg, 75 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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) $\delta$ 30.93 (q, $J$ = 33.6 Hz). MS (MALDI) m/z calcd for M$^+$ C$_{17}$H$_{23}$BF$_2$N$_2$ 304.19; found 303.96.
BODIPY 9f

Orange solid (70 mg, 28 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 164.1, 134.4, 130.1, 127.1, 117.4, 22.0, 12.6; $^{19}$F NMR (CDCl$_3$, 282 MHz) $\delta$ 31.94 (q, $J = 33.4$ Hz). MS (MALDI) m/z calcd for M$^+$ C$_{13}$H$_{15}$BF$_2$N$_2$ 248.13; found 247.85.

$^{19}$F NMR (CDCl$_3$)
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
BODIPY 9g

Red solid (65 mg, 22 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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) $\delta$ 27.29 (q, $J = 33.5$ Hz). MS (MALDI) m/z calcd for M$^+ \text{C}_{17}\text{H}_{19}\text{BF}_2\text{N}_2$ 300.16; found 299.91.

BODIPY 9h

Messy reaction, obtained complex mixtures.
$^1\text{H NMR} (\text{CDCl}_3)$

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

$^{19}\text{F NMR} (\text{CDCl}_3)$
NMR and MS study of the self-condensation reaction:

Compound 8a (25 mg, 0.2 mmol) was dissolved in 1 mL of CDCl₃ and then POCl₃ (50 μL) was added. The reaction was monitored by $^1$H, $^{13}$C and $^{31}$P over 10 min intervals at room temperature.
APPENDIX D
EXPERIMENTAL DATA FOR CHAPTER IV

Synthesis of 3,6-ditriflyxanthone 16

3,6-Dihydroxy-xanthen-9-one (6.85 g, 30 mmol) was dissolved in 150 mL CH$_2$Cl$_2$ and pyridine (24.5 mL, 300 mmol) was added slowly over 5 min at 0 °C. The mixture was stirred at 0 °C for 10 min then Tf$_2$O (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$_2$SO$_4$. The solvents were removed under reduced pressure and the residue was recrystallized from CH$_2$Cl$_2$/hexanes to afford the pure product as a white crystal (13.2 g, 89 %). $^1$H 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).

$^{19}$F 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 Na₂CO₃ (aq.) and water to give the crude product, which was recrystallized from EtOAc/hexanes to afford the pure product.

3,6-Di-piperidin-1-yl-xanthen-9-one (17)
Yellow solid (3.0 g, 89 %). ¹H NMR (300 MHz, d⁶-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); ¹³C NMR (75 MHz, d⁶-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)⁺ C₂₃H₂₆N₂O₂ 363.21; found 363.21.
$^1$H NMR (d$^6$-DMSO)

$^{13}$C NMR (d$^6$-DMSO)
3,6-Di-morpholin-4-yl-xanthen-9-one (18)

White solid (308 mg, 93 %). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$O$_4$ 367.17; found 367.17.
3,6-Bis-(4-Boc-piperazin-1-yl)-xanthen-9-one (19)

White solid (680 mg, 60%). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_4$O$_6$ 565.30; found 565.29.
3,6-Di-pyrrolidin-1-yl-xanthen-9-one (20)

Yellow solid (90 mg, 18%). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_{33}$N$_2$O$_2$ 335.18; found 335.18.
Synthesis of 3-triflyl-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 CH₂Cl₂ 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. Rₚ = 0.66 (40 % EtOAc/hexanes). ¹H NMR (300 MHz, d₆-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, 2.4Hz), 6.82 (d, 1H, J = 2.4 Hz), 3.50-3.47 (m, 4H), 1.60 (br, 6H); ¹³C NMR (75 MHz, d₆-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)+ C₁₉H₁₇F₃NO₅S 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 °C and stirred for 12 h. After cooling to room temperature, the solution was diluted with 30 mL CH₂Cl₂ and washed with saturated Na₂CO₃ (aq.) (1 x 30 mL), water (1 x 30 mL) and dried over Na₂SO₄. 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, CDCl₃) \( \delta \) 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, CDCl₃) \( \delta \) 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 \((M+H)^+\) C₂₂H₂₈N₂O₃ 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 °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 CH₂Cl₂. The organic layer was washed with water, brine, dried over Na₂SO₄ and concentrated under reduced pressure. The residue was purified by flash chromatography (5% to 10% MeOH/CH₂Cl₂) to give the pure product.

Rosamine 11a
Purple solid (71 mg, 89 %). \( R_f = 0.28 \) (10 % MeOH/CH₂Cl₂). \(^1\)H NMR (500 MHz, CDCl₃) \( \delta \) 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, CDCl₃) \( \delta \) 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)\(^+\) C₂₄H₂₉N₂O 361.2280; found 361.2287.
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
Rosamine 11b

Purple solid (66 mg, 70 %). \( R_f = 0.34 \) (10 % MeOH/CH\(_2\)Cl\(_2\)). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \( \delta \) 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); \(^1^3\)C NMR (125 MHz, CDCl\(_3\)) \( \delta \) 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.

\(^1\)H NMR (CDCl\(_3\))

\(^1^3\)C NMR (CDCl\(_3\))
Rosamine 11c

Green solid (92 mg, 100 %). $R_f = 0.35$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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-Cl)$^+$ C$_{20}$H$_{31}$N$_2$O 423.2436; found 423.2437.

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

$^{13}$C NMR (CDCl$_3$)
Rosamine 11d

Green solid (105 mg, 90 %) [Made from 4-Iodophenyl magnesium chloride, which was prepared from 1,4-di-iodobenzene with iPrMgCl]. $R_f = 0.36$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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-Cl)$^+$ C$_{29}$H$_{30}$IN$_2$O 549.1403; found 549.1399. Anal. Calcd for C$_{29}$H$_{30}$ClIN$_2$O: 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.
Rosamine 11e

Green solid (98 mg, 100 %). \( R_f = 0.33 \) (10 % MeOH/CH\(_2\)Cl\(_2\)). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 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-Cl)+ C\(_{30}\)H\(_{33}\)N\(_2\)O\(_2\) 453.2542; found 453.2546.

\(^1\)H NMR (CDCl\(_3\))

\(^{13}\)C NMR (CDCl\(_3\))
Rosamine 11f
Green solid (97 mg, 100 %).  $R_f = 0.36$ (10 % MeOH/CH₂Cl₂).  $^1$H NMR (500 MHz, CDCl₃) 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, CDCl₃) 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)$^+$ C₃₀H₃₃N₂O₂ 453.2542; found 453.2540.
Rosamine 11g
Green solid (69 mg, 74 %). \( R_f = 0.25 \) (10 % MeOH/CH\(_2\)Cl\(_2\)). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 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\(^+\)) \( C_{27}H_{29}N_2O_S \) 429.2001; found 429.1995.

\(^1\)H NMR (CDCl\(_3\))

\(^{13}\)C NMR (CDCl\(_3\))
Rosamine 11h

Green solid (100 mg, 96 %). $R_f = 0.32$ (10 % MeOH/CH$_2$Cl$_2$). [Made from 2,6-dimethoxyphenyl lithium, which was prepared from 1,3-dimethoxybenzene with nBuLi].

$^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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-Cl)$^+$ C$_{31}$H$_{35}$N$_2$O$_3$ 483.2648; found 483.2656.

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

$^{13}$C NMR (CDCl$_3$)
Rosamine 11i
Green solid (95 mg, 99 %). $R_f = 0.32$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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$_2$O 437.2593; found 437.2598.

$^1$H NMR (CD$_3$OD)

$^{13}$C NMR (CD$_3$OD)
Rosamine 12i

Green solid (94 mg, 98 %). $R_f = 0.26$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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$_2$O$_3$ 441.2178; found 441.2184.

$^1$H NMR (CD$_3$OD)

$^{13}$C NMR (CD$_3$OD)
Rosamine 13i

Green solid (47 mg, 53 %). $R_f = 0.28$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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$_2$O 409.2280; found 409.2277.

$^1$H NMR (CD$_3$OD)

$^{13}$C NMR (CD$_3$OD)
Rosamine 22i

Green solid (93 mg, 98 %). $R_f = 0.30$ (10% MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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$_2$O$_2$ 439.2386; found 439.2384.

$^1$H NMR (CD$_3$OD)

$^{13}$C NMR (CD$_3$OD)
Rosamine 14i
Purple solid (121 mg, 90 %). $R_f = 0.37$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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-Cl)$^+$ C$_{38}$H$_{47}$N$_4$O$_5$ 639.3546; found 639.3553.

$^1$H NMR (CD$_3$OD)

$^{13}$C NMR (CD$_3$OD)
Rosamine 14d

Green solid (107 mg, 68%).  $R_f = 0.31$ (10% MeOH/CH$_2$Cl$_2$).  $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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-Cl)$^+$ C$_{37}$H$_{44}$N$_4$O$_5$ 751.2356; found 751.2342.

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

$^{13}$C NMR (CDCl$_3$)
Rosamine 14h

Purple solid (135 mg, 94 %).  $R_f = 0.31$ (10 % MeOH/CH$_2$Cl$_2$).  $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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-Cl)$^+$ C$_{39}$H$_{49}$N$_4$O$_7$ 685.3601; found 685.3598.

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

$^{13}$C NMR (CDCl$_3$)
Synthesis of rosamine 15h

A solution of rosamine 14h (50 mg, 0.07 mmol) in 5 mL TFA/CH₂Cl₂ (1:1) was stirred at room temperature for 1 h. The solvents were removed with a N₂ stream. The residue was dissolved in 15 mL iPrOH/CH₂Cl₂ (1:1) and washed with saturated NaHCO₃ (aq.), water, and brine then dried over Na₂SO₄. 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 (H₂O - 50 % CH₃CN/H₂O) to give the pure product 15h (30 mg, 83 %) as a purple solid. ¹H NMR (500 MHz, CD₃OD) 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); ¹³C NMR (125 MHz, CD₃OD) 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-Cl)⁺ C₂₀H₃₃N₄O₃ 485.2553; found 485.2557.
$^{13}$C NMR (CD$_3$OD)
Synthesis of rosamine 23

Rosamine 14d (39 mg, 0.05 mmol) was dissolved in 2 mL TFA/CH₂Cl₂ (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 K₂CO₃ (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 CH₂Cl₂ and washed with water, dried over Na₂SO₄ and concentrated under reduced pressure. The residue was purified by flash chromatography (10 % MeOH/CH₂Cl₂, Rf = 0.28) to afford the pure product 23 (31 mg, 76 %) as a purple solid. ¹H NMR (500 MHz, CDCl₃) 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); ¹³C NMR (125 MHz, CDCl₃) 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)⁺ C₃₉H₄₈I₄N₄O₅ 779.2669; found 779.2671.
Synthesis of Rosamine 24

BBr\textsubscript{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\textsubscript{2}Cl\textsubscript{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 \textsuperscript{1}PrOH/CH\textsubscript{2}Cl\textsubscript{2} (3 x 15 mL). The organic layer was washed with water (1 x 20 mL), brine (1 x 20 mL), dried over Na\textsubscript{2}SO\textsubscript{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 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 flash chromatography (5 % to 10 % MeOH/CH\textsubscript{2}Cl\textsubscript{2}) to afford the product 24 (85 mg, 87 %) as a green solid. \( R_f = 0.15 \) (10 % MeOH/CH\textsubscript{2}Cl\textsubscript{2}). \textsuperscript{1}H NMR (500 MHz, CDCl\textsubscript{3}/CD\textsubscript{3}OD 1:1) \( \delta \): 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); \textsuperscript{13}C NMR (125 MHz, CDCl\textsubscript{3}/CD\textsubscript{3}OD 1:1) \( \delta \): 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-\text{Cl})\textsuperscript{+} \text{C}_{29}\text{H}_{31}\text{N}_{2}\text{O}_{3} 455.2335; found 455.2340.
$^1$H NMR (1:1 CDCl$_3$/CD$_3$OD)

$^{13}$C NMR (1:1 CDCl$_3$/CD$_3$OD)
Synthesis of Rosamine 25

Rosamine 24 (137 mg, 0.28 mmol), Cs₂CO₃ (456 mg, 1.4 mmol), Bu₄NI (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 °C for 12 h then diluted with 30 mL CH₂Cl₂, washed with H₂O (3 x 20 mL), brine (1 x 20 mL) and dried over Na₂SO₄. The solvents were removed under reduced pressure and the residue was passed through a short pad of silica gel eluting with 5 % MeOH/CH₂Cl₂ to give the crude product which was used in the next step without further purification. The crude material was dissolved in 10 mL TFA/CH₂Cl₂ (1:1) and stirred at 25 °C for 1 h. The solvents were removed with a N₂ stream. The residue was dissolved in 30 mL CH₂Cl₂, washed with H₂O (2 x 20 mL), brine (1 x 20 mL) and dried over Na₂SO₄. The solvents were removed under reduced pressure. The residue was dissolved in 20 mL MeOH/CH₂Cl₂ (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 (H₂O – 60 % CH₃CN/H₂O) to afford the pure product 25 as a green solid (128 mg, 75 %). ¹H NMR (500 MHz, CDCl₃/CD₃OD 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, 2.5 Hz), 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); ¹³C NMR (125 MHz, CDCl₃/CD₃OD 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-Cl)⁺ C₃₃H₃₅N₂O₇ 571.2444; found 571.2451.
Labeling of Avidin with Rosamine 25

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.
Rosamine 24 (68 mg, 0.15 mmol) was dissolved in 2 mL concentrated H$_2$SO$_4$. 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 iPrOH/CH$_2$Cl$_2$. The organic layer was evaporated to dryness under reduced pressure and the residue was purified by flash chromatography (5 to 15 % MeOH/CH$_2$Cl$_2$) to afford product 26 (32 mg, 35 %) as a purple solid and 27 (39 mg, 42 %) as a green solid.

26: $^1$H NMR (300 MHz, 1:1 CD$_3$OD/CDCl$_3$) $\delta$ 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$_3$OD/CDCl$_3$) $\delta$ 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$_2$O$_9$S$_2$ 615.15; found 615.15.

27: $^1$H NMR (500 MHz, 1:1 CD$_3$OD/CDCl$_3$) $\delta$ 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$_3$OD/CDCl$_3$) $\delta$ 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$_2$O$_6$S 611.09; found 611.02.
$^{1}H$ NMR (1:1 CDCl$_3$/CD$_3$OD)

$^{13}C$ NMR (1:1 CDCl$_3$/CD$_3$OD)
\[ ^{1}H\text{ NMR (1:1 CDCl}_{3}/\text{CD}_{3}\text{OD)} \]

\[ ^{13}\text{C NMR (1:1 CDCl}_{3}/\text{CD}_{3}\text{OD)} \]

![Chemical Structure](image)
$^1$H NMR (500 MHz, CD$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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-Cl)$^+$ C$_{28}$H$_{31}$N$_4$O 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 **H** at 1.4 x 10^{-7} M in a CH2Cl2/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\textsuperscript{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 mitochondria 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 CellR\textsuperscript{®} 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

Synthesized according to the procedure described in Appendix D. Light yellow solid (500 mg, 51%). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$O$_6$ 479.22; found 479.23.

$^1$H NMR (CDCl$_3$)
Tf₂O (1.0 mmol) was added dropwise to a solution of 3,6-diamino-xanthen-9-one (0.2 mmol) in 2 mL CH₂Cl₂ at 25 °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 °C. (If the reaction was incomplete, another portion of Tf₂O and nucleophile was added). The solvents were removed under reduced pressure and the residue was purified by flash chromatography (2 % to 10 % MeOH/CH₂Cl₂) to afford the pure product 30.
Pyronin 30a

Red solid (94 mg, 81%). $R_f = 0.38$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 99.2 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ C$_{28}$H$_{36}$N$_3$O 430.29; found 430.28.

$^{19}$F NMR (CDCl$_3$)
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
Pyronin 30b

Red solid (69 mg, 61 %). $R_f = 0.44$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 99.2 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ C$_{27}$H$_{36}$N$_3$O 418.29; found 418.28.

$^{19}$F NMR (CDCl$_3$)
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
Pyronin 30c

Yellow solid (112 mg, 98 %) [racemic sec-BuNH₂ was used as a nucleophile]. $R_f = 0.36$ (10 % MeOH/CH₂Cl₂). $^1$H NMR (300 MHz, CDCl₃) $\delta$ 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, CDCl₃) $\delta$ 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, CDCl₃) $\delta$ 99.1 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ C₂₇H₃₆N₃O 418.29; found 418.28.

$^{19}$F NMR (CDCl₃)
Pyronin 30d

Yellow solid (101 mg, 80 %). $R_f = 0.35$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 99.2 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ C$_{30}$H$_{40}$N$_3$O$_3$ 490.31; found 490.30.

$^{19}$F NMR (CDCl$_3$)
Pyronin 30e

Green solid (28 mg, 23%). \( R_f = 0.29 \) (10 % MeOH/CH\(_2\)Cl\(_2\)). [1.5 eq Tf\(_2\)O and 3 eq BnSH were used in the reaction and stirred for 3h at 25°C]. \(^1\)H NMR (500 MHz, CDCl\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 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\(_3\)) \( \delta \) 99.3 (s). MS (ESI) m/z calcd for (M-OTf\(^+\)) \( C_{30}H_{33}N_2OS \) 469.23; found 469.23.

\(^{19}\)F NMR (CDCl\(_3\))
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
Pyronin 30f

Red solid (120 mg, 86 %). $R_f = 0.36$ (10 % MeOH/CH$_2$Cl$_2$). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 99.2 (s). MS (ESI) m/z calcd for (M-OTf)$^+$ C$_{32}$H$_{40}$N$_5$O$_5$ 546.30; found 546.31.

$^{19}$F 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
\[ a = 12.2520(19) \, \text{Å}, \quad \alpha = 90^\circ. \]
\[ b = 14.376(2) \, \text{Å}, \quad \beta = 106.676(2)^\circ. \]
\[ c = 16.502(3) \, \text{Å}, \quad \gamma = 90^\circ. \]
Volume 2784.4(7) Å³

Density (calculated) 1.354 Mg/m³
Absorption coefficient 0.175 mm⁻¹
F(000) 1200

Crystal size 0.50 x 0.43 x 0.40 mm³
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 F²
Data / restraints / parameters 4882 / 0 / 354
Goodness-of-fit on F² 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
pH dependence of UV absorbance of 30a under: a acidic; b basic conditions (CH$_2$Cl$_2$, 7.0 x 10$^{-6}$ M). Fluorescence pH dependence of 30a under: c acidic; d basic conditions (CH$_2$Cl$_2$, 7.0 x 10$^{-7}$ M). e Fluorescence intensity of 30a under alternating acidic (TFA) and basic (Et$_3$N) conditions (CH$_2$Cl$_2$, 7.0 x 10$^{-7}$ M).
Figure E3. pH Dependence of Photophysical Properties for 30e

a. pH dependence of UV absorbance of 30e under: a acidic; b basic conditions (CH₂Cl₂, 9.6 x 10⁻⁶ M). Fluorescence pH dependence of 30e under: c acidic; d basic conditions (CH₂Cl₂, 1.4 x 10⁻⁷ M). e UV absorbance of 30e under alternating acidic and basic conditions (CH₂Cl₂, 9.6 x 10⁻⁶ M) [absorbance decreased at λ_max with addition of Et₃N and won’t recover after addition of TFA]. f Fluorescence of 30e under alternating acidic and basic conditions (CH₂Cl₂, 1.4 x 10⁻⁷ M). [Completely lost fluorescence within 10 min after addition of Et₃N and did not recover after addition of TFA]
Figure E4. Photograph of 16 and 30 in CH₂Cl₂ Under UV Irradiation
APPENDIX F
EXPERIMENTAL DATA FOR CHAPTER VI

Synthesis of BODIPY donor 35

\[
\text{I} \quad \text{O} \quad \text{CHO} \quad \begin{array}{c}
\text{N} \\
\text{H}
\end{array} \quad \begin{array}{c}
\text{I} \\
\text{NB} \\
\text{F} \\
\text{N} \\
\text{O} \\
\text{O}
\end{array} \quad \begin{array}{c}
\text{N} \\
\text{NN}
\end{array} \quad \text{SO}_3\text{Na}
\]

(i) TFA, CH\(_2\)Cl\(_2\), 1 h
(ii) chloranil, 0.5 h
(iii) Et\(_3\)N, BF\(_3\)OEt\(_2\)
25 °C, 12 h

\[
\text{N}_3 \quad \text{SO}_3\text{Na}
\]

TBTA, Cu, CuSO\(_4\)5H\(_2\)O
THF/H\(_2\)O, 50 °C, 12 h

36 53 %

35 95 %
A mixture of 2-hydroxy-4-iodobenzaldehyde (501 mg, 2.02 mmol) and K₂CO₃ (332 mg, 2.4 mmol) in 20 mL CH₃CN was stirred at 25 °C for 10 min then a solution of CH≡CCH₂(OCH₂CH₂)₃OTs (685 mg, 2.0 mmol) in 20 mL CH₃CN 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 Na₂SO₄ 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. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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.

**¹H NMR (CDCl₃)**

```
```

**¹³C NMR (CDCl₃)**

```
```
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 CH₂Cl₂ 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 CH₂Cl₂ and Et₃N (0.63 mL, 4.56 mmol) was added followed by BF₃OEt₂ (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 Na₂SO₄ 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, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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)⁺ C₂₈H₃₂BF₂INaN₂O₄ 659.14; found 659.18.
$^1$H NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
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\textsubscript{4}·5H\textsubscript{2}O (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\textsubscript{2}Cl\textsubscript{2}) to afford the product 35 (95 mg, 95 %) as an orange semi-solid. \textsuperscript{1}H NMR (500 MHz, CD\textsubscript{3}OD) \(\delta\) 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); \textsuperscript{13}C NMR (125 MHz, CD\textsubscript{3}OD) \(\delta\) 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; \textsuperscript{19}F NMR (282 MHz, CD\textsubscript{3}OD) \(\delta\) 31.81-31.02 (m); MS (ESI) m/z calcd for (M-Na)\textsuperscript{−} C\textsubscript{31}H\textsubscript{38}BF\textsubscript{2}IN\textsubscript{5}O\textsubscript{5}S 800.16; found 800.16.
\[ \text{\( ^1H \) NMR (CD\textsubscript{3}OD)} \]
\[ \text{\( ^{13}C \) NMR (CD\textsubscript{3}OD)} \]
\[ \text{\( ^{19}F \) NMR (CD\textsubscript{3}OD)} \]
Synthesis of BODIPY acceptors

(1) Syntheses of benzaldehyde derivatives

2-Hydroxy-4-iodobenzaldehyde (1.24 g, 5.0 mmol) and K₂CO₃ (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 H₂O and brine. The organic layer was dried over Na₂SO₄ 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 %). ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) δ 159.5, 155.8, 145.5, 133.0, 126.4, 124.0, 121.6, 113.6, 91.9, 52.5.
(i) Hexaethylene glycol (8.9 g, 30 mmol) was dissolved in 60 mL dichloromethane and Ag$_2$O (7.7 g, 33 mmol) was added in three portions over 10 min. The mixture was cooled to 0 °C and stirred for 10 min. 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$_2$Cl$_2$) 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$_2$CH$_2$O)$_6$CH$_2$CO$_2$Me (2.4 g, 6.8 mmol) and Me$_3$NHCl (67 mg, 0.7 mmol) in 30 mL acetonitrile was cooled to 0 °C and Et$_3$N (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 Na₂SO₄ and concentrated under reduced pressure. The residue was purified by flash chromatograph (2 to 5 % MeOH/CH₂Cl₂) to afford the pure product as yellow oil (2.5 g, 73 %). ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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 K₂CO₃ (774 mg, 5.6 mmol) in 30 mL CH₃CN was stirred at room temperature for 10 min then a solution of TsO(CH₂CH₂O)₆CH₂CO₂Me in 20 mL CH₃CN 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 Na₂SO₄ and concentrated under reduced pressure. The residue was purified by flash chromatograph (2 to 5 % MeOH/CH₂Cl₂) to afford the pure product as yellow oil (2.4 g, 92 %). ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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)⁺ C₂₂H₃₃ILiO₁₀ 591.13; found 591.13.
$^{1}H$ NMR (CDCl$_3$)

$^{13}C$ NMR (CDCl$_3$)
2-Hydroxy-5-iodobenzaldehyde (1.24 g, 5.0 mmol) and K₂CO₃ (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 H₂O and brine. The organic layer was dried over Na₂SO₄ 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, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) δ 188.0, 168.2, 159.7, 143.9, 137.3, 127.0, 114.9, 84.5, 65.5, 52.5.
(2) Syntheses of pyrrole fragments 38-40

(i) A solution of 2-iodophenol (3.3 g, 15.0 mmol) and K₂CO₃ (2.7 g, 19.5 mmol) in 30 mL CH₃CN was stirred at room temperature for 10 min then a solution of TsO(CH₂CH₂O)₃CH₃ (5.3 g, 16.5 mmol) in 20 mL CH₃CN 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 Na₂SO₄ 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 %). ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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 °C and iPrMgCl (2M in THF, 0.70 mL, 1.4 mmol) was added dropwise over 2 min. The solution was stirred at 0 °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 °C and stirred for 20 min. The mixture was extracted with EtOAc. The organic layer was washed with water, brine and dried over Na₂SO₄ 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) Na₂CO₃ (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(PPh₃)₄ (104 mg, 0.1 mmol) in 20 mL PhCH₃ and 4 mL MeOH. The mixture was degassed and filled with N₂. 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 Na₂CO₃(aq), water, brine and dried over Na₂SO₄ then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et₃N/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. ¹H NMR (300 MHz, CDCl₃) 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 °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 °C. The mixture was extracted with EtOAc. The organic layer was washed with water, brine and dried over Na₂SO₄ then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et₃N/Hexanes) silica gel (40 % EtOAc/Hexanes) to afford the product 38b (1.2 g, 87 %) as yellow oil. ¹H NMR (300 MHz, CDCl₃) 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.

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

$^{13}$C NMR (CDCl$_3$)
$^{1}H$ NMR (CDCl$_3$)
$^{1}H$ NMR (CDCl$_3$)

$^{13}C$ NMR (CDCl$_3$)
(i) A solution of 2-iodophenol (3.3 g, 15.0 mmol) and K$_2$CO$_3$ (2.7 g, 19.5 mmol) in 30 mL CH$_3$CN was stirred at room temperature for 10 min then a solution of TsO(CH$_2$CH$_2$O)$_3$CH$_3$ (5.3 g, 16.5 mmol) in 20 mL CH$_3$CN 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$_2$SO$_4$ 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 %). $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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, 2H), 3.51-3.48 (m, 2H), 3.33 (s, 3H); $^{13}$C NMR (75 MHz, CDCl$_3$) $\delta$ 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 $^1$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$_2$SO$_4$ 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$_2$CO$_3$ (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$_2$CO$_3$ (aq), water, brine and dried over Na$_2$SO$_4$ then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et$_3$N/Hexanes) silica gel (50 % EtOAc/Hexanes) to afford the product 41c (1.0 g, 94 %) as yellow oil. $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$SO$_4$ then concentrated under reduced pressure. The residue was purified on neutralized (washed with 5% Et$_3$N/Hexanes) silica gel (40 % EtOAc/Hexanes) to afford the product 38c (912 mg, 82 %) as yellow oil. $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$)
$^{1}H$ NMR (CDCl$_3$)

$^{13}C$ NMR (CDCl$_3$)
(i) 6-Hydroxy-1-tetralone (2.11 g, 13.0 mmol) and K$_2$CO$_3$ were dissolved in 30 mL CH$_3$CN and stirred for 10 min then a solution hexa(ethylene glycol) monomethyl ether tosylate (6.17 g, 13.7 mmol) in 20 mL CH$_3$CN 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$_2$O and brine. The organic layer was dried over Na$_2$SO$_4$ and evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH$_2$Cl$_2$) 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$_2$OH·HCl (2.25 g, 32.4 mmol) and Et$_3$N (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$_2$O and brine. The organic layer was dried over Na$_2$SO$_4$ and evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH$_2$Cl$_2$) to afford the product (4.57 g, 93 %) as yellow oil. $^1$H 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$_3$) $\delta$ 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 $^\circ$C for 10 h, the mixture was cooled to room temperature and diluted with EtOAc then washed with H$_2$O and brine. The organic layer was dried over Na$_2$SO$_4$ and evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH$_2$Cl$_2$) 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$_3$) $\delta$ 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).
$^1$H NMR (CDCl$_3$)
(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 NH₄Cl (aq) at -78 °C. The mixture was diluted with water and extracted with Et₂O (3 times). The organic layer was washed with water and brine, dried over Na₂SO₄ and evaporated under reduced pressure. The residue was passed through a short pad of silica gel and eluted with 50 % CH₂Cl₂/hexanes. The product was used for the next step without further purification.

(2) A solution of the above product in 30 mL PhCH₃ was heated to reflux and stirred for 20 min then cooled to 25 °C. Et₃N (1.7 mL, 12 mmol) and DMAP (122 mg, 1 mmol) were added to the reaction mixture followed by a solution of Boc₂O (2.62 g, 12 mmol) in 20 mL CH₂Cl₂. The mixture was stirred at 25 °C for 12 h then washed with water and brine. The organic layer was dried over Na₂SO₄ 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. ¹H NMR (300 MHz, CDCl₃) 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$_3$) $\delta$ 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(PPh₃)₄ (85 mg, 0.07 mmol) in 20 mL PhCH₃ and 4 mL MeOH was added Na₂CO₃ (2M in water, 4.4 mL). The mixture was degassed and filled with N₂. 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 Na₂CO₃ (aq), water, brine and dried over Na₂SO₄ 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. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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, 68.4, 67.0, 60.3, 58.4, 27.4, 13.9. MS (ESI) m/z calcd for (M+H)+ C₂₇H₃₆NO₉ 518.24; found 518.24.
$^{1}H$ NMR (CDCl$_3$)

$^{13}$C NMR (CDCl$_3$)
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 N₂ 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 Na₂SO₄ and evaporated to dryness under reduced pressure. The residue was purified on neutralized silica gel (washed with 5 % Et₃N/Hexanes) eluted with 60 % EtOAc/Hexanes to afford the pure product 40a (103 mg, 73 %) as brown oil. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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(PPh₃)₄ (162 mg, 0.14 mmol) in 40 mL PhCH₃ and 8 mL MeOH was added Na₂CO₃ (2M in water, 10.8 mL). The mixture was degassed and filled with N₂. 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 Na₂CO₃ (aq), water, brine and dried over Na₂SO₄ 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. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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.
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 N₂ 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 Na₂SO₄ and evaporated to dryness under reduced pressure. The residue was purified on neutralized silica gel (washed with 5 % Et₃N/Hexanes) eluted with 60 % EtOAc/Hexanes to afford the pure product (1.28 g, 81 %) as a gray solid. ¹H NMR (300 MHz, CDCl₃) 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); ¹³C NMR (75 MHz, CDCl₃) 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$_3$)
Compound 44 (320 mg, 1 mmol) and pyrrole 38a (346 mg, 2 mmol) were dissolved in 10 mL CH₂Cl₂ 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 CH₂Cl₂. The solvents were removed under reduced pressure. The residue was dissolved in 20 mL 1,2-dichloroethane and Et₃N (0.42 mL, 3 mmol) was added followed by BF₃OEt₂ (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 Na₂SO₄ 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. \(^1\)H NMR (500 MHz, CDCl₃) 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$_2$IN$_2$O$_5$ 695.10; found 695.11.
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$IN$_2$NaO$_{17}$ 1245.40; found 1245.34.
Blue sticky solid (915 mg, 64%), prepared from benzoaldehyde \textbf{45} (683 mg, 1.17 mmol) and pyrrole \textbf{38c} (715 mg, 2.34 mmol) using similar conditions as described for \textbf{47a}.

$^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$IN$_2$O$_{17}$ 1222.41; found 1222.36.
NMR spectra for the following compounds:

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

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

$^{19}F$ NMR (CDCl$_3$)
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$IN$_2$O$_5$ 694.09; found 694.30.
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 40.65-39.12 (m); MS (MALDI) m/z calcd for M$^+$ C$_{34}$H$_{26}$BF$_2$IN$_2$O$_3$ 686.10; found 686.02.
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 40.11-39.42 (m, 1F), 39.19-38.50 (m, 1F); MS (MALDI) m/z calcd for M$^+$ C$_{36}$H$_{30}$BF$_2$IN$_2$O$_5$ 746.13; found 746.03.
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$_3$) $\delta$ 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, 6H), 3.29 (s, 6H), 2.79-2.76 (m, 4H), 2.57-2.54 (m, 4H); $^{13}$C NMR (125 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 40.80-40.16 (m, 1F), 38.38-37.75 (m, 1F); MS (MALDI) m/z calcd for M$^+$ C$_{72}$H$_{102}$BF$_2$IN$_2$O$_{23}$ 1538.60; found 1538.53.
$\text{NMR (CDCl}_3\text{)}$

$\text{1H NMR (CDCl}_3\text{)}$

$\text{13C NMR (CDCl}_3\text{)}$

$\text{19F NMR (CDCl}_3\text{)}$
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 29.09-28.65 (m); MS (MALDI) m/z calcd for M$^+$ C$_{60}$H$_{74}$BF$_2$IN$_2$O$_9$ 1302.40; found 1302.16.
$^{1}H\text{ NMR (CDCl}_3\text{)}$

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

$^{19}F\text{ NMR (CDCl}_3\text{)}$
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$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 28.75-28.48 (m); MS (MALDI) m/z calcd for (M-F)$^+$ C$_{60}$H$_{74}$BF$_2$IN$_2$O$_{19}$ 1283.40, found 1283.30; calcd for M$^+$ C$_{60}$H$_{74}$BF$_2$IN$_2$O$_{19}$ 1302.40, found 1302.29; calcd for (M+Na)$^+$ C$_{60}$H$_{74}$BF$_2$INaN$_2$O$_{19}$ 1325.39, found 1325.28.
$^1$H NMR (CDCl$_3$)

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

$^{19}$F NMR (CDCl$_3$)
(4) Syntheses of BODIPYs 48

Typical procedures:

BODIPY 47a (208 mg, 0.3 mmol), trimethylsilylacetylene (0.4 mL, 3.0 mmol), PdCl₂(PPh₃)₂ (11 mg, 0.02 mmol), CuI (3 mg, 0.02 mmol) were dissolved in 20 mL CH₂Cl₂ then Et₃N (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, CDCl₃) 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); ¹³C NMR (125 MHz, CDCl₃) 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⁺ C₃₇H₃₅BF₂N₂O₅Si 664.24; found 664.31.
Purple semi-solid (271 mg, 86 %), prepared from 47b (322 mg, 0.26 mmol). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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\(^+\)) \(^{C_{61}H_{83}BF_2N_2O_17Si}\) 1215.54; found 1215.47.
$^{1}H$ NMR (CDCl$_3$)

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

$^{19}F$ NMR (CDCl$_3$)
Blue semi-solid (819 mg, *quant*.), prepared from 47c (840 mg, 0.69 mmol). \(^1^H\) NMR (500 MHz, CDCl\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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\(_3\)) \(\delta\) 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\(^+\) \(C_61H_83BFN_2O_{17}Si\) 1173.55; found 1173.55.
$^{1}$H NMR (CDCl$_3$)

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

$^{19}$F NMR (CDCl$_3$)
Red solid (194 mg, 83 %), prepared from 47d (245 mg, 0.35 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$N$_2$O$_5$Si 664.24; found 664.12.
Blue solid (174 mg, 88 %), prepared from 47e (206 mg, 0.3 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 40.67-39.08 (m); MS (MALDI) m/z calcd for M$^+$ C$_{39}$H$_{35}$BF$_2$N$_2$O$_3$Si 656.25; found 656.16.
$^{1}H$ NMR ($CDCl_{3}$)

$^{13}C$ NMR ($CDCl_{3}$)

$^{19}F$ NMR ($CDCl_{3}$)
Blue solid (124 mg, 87 %), prepared from 47f (149 mg, 0.2 mmol).  $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$N$_2$O$_5$Si 716.27; found 716.13.
$^{1}H$ NMR (CDCl$_3$)

$^{13}C$ NMR (CDCl$_3$)
Green sticky solid (566 mg, 80 %), prepared from 47f (724 mg, 0.47 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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, 6H), 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$_3$) $\delta$ 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$_3$) $\delta$ 40.84-40.19 (m, 1F), 38.31-37.65 (m, 1F); MS (MALDI) m/z calcd for M$^+$ C$_{77}$H$_{111}$BF$_2$N$_2$O$_{23}$Si 1508.74; found 1508.68.
Green semi-solid (443 mg, \textit{quant.}), prepared from 47h (458 mg, 0.35 mmol). \textsuperscript{1}H NMR (300 MHz, CDCl\textsubscript{3}) \(\delta\) 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); \textsuperscript{13}C NMR (75 MHz, CDCl\textsubscript{3}) \(\delta\) 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); \textsuperscript{19}F NMR (282 MHz, CDCl\textsubscript{3}) \(\delta\) 29.01-28.63 (m); MS (MALDI) m/z calcd for (M-F)\textsuperscript{+} \(C_{65}H_{83}BF_{2}N_{2}O_{19}Si\) 1253.54, found 1253.49; calcd for (M+Na)\textsuperscript{+} \(C_{65}H_{83}BF_{2}NaN_{2}O_{19}Si\) 1295.53, found 1295.47.
Green semi-solid (1.08 g, 94 %), prepared from 47i (1.18 g, 0.91 mmol). \( ^1H \) NMR (300 MHz, CDCl\(_3\)) \( \delta \) 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); \( ^13C \) NMR (75 MHz, CDCl\(_3\)) \( \delta \) 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); \( ^19F \) NMR (282 MHz, CDCl\(_3\)) \( \delta \) 28.66-28.39 (m); MS (MALDI) m/z calc'd for (M-F)\(^+\) C\(_{65}H_{83}BFN_2O_{18}Si\) 1253.54, found 1253.58; calc'd for M\(^+\) C\(_{65}H_{83}BF_2N_2O_{18}Si\) 1272.54, found 1272.58; calc'd for (M+Na)\(^+\) C\(_{65}H_{83}BF_2NaN_2O_{18}Si\) 1295.53, found 1295.54.
BODIPY 48a (66 mg, 0.1 mmol) and K$_2$CO$_3$ (69 mg, 0.5 mmol) were dissolved in 5 mL MeOH + 5 mL THF + 1 mL H$_2$O. After stirring at room temperature for 24 h, the reaction mixture was diluted with H$_2$O, acidified with 2 M HCl and extracted with CH$_2$Cl$_2$. The organic layer was washed with H$_2$O, brine and dried over Na$_2$SO$_4$ then evaporated to dryness. The residue was purified by flash chromatography (2 to 5 % MeOH/CH$_2$Cl$_2$) to afford the product 37a (49 mg, 85 %) as red solid. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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-H)$^-$$C_{33}H_{24}BF_2N_2O_5$ 577.17; found 577.21.
Purple semi-solid (191 mg, 76 %), prepared from 48b (271 mg, 0.23 mmol).  
$^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$N$_2$O$_{17}$ 1105.49; found 1105.45.
$^1$H NMR (CDCl$_3$)

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

$^{19}$F NMR (CDCl$_3$)
Blue semi-solid (512 mg, 67 %), prepared from 48c (815 mg, 0.69 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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-H)$^-\text{C}_{57}\text{H}_{72}\text{BF}_2\text{N}_2\text{O}_{17}$ 1105.49; found 1105.50.
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$_2$N$_2$O$_5$ 577.1746; found 577.1743.
Blue solid (36 mg, 80 %), prepared from 48e (52 mg, 0.08 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 40.14-39.64 (m); MS (ESI) m/z calcd for (M-H)$^-$ C$_{35}$H$_{24}$BF$_2$N$_2$O$_3$ 569.18; found 569.17.
Blue solid (63 mg, quant.), prepared from 48f (72 mg, 0.1 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$172.8, 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$_3$) $\delta$ 39.50 (br); MS (ESI) m/z calcd for (M-H)$^-\text{C}_{37}\text{H}_{28}\text{BF}_2\text{N}_2\text{O}_5$ 629.21; found 629.18.
Green sticky solid (315 mg, 80 %), prepared from 48g (416 mg, 0.28 mmol). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 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$_3$) $\delta$ 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$_3$) $\delta$ 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$_2$NaN$_2$O$_{23}$ 1445.68; found 1445.79.
$^{1}H$ NMR (CDCl$_3$)

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

$^{19}F$ NMR (CDCl$_3$)
Syntheses of through-bond energy transfer cassettes

(1) Syntheses of cassettes 31a-g

Typical procedures:

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 CuSO₄·5H₂O (0.3 mg, 0.001 mmol) were dissolved in 5 mL THF + 1 mL H₂O 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 (H₂O to 50 % CH₃CN/H₂O) to afford the product (9 mg, 75 %) as red solid. ¹H NMR (500 MHz, d⁶-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); ¹³C NMR (125 MHz, d⁶-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-2Na)⁻ C₅₂H₄₁B₂F₄N₇O₁₁S₂ m/z 550.62, found 550.58; calcd for (M-2Na-H)⁻ C₅₂H₄₀B₂F₄N₇O₁₁S₂ m/z 366.74, found 366.73.
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) $\delta$ 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) $\delta$ 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-2Na)$^{2-}$ C$_{76}$H$_{89}$B$_2$F$_4$N$_7$O$_{23}$S$_2$ M/Z: 814.78; found 814.71.
$^{1}H$ NMR ($d^6$-DMSO)

$^{13}C$ NMR ($d^6$-DMSO)
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) $\delta$ 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) $\delta$ 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-2Na)$^2$: C$_{76}$H$_{89}$B$_2$F$_4$N$_7$O$_{23}$S$_2$ m/z 814.78; found 814.75.
$^1$H NMR ($d^6$-DMSO)

$^{13}$C NMR ($d^6$-DMSO)
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$_3$OD) $\delta$ 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-2Na)$^2^-$ C$_{52}$H$_{41}$B$_2$F$_4$N$_7$O$_{11}$S$_2$ m/z 550.62, found 550.59; calcd for (M-2Na+H)$^-^C$$_{52}$H$_{42}$B$_2$F$_4$N$_7$O$_{11}$S$_2$ m/z 1102.25, found 1102.23.
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) $\delta$ 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) $\delta$ 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-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)

$^{13}$C NMR ($d^6$-DMSO)
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-2Na)$^{2-}$ C$_{56}$H$_{45}$B$_2$F$_4$N$_7$O$_{11}$S$_2$ m/z 576.64; found 576.62.
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$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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-2Na)$^{2-}$ C$_{92}$H$_{117}$B$_2$F$_4$N$_7$O$_{29}$S$_2$ M/Z: 972.87, found 972.86; calcd for (M-2Na-H)$^{3-}$ C$_{92}$H$_{116}$B$_2$F$_4$N$_7$O$_{29}$S$_2$ M/Z: 648.26, found 648.22.
$^1$H NMR (CDCl$_3$/CD$_3$OD)

$^{13}$C NMR (CDCl$_3$/CD$_3$OD)
(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$_3$N (35 μL, 0.25 mmol) was added. The reaction mixture was stirred at 25 °C for 24 h then concentrated under reduced pressure. The residue was purified by flash chromatography (20 % MeOH/CH$_2$Cl$_2$) 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-Na)⁺ C₅₀H₃₈B₂F₄N₅O₉S 982.25; found 982.25. m/z calcd for (M-Na-H)²⁻ C₅₀H₃₇B₂F₄N₅O₉S 490.62; found 490.62.

1H NMR (d₆-DMSO)

13C NMR (d₆-DMSO)
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$_3$OD) $\delta$ 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$_3$OD) $\delta$ 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).
Table F1. Photophysical properties of cassette 33

<table>
<thead>
<tr>
<th>solvent</th>
<th>$\lambda_{abs\ D}$ (nm)</th>
<th>$\lambda_{abs\ A}$ (nm)</th>
<th>$\lambda_{emiss\ A}$ (nm)</th>
<th>$\log(\varepsilon_{max})$</th>
<th>$\Phi$ of acceptor</th>
<th>E.T.E (%)$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MeOH</td>
<td>503</td>
<td>586</td>
<td>625</td>
<td>4.97</td>
<td>4.78</td>
<td>0.63</td>
</tr>
</tbody>
</table>

$\Phi$ of acceptor: Ex. D$^a$, Ex. A$^b$

$^a$ Rhodamine 6G ($\Phi = 0.94$ in EtOH); and $^b$ Cresyl Violet ($\Phi = 0.54$ 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™ 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™ 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™ 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™ 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™ 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 \lambda_{max} for the dye (A_{max})

(ii) correct for the contribution of the dye to the absorbance at A_{280}:

\[ A_{protein} = A_{280} - A_{max} \times \left( \frac{A_{280\ free\ dye}}{A_{max\ free\ dye}} \right) \]

(iii) Dye/protein ration = \left( \frac{A_{max}}{\varepsilon_{dye}} \right) / \left( \frac{A_{protein}}{\varepsilon_{protein}} \right)

**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}

(\Delta A_{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).
\( \lambda_{\text{max abs}(D)}: 502 \text{ nm}; \ \lambda_{\text{max abs}(A)}: 556 \text{ nm}; \ \lambda_{\text{emiss}(A)}: 604 \text{ nm}; \)

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). $\lambda_{\text{max \, abs}}(\text{D})$: 505 nm; $\lambda_{\text{max \, abs}}(\text{A})$: 558 nm; $\lambda_{\text{emiss}}(\text{A})$: 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; $\lambda_{\text{emiss(A)}}$: 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; $\lambda_{\text{emiss(A)}}$: 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; $\lambda_{\text{emiss(A)}}$: 481 nm; 

biotin/protein/dye ratio = 6.6 : 1.0 : 5.2
In-vitro three protein interactions

Experiment (ii) (the key experiment):
In a fluorescence cuvette, 1-streptavidin (3.0 \times 10^{-7} \text{ M}) and 2-avidin (9.0 \times 10^{-7} \text{ M}) were pretreated with BSA (1.8 \times 10^{-5} \text{ 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 \times 10^{-7} \text{ 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 \text{ after}}).

![Fluorescence spectra for exp. (ii) after FRET (ex. 430 nm)](image)

The following strategies were used to deduce the approximate fluorescence of the reaction mixture before FRET:
(i) Streptavidin (3.0 \times 10^{-7} \text{ M}), avidin (9.0 \times 10^{-7} \text{ M}) and BSA (1.8 \times 10^{-5} \text{ M}) were mixed in 0.1 M phosphate buffer (pH 7.4) for 20 min, then Atto-425-BSA-biotin (2.0 \times 10^{-7} \text{ 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_{\text{baseline-1}}). This gives a baseline for one component (Atto-425-BSA-biotin).
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 x 10^{-7} M), 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) for 20 min. Fluorescence of this solution was taken when excited at 430 nm (F_{430}) and 488 nm (F_{488 \text{ before}}).

(b) Since Atto-425-BSA-biotin is not excited at 488 nm, the fluorescence changes observed between F_{488 \text{ after}} and F_{488 \text{ 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:

\[ F_{\text{baseline-2}} = F_{430} \times \left( F_{488 \text{ after}} / F_{488 \text{ before}} \right) \]

(iii) The approximate fluorescence of the reaction mixture before FRET was then the sum of the two components defined above:

\[ F_{\text{baseline}} = F_{\text{baseline-1}} + F_{\text{baseline-2}} \]
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 ($3.0 \times 10^{-7}$ M) and 2-avidin ($9.0 \times 10^{-7}$ M) were pretreated with BSA ($1.8 \times 10^{-5}$ 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 ($2.0 \times 10^{-7}$ 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 ($3.0 \times 10^{-7}$ M) and 2-avidin ($9.0 \times 10^{-7}$ M) were pretreated with BSA ($1.8 \times 10^{-5}$ M) and biotin ($1.0 \times 10^{-4}$ 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 ($2.0 \times 10^{-7}$ 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 \times 10^{-7} \text{ M}) and 2-avidin (9.0 \times 10^{-7} \text{ M}) were pretreated with BSA (1.8 \times 10^{-5} \text{ M}) in 0.1 \text{ M} phosphate buffer (pH 7.4). The mixture was incubated in the dark for 20 min, and then Atto-425-BSA-biotin (2.0 \times 10^{-7} \text{ 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).

\textbf{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 \( (F_{\text{before, 1-streptavidin, 606}}) \) can be read from the baseline \( (F_{\text{baseline-2}}) \) described above;

(b) The fluorescence intensity of 1-streptavidin after FRET is given by:

\[
F_{\text{after, 1-streptavidin, 606}} = F_{\text{after, 606}} - F_{\text{after, 480}} \times (F_{606}/F_{480})_{\text{Atto-425-BSA-biotin}}
\]

where \( F_{\text{after, 606}} \), \( F_{\text{after, 480}} \) are the fluorescence intensity for the reaction mixture after FRET at 606 nm and 480 nm, respectively; \( (F_{606}/F_{480})_{\text{Atto-425-BSA-biotin}} \) is the ratio of fluorescence intensity at 606 nm and 480 nm for Atto-425-BSA-biotin;

(c) The percentage fluorescence change for 1-streptavidin at 606 nm is given by:

\[
(F_{\text{after, 1-streptavidin, 606}} - F_{\text{before, 1-streptavidin, 606}}) \times 100 / F_{\text{before, 1-streptavidin, 606}}
\]

(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:

\[
F_{\text{before, 2-avidin, 654}} = F_{\text{before, 654}} - F_{\text{before, 1-streptavidin, 606}} \times (F_{654}/F_{606})_{1\text{-streptavidin}}
\]
where $F_{\text{before, 1-streptavidin}, 606}$ is the fluorescence intensity of 1-streptavidin at 606 nm before FRET; $(F_{654}/F_{606})_{1\text{-streptavidin}}$ 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:

$$F_{\text{after, 2-avidin}, 654} = F_{\text{after, 654}} - F_{\text{after, 480}} \times (F_{654}/F_{480})_{\text{Atto-425-BSA-biotin}} - F_{\text{after, 1-streptavidin}, 606} \times (F_{654}/F_{606})_{1\text{-streptavidin}}$$

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 654 nm is calculated from:

$$\frac{(F_{\text{after, 2-avidin}, 654} - F_{\text{before, 2-avidin}, 654}) \times 100}{F_{\text{before, 2-avidin}, 654}}$$

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:

\[ E = 1 - \frac{F_{\text{after}}}{F_{\text{before}}} \]

where \( F_{\text{after}} \) and \( F_{\text{before}} \) are the FRET donor (Atto 425) fluorescence intensity after and before FRET, respectively.

<table>
<thead>
<tr>
<th>experiment (i)</th>
<th>experiment (ii)</th>
<th>experiment (ii)</th>
<th>experiment (iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRET efficiency (%)</td>
<td>15</td>
<td>74</td>
<td>16</td>
</tr>
</tbody>
</table>
(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.

![Fluorescence spectra for exp. (ii), before and after FRET (ex. 430 nm)](image)

**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, I-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 (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 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:

- 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.
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).

**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:

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
Figure F33. Biotin-BSA-Atto425 (FRET Donor) only
Figure F34. Streptavidin-1 (Acceptor 1) only
**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